Rolf Steyer

Probability and Causality.

Conditional and Average Total Effects

Version: September 5, 2019

University of Jena

Preface

What can we do to reduce global warming? How can we prevent another global financial crisis? How to fight AIDS? How can we reduce hunger in the world? These questions ask about causal effects of interventions. Obviously, interventions based on the wrong causal theories and hypotheses will cost the life of many and huge amounts of money that could be spent more appropriately. Even if our daily problems are less dramatic, they are of the same nature. Just think about your own actions that you have to chose in your responsibilities as a student, scientist, teacher, physician, psychologist, politician, or as a parent! Whatever you do has effects, and these effects might be different if you take one action instead of another one. It is these kind of thoughts that makes us believe that there is no other issue in the methodology of empirical sciences that deserves and needs more attention and effort than causality. And because the dependencies we are investigating are of a nondeterministic nature, we need a *probabilistic theory of causality*. In other words, we need to understand *probability* and *causality*.

What this book is about

Empirical causal research involves several inferences and interpretations. Among these are:

(a) statistical inference, that is, the inference from a data sample to parameters characterizing the distributions of random variables,
(b) causal inference, that is, the inference from parameters characterizing the distributions of random variables to causal effects and/or dependencies,
(c) interpretation of the putative cause,
(d) interpretation of the outcome variable,
(e) interpretation of the random experiment considered.

This book does not deal with all these points. We will neither discuss the mathematics of statistical inference nor the content issues of construct validity or external validity (Campbell & Stanley, 1963; T. D. Cook & Campbell, 1979; Shadish, Cook, & Campbell, 2002) involved in points (c) to (e). Instead we will focus on the second point: causal inference, that is, the inference from parameters (such as the conditional expectation values of an outcome variable in two treatment conditions, which, per se, have no causal interpretation) to causal effects. This is what the probabilistic theory of causal effects presented in this book is about. As will be shown, causal effects are also parameters that characterize the joint distributions of the random variables considered in a random experiment. However, their definitions are less obvious than 'ordinary' conditional expectation values and their differences.

Basic idea

In order to get a first impression of what this means, let us briefly formulate the basic idea that can most easily be explained if the putative (or presumed) cause is a treatment variable. Suppose an individual, or in more general terms, an observational unit, could be treated by condition 1 or it could be treated by condition 0, *everything else being invariant*. If there is a difference in the outcome considered (some measure of success of the treatment), then this difference is due to the difference in the two treatment conditions. This conception goes back at least to Mill (1843/1865).

Multiple determinacy

The problem with this first version of the basic idea is that most outcomes are *multiply determined*, that is, they are not only influenced by the treatment variable, but by many other variables as well. In the field of agricultural research, for example, the *yield* (*outcome*) of a *variety* does not only depend on the variety (*treatment*) itself, but it also depends on the quality of the *plot* (*observational unit*), such as the average hours of sunshine on the plot per day, the amount of water reaching the plot, and the number of microbes in the plot, and so on. Although Mill's idea sounds perfect, it is not immediately clear which implications it has for practice, because the number of other causes is often too large for keeping constant all of them. Furthermore, Mill's idea fails to distinguish between potential confounders and intermediate variables. Holding constant all intermediate variables as well — and not only all pretreatment variables — would imply that there is no treatment effect any more, if we assume that all treatment effects have to be transmitted by some intermediate variables.

Because of the problem of multiple determinacy, Mills conception has been complemented by Sir Ronald A. Fisher (1925/1946) and by Jerzy S. Neyman (1923/1990) in the second and third decades of the last century. Simply speaking, emphasizing and propagating the randomized experiment, Fisher replaced the *ceteris paribus* clause ('everything else invariant') by the *ceteris paribus distributionibus* clause: *all other possible causes (the 'pretreatment variables') having the same distribution*. This is what randomized assignment of units to treatment conditions, for example, based on a coin flip, secures.

A metaphor — The invisible man and his shadow

Imagine an invisible man. Although we cannot see him, suppose we know that he is there, because we can see his shadow. Furthermore, suppose we would like to measure his size. Doing that, we have two problems, a theoretical and a practical one. The *theoretical problem* is to define *size*. We have to clarify that we do not mean 'volume' or 'weight', but 'height' — without shoes, and without hat and hair. Unfortunately, actual height varies slightly in the course of a day. Hence, we define *size* to be the expectation (with respect to the uniform distribution over the 24 hours) of the momentary heights. This solves the theoretical problem; now we know what we want to measure.

However, because the man is invisible, we cannot measure his *size* directly — and this is not only because his size slightly varies over the day. The crucial problem is that we can only observe his shadow. And this is the *practical problem*: How to determine his size from his shadow? Sometimes, there is almost no shadow at all, sometimes it is huge. Some

geometrical reflection yields a first simple solution: measuring the shadow when the sun has an angle of 45°. But what if it is winter and the sun does not reach this angle? Now we need more geometrical knowledge, taking into account the actual angle of the sun and the observed length of the shadow. This will yield an exact measure of the *size* of the invisible man at this time of the day as well.

Determining a causal effect we face the same kind of problems. First, we have to define a *causal effect*, and second, we have to find out how to determine it from empirically estimable parameters such as true means, that is, from conditional expectation values. The simple solution — corresponding to the 45° angle of the sun in the metaphor — is the perfect randomized experiment. The sample mean differences we observe in a randomized experiment only randomly deviate from the causal effect (due to random sample variation). In contrast, in quasi-experiments and observational studies, solutions to the practical problem are more sophisticated. They are also more sophisticated than in the metaphor of the invisible man, because it is not only *one* other variable (the angle) that determines the length of the shadow; instead there often are *many* other variables systematically determining the sample means as well as the true means that are estimated by these sample means. This is again the problem of multiple determinacy. Furthermore, an observed effect may even be negative although the true causal effect is positive, and vice versa. And this reversal of effects can be systematic, and not only be due to sampling error.

This book presents a solution to the theoretical and the practical problems mentioned above. Unfortunately, both solutions are not as simple and obvious as in our metaphor. Furthermore, there is not only one single kind of causal effects, even if we restrict ourselves to total causal effects and do not consider direct and indirect effects).

Total individual and average causal effects

To our knowledge, the first pioneer tackling the theoretical *and* the practical problems was Jerzy S. Neyman (1923/1990). While Fisher propagated the design technique of randomization, Neyman introduced the concepts of total individual and average causal effects, thus attempting a first solution to the theoretical problem mentioned above. (Note, however, that he used different terms for these concepts). Developing statistical methods for agricultural research, he assumed that, for each individual plot, there is an intra-individual (i.e., plot-specific) distribution of the outcome variable, say Y , under each treatment. He then defined the *individual causal effect of treatment x compared to treatment x'* to be the difference between the intra-individual (plot-specific) expectation of Y (the "true yield") given treatment ("variety") x and the intra-individual (plot-specific) expectation of Y given treatment ("variety") x'. Once the individual causal effect is defined, the *average treatment effect* of x compared to x' on Y is simply the expectation (true mean) of the corresponding individual (plot-specific) causal effects in the set (population) of observational units (plots). Similarly, several kinds of *conditional effects* can be defined, conditioning, for instance, on covariates, that is, on other causes of Y that cannot be affected by X, such as measures of the *quality of the soil* before treatment, *average hours of sunshine, average hours of rain*, and so on.

Total, direct, and indirect effects

At about the same time as Neyman and Fisher developed their ideas, Sewall Wright (Wright, 1918, 1921, 1923, 1934, 1960a, 1960b) developed his ideas on path analysis and the concepts of total, direct, and indirect effects. While his *total effect* aims at the same idea as the causal total average effect, his *direct* and *indirect effects* were new. Simply speaking, in the context of an experiment or quasi-experiment, a direct effect of the treatment is the effect that is not transmitted through the intermediate variables; it is the conditional effect of the treatment variable holding constant the intermediate variables on one of their values. In contrast, the *indirect effect* is the difference between the total effect and the direct effect.

Fundamental problem of causal inference

Whereas the basic ideas outlined above are relatively simple and straightforward, trying to put them into practice — that is, solving the practical problem mentioned above — is often difficult and needs considerable sophistication. The "fundamental problem of causal inference" (Holland, 1986) is that we cannot expose an observational unit to treatment 1 and, at the same time, to treatment 0. However, this is exactly what is necessary if we want to be sure that 'everything else is invariant', a clause that is also an implicit assumption in the solution proposed by Neyman. Comparing the true yield of treatment 1 to treatment 0 *within the same plot* at the same time and identical conditions is an ideal version of the *ceteris paribus* clause, which unfortunately is rarely accomplishable.

Pre-post designs

If we choose to first observe a unit under 'no treatment' and then observe it again after 'treatment', we may be tempted to interpret the pre-post differences as estimates of the individual causal effects of the treatment given in between. However, this interpretation might be wrong, because the unit may have developed (maturated, learned), may have suffered from critical life events, may have experienced historical change, and so on (see, e. g., Campbell & Stanley, 1963; T. D. Cook & Campbell, 1979; Shadish et al., 2002). Hence, in these *pre-post designs* or synonymously, *within-group designs*, we have to make assumptions on the nature of these possible alternative interpretations of the pre-post comparisons, for example, that they do not hold in the application considered or that they have a certain structure that can be taken into account when making causal inferences based on pre-post comparisons.

Between-group designs

If, instead of making comparisons within a unit, we compare different units to each other in *between-group experiments*, we certainly lose the possibility of estimating the *individual* causal effects. However, what we can hope for is that we are still able to estimate the *causal average total effect* and certain *causal conditional total effects*. But how to estimate the average of the causal individual total effects if, due to the fundamental problem of causal inference, the causal individual total effects are not estimable? Both, between-group experiments and quasi-experiments, have a set of (observational) units, at least two

experimental conditions ('treatment conditions', 'expositions', 'interventions', etc.), and at least one outcome variable ('response', 'criterion', 'dependent variable') Y. In the medical sciences, the units are usually patients. In psychology the observational units are often persons, but it could be persons-in-a-situation, or groups as well. In economics it could be subjects, companies, or countries, for instance. In educational sciences the units might be school classes, schools, communities, districts, or countries. In sociology and the political sciences, the units could be persons, but also communities, countries, and so on. In this book we show how to define and also how to make inferences about the average of the causal individual total effects in such sets (and subsets) of observational units and about causal conditional total effects, conditioning on attributes of the observational units or on pretest scores, for instance.

Scope of the theory

In order to delineate the scope of the theory, consider the following kind of *random experiment*: Draw an observational unit u (e. g., a person) out of a set of units, observe the value z of a (possibly multivariate qualitative or quantitative) potential confounder Z for this unit, assign the unit or observe its assignment to x, one of several experimental conditions, and record the numerical value y of the outcome variable Y. We will use U to denote the random variable representing with its value u the unit drawn. Note that many observations can be made additionally to observing U, Z, X, and Y. Although this single-unit trial is a prototype of the kind of empirical phenomena the theory is dealing with, there are other single-unit trials in which the theory can be applied as well (see ch. 2). In fact, the theory is applicable far beyond the true (i.e., the randomized experiment) and the quasi-experiment. This includes applications in which the putative causes are *not* manipulable. However, in this volume we do not treat the case in which the putative cause is a continuous random variable. Otherwise, the theory has its limitations only if there is no clear time order of the random variables considered as putative causes or outcomes.

True experiments and quasi-experiments

The single-unit trial described above is a random experiment, but not necessarily a randomized experiment. A *randomized experiment* is a special random experiment in which the drawn unit is assigned to one of the treatment conditions *via randomization*, for example, depending on the outcome of a coin flip. (In empirical applications, the single-unit trials are repeated n times, where n denotes the sample size.) Referring to single-unit trials, we can distinguish the *true experiment* from the *quasi-experiment* as follows: In the *true experiment*, there are at least two treatment conditions and the assignment to one of the treatment conditions is randomized, for example, by flipping a coin. In a traditional *randomized experiment*, for instance, the treatment probabilities are chosen to be equal for all units. However, equal treatment probabilities for all units are neither essential for the definition of the true experiment nor for drawing valid causal inferences. We may as well have treatment probabilities depending on the units and/or on another potential confounder (for more details, see, e. g., Rem. 8.31), *as long as these treatment probabilities are fixed or known by the researcher*. Note, however, that in designs, in which different units have different treatment probabilities, standard techniques of data analysis such as t-tests or analysis of variance do not test the correct hypotheses any more.

For between-group designs, the *quasi-experiment* may be defined such that there are at least two treatment conditions; however, in contrast to the true experiment, the treatment probabilities are unknown. Nevertheless, valid causal inferences can be drawn in quasi-experiments *provided that we can rely on certain assumptions* (see ch. 6). In specific applications these assumptions might be wrong. If they are actually wrong, causal inferences can be completely wrong as well.

Who should study this book?

The methodologist

In the first place, we would like to address the *methodologist*, that is, the expert in empirical research methodology, especially in the social, economic, behavioral, cognitive, medical, agricultural, and biological sciences. This book provides answers to some of the most important and fundamental questions of these empirical sciences: What do we mean by terms like 'X affects Y', 'X has an effect on Y', 'X influences Y', 'X leads to Y', and so on used in our informal theories and hypotheses? How can we translate these terms into a precise language (i. e., probability theory) that is compatible with the statistical analysis of empirical data? How to design an empirical study and how to analyze the resulting data if we want to probe our theories and learn from such data about the causal dependencies postulated in our theories and hypotheses? And last but not least: How to evaluate interventions, treatments, or expositions to (possibly detrimental) environments, and learn about how which effects they have for which kind of subjects or observational-units, and under which circumstances?

The statistician

Many statisticians believe that causality is beyond the horizon of their profession. Causality might be a matter of empirical researchers and philosophers, they say, but not their own. They think that it cannot be treated mathematically and therefore a statistician should refrain from causal interpretations. As a consequence, they ignore the issue of causality. This book proves that these beliefs are prejudices. The theory of causal effects, as presented here, is a branch of probability theory, which itself, at least since Kolmogorov (1933/1977), is a part of pure mathematics — although with an enormous potential for applications in many empirical sciences and even beyond. The main purpose of this book is to translate the informal concepts about causal effects shared by many methodologists and applied statisticians into well-defined terms of mathematical probability theory. The principle is not to use any term that itself is not defined in other mathematic terms, and the result is a purely mathematical theory of causal effects. Of course, this will make it harder to read this book for the methodologist and those not yet trained in probability theory. However, the reward is a much deeper understanding of what is essential and a much better grasp of the nature of our theories about the real world.

Of course, undefined terms are still used in this book, but only in the examples, in the interpretations, and in the motivations of the definitions. The theory itself is pure mathematics, just in the same way as Kolmogorov's probability theory presented in 1933, which explicated the mathematical, measure-theoretical structure of probabilistic concepts. Substantive meaning results, for example, if we interpret the core components of

the formal structure in a specific random experiment considered. And this is also true for the theory of causal effects presented in this book.

The empirical scientist

The empirical scientist in the fields mentioned above has at least three good reasons to study this book. The *first* is that some crucial parts of his theories and hypotheses are explicated, at least when it comes to considering a concrete experiment or study. The ambiguity in causal language such as 'X affects Y', 'X has an effect on Y', 'X influences Y', 'X leads to Y' are not necessary any more. Reading this book will make it possible to replace these ambiguous terms by well-understood and well-defined terms, improving the precision of empirical research and theories.

The *second* motivation of the empirical scientist is that even if he knows his own theoretical concepts and hypotheses, he still has to know how to design experiments and studies that enable him to test them empirically.

Third, the standard ways of analyzing data offered in the textbooks of applied statistics and in the available computer programs often do not estimate and test the causal effects and dependencies we refer to in our theories. And this is not only bad for the empirical scientist but also for all those relying on the validity of his inferences and his expertise. Just think about all the harmful consequences of wrong causal theories in various empirical research fields, if they are applied to solving concrete problems!

The experimental scientist

There are two messages for those who do their research with experiments, a good one and a bad one. The good news is that, in a perfect randomized experiment, the causal average total treatment effect is indeed estimated when comparing sample means between two different treatment conditions. The bad news is that *we can not rely on randomized assignment of units to treatment conditions* when it comes to estimating *direct* and *indirect* effects. More specifically, in such an analysis it is usually not sufficient to consider intermediate variables, treatment and outcome variables. Instead we also have to include in our analysis *pre-treatment variables* such as a pre-test of the intermediate variable and a pre-test of the outcome variable and apply adjustment methods, very much in the same way as we have to use these techniques in quasi-experiments. Hence, if we want to study the black box between the treatment and the outcome variables, we have to adopt the techniques of causal modeling that are far beyond traditional comparisons of means and analysis of variance. (For more details see, e. g., ?, ?).

The philosopher of science

Philosophers of science study and teach the methodology of empirical sciences. In that respect, their task is very similar to that of the methodologist, perhaps only more general and less specific for a certain discipline. Therefore, it is not surprising that probabilistic causality has also been tackled by philosophers of science (see, e. g., Cartwright, 1979; Spohn, 1980; Stegmüller, 1983; Suppes, 1970). Compared to these approaches, our emphasis is more on those parts of the theory that have implications for the *design* of empirical studies and the *analysis of data* resulting from such studies.

The students in these fields

We believe that probabilistic causal effects is the most rewarding topic in methodology. Although it is tough to get into it, you will get insights why all this methodology stuff was useful and what it was good for. At least this is what our students say at the end of our curriculum, even if they did not have the choice whether or not to take our course on causal effects.

Research traditions in causal effects

Several research traditions have been contributing to the theory of causal effects in various ways. From the *Neyman-Rubin tradition*, we adopted the idea that it is important to define various causal effects such as individual, conditional, and average total effects, even though we modified and extended these concepts in important aspects. Defining causal effects is important for proving that certain methods of data analysis yield unbiased estimates of these effects if certain assumptions can be made. Are there conditions under which the analysis of change scores (between pre- and post-tests) and repeated-measures analysis of variance yield causal effects? Under which conditions do we estimate and test causal effects in the analysis of covariance? Which are the assumptions under which propensity score methods yield estimates of causal effects? Which are the assumptions under which an instrumental variable analysis estimates a causal effect? All these questions and their answers presuppose that we have a mathematical definition of causal effects. Simply speaking, in our theory we replace Rubin's potential outcome variables by our true-outcome variables, thus allowing for variance in the outcome (or response) variables given treatment and an observational unit. Many important results of the theory, for example, about *strong ignorability* and about *propensity scores* remain unchanged, while other results are new, giving more insights, and open the floor for new research techniques.

From the *Campbellian tradition* (see, e. g., Campbell & Stanley, 1966; T. D. Cook & Campbell, 1979; Shadish et al., 2002) we learned that there are questions and problems beyond the theory causal effects itself that are relevant in empirical causal research, such as: How to generalize beyond the study? What does the treatment variable mean? What is the meaning of the outcome variable? And, perhaps the most general question: Are there alternative explanations for the effect? The vast majority of social scientists (including ourselves) have been educated in this research tradition to some degree. Although this training is still very useful as a general methodology framework, it lacks precision and clarity in a number of issues — and the definition of a causal effect is one of them that remains unnecessarily vague in their ideas dealing with interval validity.

From the *graphical modeling tradition* (see, e. g., Cox & Wermuth, 2004; Pearl, 2009; Spirtes, Glymour, & Scheines, 2000), we learned that conditional independence plays an important role in causal modeling. This research tradition has also been developing techniques to estimate causal effects and to search for causal models if specific assumptions can be made. The fact that randomization in a true experiment in no way guarantees the validity of causal inferences on *direct* effects has been brought up by this research tradition.

Structural equation modeling and *psychometrics* have been teaching us how to use latent variables and structural equation modeling in testing causal hypotheses. Due to a number of statistical programs such as AMOS (Arbuckle, 2006), EQS (Bentler, 1995), lavaan

(Rosseel, 2012), LISREL (Jöreskog & Sörbom, 1996/2001), Mplus (Muthén & Muthén, 1998-2007), OpenMx (OpenMx, 2009), RAMONA (Browne & Mels, 1998), structural equation modeling became extremely popular in the social sciences. Although many users of these programs hope to find causal answers, it should be clearly stated that structural equation modeling — and this is true for all kinds of statistical models (including analysis of variance) — does neither automatically estimate and test causal effects, nor does it provide a satisfactory *theory* of causal effects and dependencies. Nevertheless, this research tradition contributes — just like other areas of statistics — a number of statistical techniques that can be very useful in causal modeling.

In this book, we also aim at embedding — and, where necessary, extending — conventional statistical procedures such as analysis of covariance, nonorthogonal analysis of variance, and latent variable modeling, but also more recent techniques based on propensity scores into a coherent theory of probabilistic causality.

How to use this book

This book is written such that standard mathematical probability theory is sufficient for a complete understanding, provided one takes the time that these topics require. In many parts, this is not a book one can just *read;* instead it is a book to be *studied*. This includes working on the questions and exercises provided in each chapter. We presume that the reader is familiar with — or learns while studying this book — the essentials of probability theory, including conditional expectations, conditional independence, and conditional distributions. These essentials of probability theory are dealt with in Steyer and Nagel (2017). That book is also referred to very often for definitions, theorems, and other propositions used in this text. These references are abbreviated by SN-Definition, SN-Theorem, SN-Remark, or SN-(10.32), the latter referring to an equation or a proposition in that book.

We devoted this book almost entirely to the *theory* of causal effects. We also developed the PC program *Causal Effects Explorer* (Nagengast, Kröhne, Bauer, & Steyer, 2007) that can be used for exploring prima facie effects, conditional and average total effects given certain parameters. We believe that this program is useful for teaching and learning the fundamentals of the theory. Furthermore, the program *EffectLiteR* (Mayer, Dietzfelbinger, Rosseel, & Steyer, 2016), can be used to estimate conditional and average total effects from empirical data in experiments and quasi-experiments. Both programs, which are available at *www.causal-effects.de*, may be used together with this book in a course on causal modeling. In fact, this is the content of our workshops on the analysis of causal conditional and average total effects, which are available both as videos-on-demand on the internet and on DVDs, again at *www.causal-effects.de*.

Acknowledgements

This book has been written with the help of several colleagues and students. Werner Nagel (FSU Jena) helped whenever we felt lost in probability spaces. Stephen G. West (Arizona State University) and Felix Thömmes (Cornell University) made detailed suggestions for improving readability of the book. Safir Yousfi and Sonja Hahn contributed and/or suggested concrete ideas, Sonja being extremely helpful also in checking some of the mathematics. Our students Franz Classe, Lisa Dietzfelbinger, Niclas Heider, Marc Heigener, Remo Kamm, Lawrence Lo, David Meder, Marita Menzel, Yuka Morikawa, Sebastian Nit-

sche, Fabian Schäfer, Michael Temmerman, Sebastian Weirich, Anna Zimmermann, and other students critically commented on previous versions, helped minimizing errors, or organizing the references. Uwe Altmann, Linda Gräfe, Sven Hartenstein, Ulf Kröhne, Axel Mayer, Marc Müller, Christof Nachtigall, Benjamin Nagengast, Andreas Neudecker, Ivailo Partchev, Jan Plötner, Steffi Pohl, Norman Rose, Marie-Ann Sengewald, and Andreas Wolf together with the others mentioned above provided the intellectual climate in which this book could be written. We are also grateful to the students and colleagues participating at our courses on the analysis of causal effects asking questions and making important comments. Over the years, this helped a lot to improve this book.

Jena, June 6, 2018

Contents

Part I Introduction

1 Introductory Examples 3
1.1 Example 1 — Joe and Ann With Self-Selection 3
1.1.1 Joint probabilities $P(X=x, Y=y)$ 6
1.1.2 Marginal probabilities $P(X=x)$ and $P(Y=y)$ 7
1.1.3 Prima facie effect 8
1.1.4 Individual total effects 8
1.1.5 Prima facie effect versus expectation of the individual total effects .. 11
1.1.6 How to evaluate the treatment? 12
1.2 Example 2 — Nonorthogonal Two-Factorial Experiment 13
1.2.1 Prima facie effects 14
1.2.2 Prima facie effects controlling for the qualitative covariate Z 15
1.2.3 Individual effects 16
1.2.4 Average of the individual effects 16
1.2.5 Average of the $(Z=z)$-conditional total effects 17
1.2.6 How to evaluate the treatment? 18
1.3 Summary and Conclusion 18
1.4 Exercises 21

2 Some Typical Kinds of Random Experiments 27
2.1 Simple Experiments 28
2.2 Experiments With Fallible Covariates 32
2.3 Two-Factorial Experiments 34
2.4 Multilevel Experiments 36
2.5 Experiments With Latent Outcome Variables 38
2.6 Summary and Conclusion 39
2.7 Exercises 40

3 Probability and Conditional Expectation 43
3.1 Probability Space 43
3.1.1 Set of Possible Outcomes 44
3.1.2 Set of Possible Events 45
3.1.3 Probability Measure 46
3.1.4 Conditional Probability 48
3.1.5 Conditional Probability Measure 49
3.2 Random Variable 52
3.2.1 σ-Algebra Generated by a Random Variable 55

3.2.2 Distribution of a Random Variable . . . 56
3.2.3 Expectation of a Numerical Random Variable . . . 58
3.2.4 Variance, Covariance, and Correlation . . . 61
3.3 Conditional Expectation and Related Concepts . . . 63
3.3.1 Conditional Expectation $E(Y|X)$ . . . 63
3.3.2 First Examples . . . 65
3.3.3 Regression and Factorization of a Conditional Expectation . . . 66
3.3.4 Conditional Expectation Value $E(Y|X=x)$ . . . 66
3.3.5 Partial Conditional Expectation $E(Y|X=x,Z)$ . . . 68
3.3.6 Conditional Expectation $E^{X=x}(Y|Z)$ With Respect to $P^{X=x}$ . . . 69
3.4 Summary and Conclusions . . . 73
3.5 Proofs . . . 74
3.6 Exercises . . . 74

Part II Basic Concepts

4 Potential Confounder and True Outcome Variable . . . 81
4.1 Potential Confounder and Global Potential Confounder . . . 81
4.2 True Outcome Variable . . . 88
4.3 Summary and Conclusions . . . 91
4.4 Exercises . . . 92

5 Causal Total Effects . . . 95
5.1 True Total Effect Variable . . . 96
5.2 Causal Average Total Effect . . . 98
5.3 Conditional Total Effect and Conditional Total Effect Function . . . 100
5.3.1 Notation, Assumptions and Definitions . . . 101
5.3.2 Causal $(X=x^*)$-Conditional Total Effect . . . 103
5.3.3 Reaggregation . . . 105
5.4 Example: Joe and Ann With Bias at the Individual Level . . . 107
5.4.1 The Random Experiment . . . 107
5.4.2 Choosing the Focused Treatment Variable . . . 107
5.4.3 True Outcome Variables and True Total Effects . . . 108
5.4.4 Causal Average Total Effect . . . 109
5.4.5 $(U=u)$-Conditional Prima Facie Effects . . . 109
5.4.6 Causal $(U=u)$-Conditional Total Effects . . . 111
5.4.7 Causal $(Z=z)$-Conditional Total Effects . . . 112
5.4.8 Causal $(X=x)$-Conditional Total Effects . . . 112
5.4.9 $(X=x^*,V=v)$-Conditional Total Effects . . . 113
5.4.10 Causal $(X=x,Z=z)$-Conditional Total Effects . . . 114
5.5 Summary and Conclusions . . . 115
5.6 Proofs . . . 118
5.7 Exercises . . . 118

6 Unbiasedness and Identification of Causal Effects 123
6.1 Unbiasedness of $E(Y|X)$ and its Values $E(Y|X=x)$ 124
6.2 Unbiasedness of $E(Y|X,Z)$ and its Values $E(Y|X=x,Z=z)$ 126
6.3 Unbiasedness of Prima Facie Effects 129
6.4 Identification of Causal Total Effects 131
6.4.1 Identification of the Causal Average Total Effect 131
6.4.2 Identification of a Causal Conditional Total Effect Function 132
6.4.3 Identification of a Causal Conditional Total Effect 134
6.5 Three Examples 135
6.5.1 Commonalities of all Examples 135
6.5.2 Description of the Examples 137
6.5.3 $(X=x)$-Conditional Expectation Values 139
6.5.4 $(X=x,Z=z)$-Conditional Expectation Values 141
6.5.5 Conditional Total Effects 142
6.5.6 Computing the Causal Average Total Effect From Conditional Total Effects 142
6.5.7 First Conclusions 142
6.6 Example With Accidental Unbiasedness 143
6.7 Summary and Conclusions 146
6.8 Proofs 146
6.9 Exercises 151

7 Rosenbaum-Rubin Conditions 155
7.1 Independence and Conditional Independence of Random Variables 155
7.2 Mean-Independence of τ_x From X 157
7.3 Conditional Mean-Independence of τ_x From X 159
7.4 Z-Conditional Independence of τ and X (Strong Ignorability) 165
7.5 Summary and Conclusions 173
7.6 Proofs 174
7.7 Exercises 176

8 Fisher Conditions 179
8.1 Fisher Conditions 179
8.1.1 Independence of C_X and X 180
8.1.2 Z-Conditional Independence of C_X and X 180
8.2 Implications of the Fisher Conditions 182
8.2.1 Implications of Independence of C_X and X 182
8.2.2 Unbiasedness of the Prima Facie Effects 184
8.2.3 Implications of Z-Conditional Independence of C_X and X 184
8.2.4 Unbiasedness of a Z-Conditional Prima Facie Effect Function 186
8.3 Examples 187
8.4 Methodological Conclusions 189
8.5 Summary and Conclusions 193
8.6 Proofs 195
8.7 Exercises 197

9 Reichenbach-Suppes Conditions 201
9.1 Reichenbach-Suppes Conditions 201
9.1.1 X-Conditional Independence and Mean-Independence of Y and C_X 201
9.1.2 (X,Z)-Conditional Independence of Y and C_X 203
9.2 Implication Structure Among the Reichenbach-Suppes Conditions 204
9.3 Further Implications of the Reichenbach-Suppes Conditions 206
9.3.1 Implications of X-Conditional Mean-Independence of Y from C_X ..206
9.3.2 Implications of (X,Z)-Conditional Mean-Independence of Y from C_X208
9.4 Example 210
9.5 Methodological Conclusions 212
9.5.1 Methodological Conclusions From $Y \vdash C_X | X$ 212
9.5.2 Methodological Conclusions From $Y \vdash C_X | (X,Z)$ 213
9.6 Summary and Conclusions 214
9.7 Proofs 216
9.8 Exercises 218

References 221

List of Figures

1.1 Probability of success given treatment conditions 7
1.2 Conditional probabilities of success given treatment and person 10
1.3 Conditional probabilities of success given treatment and person 12
1.4 Conditional expectation values of Y given treatment and status 17

2.1 A simple experiment or quasi-experiment 29
2.2 Experiment or quasi-experiment with a fallible covariate 33

5.1 Conditional probabilities of success given treatment and person 100

List of Tables

1.1 Joe and Ann with self-selection – compressed table 4
1.2 Joe and Ann with self-selection – explicit table 5
1.3 Joint and marginal probabilities of treatment and success 6
1.4 Joint and marginal probabilities of all three observables 9
1.5 Random experiment of a two-factorial nonorthogonal design 13
1.6 Conditional expectation values of the outcome variable given treatments 14
1.7 Conditional expectation values $E(Y|X=x,Z=z)$ given treatment and status . . 15

3.1 Joe and Ann with randomized assignment 44
3.2 No treatment for Joe 52

4.1 Joe with two independent treatments 87

5.1 Joe and Ann with bias at the individual level 108

6.1 Self-selection of the unit to a treatment 138
6.2 Randomized assignment of the unit to a treatment 139
6.3 Conditionally randomized assignment of the unit to a treatment 140
6.4 Accidental unbiasedness 144

7.1 Z-conditional mean-independence of τ_x from $1_{X=x}$ 160
7.2 Conditional probabilities $P(U=u|X=x,Z=z)$ for Table 7.1 161
7.3 Z-conditional mean-independence of τ from X 164
7.4 Strong ignorability 166
7.5 Implication relations among causality conditions for $E(Y|X=x)$ and $E(Y|X)$. 173
7.6 Implication relations among causality conditions for $E^{X=x}(Y|Z)$ and $E(Y|X,Z)$ 174

8.1 Implications among causality conditions 195

9.1 (X,Z)-conditional mean-independence of Y on C_X 211
9.2 Implications among Reichenbach-Suppes conditions 215

Part I
Introduction

Chapter 1
Introductory Examples

For more than a century there have been examples in the statistical literature showing that comparing means or comparing probabilities (e. g., of success of a treatment) between a group exposed to a treatment and a comparison group (unexposed or exposed to a different treatment) does not necessarily answer our questions: 'Which treatment is better overall?' or 'Which treatment is better for which kind of person?' Differences between true means and differences between probabilities (or any other comparison between probabilities such as odds ratios, log odds ratios, or relative risk) are usually not the treatment effects we are looking for (see, e. g., Pearson, Lee, & Bramley-Moore, 1899; Yule, 1903; Simpson, 1951). They are just *effects at first sight* or "prima facie effects" (Holland, 1986).

Just like the shadow in the metaphor of the invisible man (see the preface), prima facie effects reflect the effects of the treatment (the size of the invisible man), but also the effects of other causes (the angle of the sun). The goal of analyzing *causal* effects is to estimate the effect of the treatment alone, isolating it from other potential influences, for example, from the effects of sex, educational background, socio-economic status, and so on. The general idea is to define and, in applications, estimate a treatment effect that is not biased by preexisting differences between treatment groups that would also be observed after treatment if there were no treatment effect at all.

Overview

We illustrate systematic bias in determining *total* (as opposed to direct or indirect) treatment effects in quasi-experiments by two examples. The first one deals with a dichotomous outcome variable, the second with a quantitative one. Note that the problems described in these two examples cannot occur in a randomized experiment, but they are ubiquitous in nonrandomized quasi-experimental studies.

1.1 Example 1 — Joe and Ann With Self-Selection

In this example, the prima facie effect reverses if we switch from comparing the conditional probabilities of success between treatment and control, that is, from comparing

$$P(Y=1\,|\,X=1) \quad \text{to} \quad P(Y=1\,|\,X=0)$$

to comparing the corresponding probabilities additionally conditioning on the person variable U with values $u = \mathit{Joe}, \mathit{Ann}$, that is, to comparing

$$P(Y=1\,|\,U=u, X=1) \quad \text{to} \quad P(Y=1\,|\,U=u, X=0).$$

Table 1.1. Joe and Ann with self-selection – compressed table

U	$P(U=u)$	$P(X=1\,\vert\,U=u)$	$P(Y=1\vert U=u,X=0)$	$P(Y=1\vert U=u,X=1)$
Joe	.5	.04	.7	.8
Ann	.5	.76	.2	.4

This kind of phenomenon, which is already known at least since Yule (1903), is called *Simpson's paradox* (Simpson, 1951), and it is still being debated (see, e. g., Hernán, Clayton, & Keiding, 2011). RST Neuere Literatur?

Table 1.1 shows the compressed table of a random experiment that is composed of three parts.

(1) A person is sampled from a set of two persons, Joe and Ann, with identical probabilities for each person u, that is, with probability $P(U=u) = .5$.

(2) If Joe is sampled, then he obtains treatment ($X=1$) with probability $P(X=1\,|\,U=Joe) = .04$. In contrast, if Ann is sampled, then she obtains treatment with probability $P(X=1\,|\,U=Ann) = .76$. (These numbers may reflect self-selection to treatment and the different inclinations of the two persons to go to treatment.)

(3) If Joe is sampled and *not treated*, then his probability $P(Y=1\,|\,U=Joe, X=0)$ of success is .7. If he is sampled and *treated*, then his probability $P(Y=1\,|\,U=Joe, X=1)$ of success is .8. In contrast, if Ann is sampled and *not treated*, then her probability $P(Y=1\,|\,U=Ann, X=0)$ of success is .2, and if she is sampled and *treated*, then her probability $P(Y=1\,|\,U=Ann, X=1)$ of success is .4.

This table describes a random experiment and it contains all information we need to compute the (causal) total effects of the treatment on the outcome variable Y (success), including the (causal) conditional total effects given the person and the (causal) average total effect of the treatment.

Note that Table 1.1 does not describe a *randomized experiment*, in which, by definition, the treatment probabilities $P(X=1\,|\,U=u)$ would be identical for all observational units u. Instead, it describes a *random experiment*, which is that kind of empirical phenomenon that we usually consider when we apply probability theory using terms such as random variables, their expectations, variances, distribution, correlations, etc. In inferential statistics it is those concepts about which we formulate our hypothesis and that we try to estimate in a sample.

In probability theory, we consider such a random experiment from the *pre facto perspective*. Hence, we do not consider data that would result from actually conducting such a random experiment. Data are only important in order to learn about the laws of a random experiment from observations. Data analysis is only a way to learn about these laws. But it these laws of the random experiment that are of primary interest. More precisely, if we know the eight probabilities displayed in Table 1.1, then we have all the information that we need to compute the causal conditional and average total effects of the treatment on the outcome (success). All it needs is to define these concepts in terms of probability theory, and this is what this book is about.

Table 1.2. Joe and Ann with self-selection – explicit table

Possible outcomes ω_i		Observables			Conditional probabilities		
Unit Treatment Success	$P(\{\omega_i\})$	Person variable U	Treatment variable X	Response variable Y	$P(X=1\,\vert\,U)$	$P(Y=1\,\vert\,X)$	$P(Y=1\,\vert\,X,U)$
$\omega_1 = (Joe, no, -)$	.144	*Joe*	0	0	.04	.6	.7
$\omega_2 = (Joe, no, +)$	.336	*Joe*	0	1	.04	.6	.7
$\omega_3 = (Joe, yes, -)$	.004	*Joe*	1	0	.04	.42	.8
$\omega_4 = (Joe, yes, +)$	.016	*Joe*	1	1	.04	.42	.8
$\omega_5 = (Ann, no, -)$	.096	*Ann*	0	0	.76	.6	.2
$\omega_6 = (Ann, no, +)$	.024	*Ann*	0	1	.76	.6	.2
$\omega_7 = (Ann, yes, -)$	.228	*Ann*	1	0	.76	.42	.4
$\omega_8 = (Ann, yes, +)$	.152	*Ann*	1	1	.76	.42	.4

Table 1.2 describes the same random experiment as Table 1.1, but in different way. The eight triples such as (*Joe*, *no*, −) or (*Ann*, *yes*, +) represent one of the eight possible outcomes $\omega_1, \ldots, \omega_8$ of the random experiment that are gathered in the set Ω of possible outcomes. Remember, an event A is a subset of Ω that has a probability $P(A)$, which is assigned by the probability measure P to each element A in the set $\mathscr{A}$ of all events (see SN-ch. 4 for these elementary concepts of probability theory). The eight probabilities of the elementary events contain the same information as the eight probabilities in Table 1.1. All conditional probabilities appearing in Table 1.2 and all probabilities and all conditional probabilities presented in Table 1.1 can be computed from these eight probabilities of the elementary events.

Table 1.2 has the virtue of explicitly showing all possible outcomes of the random experiment considered. Furthermore, it shows how the random variables U, X, and Y are defined, showing the assignments of their values to each of the eight possible outcomes of the random experiment (see SN-ch. 5 for the definition of a random variable). It also displays the conditional probabilities $P(Y=1\,|\,X,U)$, $P(Y=1\,|\,X)$, and $P(X=1\,|\,U)$, which are random variables on the same probability space as the observables, that is, the random variables U, X, and Y (see Def. 3.52 and Rem. 3.58 for the definition of such a conditional probability).

The crucial point is that each of the conditional probabilities mentioned above also assigns a value to each of the eight possible outcomes of the random experiment. For example, the values assigned by $P(X=1\,|\,U)$ to each outcome $\omega_i \in \Omega$ are the conditional probabilities $P(X=1\,|\,U=u)$. More precisely,

$$P(X=1\,|\,U)(\omega_i) \;=\; P(X=1\,|\,U=u), \quad \text{if } \omega_i \in \{\omega \in \Omega\colon U(\omega) = u\}. \tag{1.1}$$

Similarly,

Table 1.3. Joint and marginal probabilities of treatment and success

	Treatment		
Success	No ($X=0$)	Yes ($X=1$)	
No ($Y=0$)	.240	.232	.472
Yes ($Y=1$)	.360	.168	.528
	.600	.400	1.000

Note. The entries in the four cells are the joint probabilities $P(X=x, Y=y)$, the other entries are the marginal probabilities $P(X=x)$ (last row) and $P(Y=y)$ (last column).

$$P(Y=1\,|\,X)(\omega_i) \;=\; P(Y=1\,|\,X=x), \quad \text{if } \omega_i \in \{\omega \in \Omega\colon X(\omega) = x\}, \tag{1.2}$$

and

$$P(Y=1\,|\,X,U)(\omega_i) \;=\; P(Y=1\,|\,X=x, U=u), \quad \text{if } \omega_i \in \{\omega \in \Omega\colon X(\omega) = x, U(\omega) = u\} \tag{1.3}$$

(see Table 1.2 in order to check these assignment rules).

1.1.1 Joint probabilities $P(X=x, Y=y)$

Unfortunately, in realistic applications, we cannot estimate the probabilities of all elementary events displayed in Table 1.2 because, usually, we cannot repeat this random experiment. Once, a person is treated, we cannot undo treatment and repeat this process. This has been called the "fundamental problem of causal inference" (Holland, 1986). Oftentimes a treatment is irreversible and the time between treatment and the assessment of the outcome variable Y can take months and even years. Nevertheless, it is meaningful to consider the random experiment described by Table 1.2 including the person-specific probabilities of treatment and the probabilities of success given treatment and person. As will be shown in chapters 4 and 5, we even *have to* consider this kind of random experiment when we want to *define* causal effects.

However, what can be estimated in empirical applications are the joint probabilities such as $P(X=x, Y=y)$ and the conditional probabilities such as $P(Y=1\,|\,X=1)$ and $P(Y=1\,|\,X=0)$, that is, the conditional probabilities of success given treatment and given control, respectively. In order to estimate $P(X=x, Y=y)$ we only have to observe the relative frequencies of the joint occurrence of treatment x and outcome y. In a simulation, we can easily repeat this random experiment n times in order to generate a data sample of size n (see Exercise 1-7). In contrast, in an empirical application, estimating $P(X=x, Y=y)$ requires to sample from a very large set of persons (observational units), and not just from the set $\{Joe, Ann\}$ of two persons. (See Splawa-Neyman, 1923/1990 for an early sampling model dealing with the problem of non-replacement).

Table 1.3 shows the joint probabilities $P(X=x, Y=y)$ of treatment and success, as well as the marginal probabilities $P(X=x)$ and $P(Y=y)$ of treatment and success, respectively. These probabilities are easily computed from the probabilities of the elementary events displayed in the second column of Table 1.2. For example, the probability $P(X=0, Y=1)$

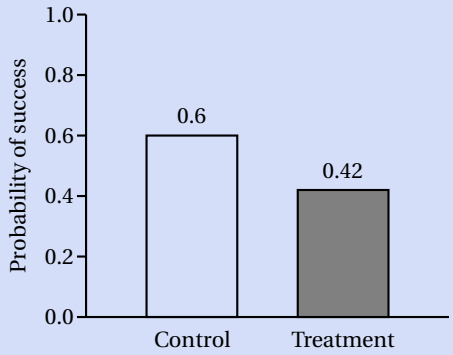

Figure 1.1. Probability of success given treatment conditions

that the sampled person receives no treatment and is successful is the sum of the probabilities of the elementary events $\{\omega_2\} = \{(Joe, no, +)\}$ and $\{\omega_6\} = \{(Ann, no, +)\}$, that is,

$$P(X=0, Y=1) \;=\; P(\{(Joe, no, +)\}) + P(\{(Ann, no, +)\}) \;=\; .336 + .024 \;=\; .36.$$

Similarly, the probability $P(X=1, Y=1)$ that the sampled person receives treatment and is successful is the sum of the probabilities of the two elementary events $\{\omega_4\} = \{(Joe, yes, +)\}$ and $\{\omega_8\} = \{(Ann, yes, +)\}$, that is,

$$P(X=1, Y=1) \;=\; P(\{(Joe, yes, +)\}) + P(\{(Ann, yes, +)\}) \;=\; .016 + .152 \;=\; .168.$$

Table 1.3 is the theoretical analog to a contingency table that would be observed in a data sample. More precisely, if we multiply the displayed numbers by the sample size, then we receive the *expected frequencies* of the corresponding events. For example, if the sample size is 1000, then we expect 240 cases in cell $(X=0, Y=0)$ and 360 cases in cell $(X=0, Y=1)$, etc. Of course, in a data sample, the *observed frequencies* would fluctuate around these expected frequencies (see Exercise 1-7).

1.1.2 Marginal probabilities $P(X=x)$ and $P(Y=y)$

The marginal probabilities $P(X=x)$ and $P(Y=y)$ are also easily computed from the probabilities of the elementary events displayed in the second column of Table 1.2. For example, the probability $P(X=0)$ that the sampled person receives no treatment is the sum of the probabilities of the four elementary events $\{\omega_1\} = \{(Joe, no, -)\}$, $\{\omega_2\} = \{(Joe, no, +)\}$, $\{\omega_5\} = \{(Ann, no, -)\}$, and $\{\omega_6\} = \{(Ann, no, +)\}$, that is,

$$\begin{aligned} P(X=0) \;&=\; P(\{(Joe, no, -)\}) + P(\{(Joe, no, +)\}) + P(\{(Ann, no, -)\}) + P(\{(Ann, no, +)\}) \\ &=\; .144 + .336 + .096 + .024 \;=\; .6, \end{aligned}$$

and this implies

$$P(X=1) \;=\; 1 - P(X=0) \;=\; .4$$

Similary, the probability $P(Y=1)$ that the sampled person is successful is the sum of the probabilities of the four elementary events $\{\omega_2\} = \{(Joe, no, +)\}$, $\{\omega_4\} = \{(Joe, yes, +)\}$, $\{\omega_6\} = \{(Ann, no, +)\}$, and $\{\omega_8\} = \{(Ann, yes, +)\}$, that is,

$$\begin{aligned} P(Y=1) &= P(\{(Joe, no, +)\}) + P(\{(Joe, yes, +)\}) + P(\{(Ann, no, +)\}) + P(\{(Ann, yes, +)\}) \\ &= .336 + .016 + .024 + .152 = .528, \end{aligned}$$

which implies

$$P(Y=0) = 1 - P(Y=1) = .472.$$

1.1.3 Prima facie effect

Comparing the conditional probability $P(Y=1 \mid X=1)$ of success given the *treatment condition* to the conditional probability $P(Y=1 \mid X=0)$ of success given the *control condition* would lead us to the (wrong) conclusion that the *treatment is harmful*. These two conditional probabilities can be computed by

$$P(Y=1 \mid X=1) = \frac{P(Y=1, X=1)}{P(X=1)} = \frac{.168}{.4} = .42$$

and

$$P(Y=1 \mid X=0) = \frac{P(Y=1, X=0)}{P(X=0)} = \frac{.36}{.6} = .6,$$

respectively (see, e. g., SN-section 4.2). Figure 1.1 displays both conditional probabilities in a bar chart.

These two conditional probabilities can be compared to each other in different ways. The simplest one is looking at the *difference* $P(Y=1 \mid X=1) - P(Y=1 \mid X=0)$. This is a particular case of the difference $E(Y \mid X=1) - E(Y \mid X=0)$ between two conditional expectation values, in which the outcome variable Y is dichotomous with values 0 and 1 (see SN-Remark 9.8). Following Holland (1986), we will call this difference the (unconditional) *prima facie effect* and use the notation

$$PFE_{10} = E(Y \mid X=1) - E(Y \mid X=0) = P(Y=1 \mid X=1) - P(Y=1 \mid X=0).$$

Hence, in this example,

$$PFE_{10} = P(Y=1 \mid X=1) - P(Y=1 \mid X=0) = .42 - .6 = -.18.$$

Other possibilities of comparing the two conditional probabilities are to compute the odds ratio, its logarithm, or the risk ratio (see, e. g., SN-Remarks 13.14 to 13.16 or chapter 4 of Rothman, Greenland, & Lash, 2008, for a detailed discussion of these and other effect parameters). No matter which of these effect parameters we choose, they all lead to the conclusion that the *treatment is harmful* (see Exercise 1-8). As shown in the following section this conclusion is utterly wrong.

1.1.4 Individual total effects

The conclusion about the effect of the treatment is completely different if we look at the treatment effects separately for Joe and Ann. Table 1.4 shows the joint distributions of

Table 1.4. Joint and marginal probabilities of all three observables

Joe ($U=Joe$)

	Treatment		
Success	No ($X=0$)	Yes ($X=1$)	
No ($Y=0$)	.144	.004	.148
Yes ($Y=1$)	.336	.016	.352
	.48	.02	.5

Ann ($U=Ann$)

	Treatment		
Success	No ($X=0$)	Yes ($X=1$)	
No ($Y=0$)	.096	.228	.324
Yes ($Y=1$)	.024	.152	.176
	.12	.38	.5

Note. The entries in the four cells for Joe and in the four cells for Ann are the joint probabilities $P(U=u, X=x, Y=y)$. The other entries are the joint probabilities $P(U=u, X=x)$ (third and last row) and $P(U=u, Y=y)$ (last column), respectively, and the two marginal probabilities $P(U=u)$.

treatment, success, and the person variable U with values Joe and Ann. The probabilities of sampling Joe and of sampling Ann are identical, that is, $P(U=Joe) = P(U=Ann) = .5$. Furthermore, the joint probabilities $P(U=u, X=x, Y=y)$ are the probabilities of the elementary events displayed in the second column of Table 1.2. These joint probabilities are displayed again in a form analog to $(2\times2\times2)$-contingency table in Table 1.4.

As already mentioned in section 1.1.1, in empirical applications, this random experiment cannot be repeated in order to obtain a data sample. However, we can repeat it in s simulation (see Exercise 1-7). If, in such a simulation, we multiply the numbers displayed in Table 1.4 by the sample size, then we receive the *expected frequencies* of the corresponding events. For example, if the sample size is 1000, then we expect 144 cases in cell $(U=Joe, X=0, Y=0)$ and 336 cases in cell $(U=Joe, X=0, Y=1)$, etc. Of course, in data samples, the *observed frequencies* fluctuate around these expected frequencies.

Using the joint probabilities displayed in Table 1.4, the conditional probability of success for Joe in the treatment condition can be computed as follows:

$$P(Y=1 \mid X=1, U=Joe) = \frac{P(U=Joe, X=1, Y=1)}{P(U=Joe, X=1)} = \frac{.016}{.016+.004} = .8$$

(see Exercise 1-9). In contrast, Joe's conditional probability of success in the control condition is

$$P(Y=1 \mid X=0, U=Joe) = \frac{P(U=Joe, X=0, Y=1)}{P(U=Joe, X=0)} = \frac{.336}{.336+.144} = .7.$$

Hence,

$$P(Y=1\,|\,X=1,U=Joe)\;-\;P(Y=1\,|\,X=0,U=Joe)\;=\;.8\;-\;.7\;=\;.1,$$

which may lead us to conclude that *the treatment is beneficial for Joe*. Again, because Y is dichotomous with values 0 and 1, this difference is a special case of the difference

$$E(Y\,|\,X=1,U=Joe)-E(Y\,|\,X=0,U=Joe),$$

which we call the *individual total (treatment) effect* of Joe, using the notation $ITE_{U;10}(Joe)$. Hence,

$$\begin{aligned} ITE_{U;10}(Joe) \;&=\; E(Y\,|\,X=1,U=Joe)\;-\;E(Y\,|\,X=0,U=Joe) \\ &=\; P(Y=1\,|\,X=1,U=Joe)\;-\;P(Y=1\,|\,X=0,U=Joe)\,. \end{aligned} \tag{1.4}$$

What about the individual total effect of Ann? Table 1.4 shows that the conditional probability of success for Ann in the treatment condition is

$$P(Y=1\,|\,X=1,U=Ann)\;=\;\frac{P(U=Ann\,,X=1,Y=1)}{P(U=Ann\,,X=1)}\;=\;\frac{.152}{.152+.228}\;=\;.4,$$

whereas it is

$$P(Y=1\,|\,X=0,U=Ann)\;=\;\frac{P(U=Ann\,,X=0,Y=1)}{P(U=Ann\,,X=0)}\;=\;\frac{.024}{.024+.096}\;=\;.2$$

in the control condition. Figure 1.2 shows these conditional probabilities in a bar chart. Considering the individual total effect

$$\begin{aligned} ITE_{U;10}(Ann) \;&=\; P(Y=1\,|\,X=1,U=Ann)\;-\;P(Y=1\,|\,X=0,U=Ann) \\ &=\; .40\;-\;.20\;=\;.20 \end{aligned} \tag{1.5}$$

of Ann may lead us to conclude that *the treatment is also beneficial for Ann*.

Hence, it seems that the treatment is *beneficial* for Joe and Ann. This, however, seems to contradict our finding ignoring the person variable. Just considering the prima facie effect

$$PFE_{10}\;=\;E(Y\,|\,X=1)-E(Y\,|\,X=0)\;=\;P(Y=1\,|\,X=1)-P(Y=1\,|\,X=0)\;=\;-.18,$$

ignoring the person variable U, the treatment seems *to be harmful*.

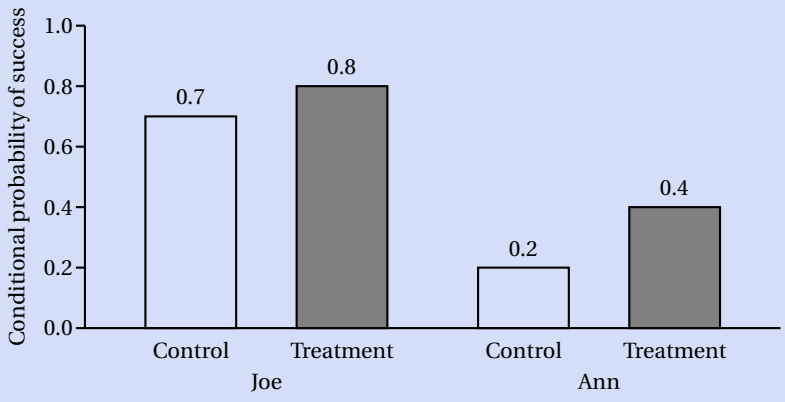

Figure 1.2. Conditional probabilities of success given treatment and person

1.1.5 Prima facie effect versus expectation of the individual total effects

In contrast to our intuition, the *prima facie effect* $E(Y|X=1)-E(Y|X=0)$ is *neither* the *simple average nor any weighted average of the corresponding individual total effects*

$$ITE_{U;10}(u) \;=\; E(Y|X=1,U=u) \;-\; E(Y|X=0,U=u)\,.$$

This is studied in more detail in the sequel.

Prima facie effect

The conditional probability $P(Y=1|X=0)$ of success given control is the sum of the corresponding probabilities $P(Y=1|X=0,U=Joe)$ and $P(Y=1|X=0,U=Ann)$, *weighted by the conditional probabilities* $P(U=Joe|X=0)$ and $P(U=Ann|X=0)$, respectively, that is,

$$\begin{aligned} P(Y=1|X=0) \;&=\; P(Y=1|X=0,U=Joe)\cdot P(U=Joe|X=0) + \\ &\qquad P(Y=1|X=0,U=Ann)\cdot P(U=Ann|X=0) \\ &= .7\cdot\frac{.48}{.6} \;+\; .2\cdot\frac{.12}{.6} \;=\; .6 \end{aligned}$$

[see SN-Box 9.2 (ii) and Exercise 1-10]. Because the difference between the conditional probabilities $P(U=Joe|X=0)=.48/.6$ and $P(U=Ann|X=0)=.12/.6$ is large, the probability of success in treatment 0 is much closer to .7 than to .2 (see the dots above $X=0$ in Fig. 1.3).

Similarly, the conditional probability $P(Y=1|X=1)$ of success given treatment condition $(X=1)$ is the sum of the two corresponding individual conditional probabilities $P(Y=1|X=1,U=Joe)$ and $P(Y=1|X=1,U=Ann)$, *weighted by the conditional probabilities* $P(U=Joe|X=1)$ and $P(U=Ann|X=1)$, respectively, that is,

$$\begin{aligned} P(Y=1|X=1) \;&=\; P(Y=1|X=1,U=Joe)\cdot P(U=Joe|X=1) + \\ &\qquad P(Y=1|X=1,U=Ann)\cdot P(U=Ann|X=1) \\ &= .8\cdot\frac{.02}{.4} + .4\cdot\frac{.38}{.4} \;=\; .42. \end{aligned}$$

Hence, the *prima facie effect* is

$$\begin{aligned} PFE_{10} \;&=\; P(Y=1|X=1)-P(Y=1|X=0) \\ &= \sum_u P(Y=1|X=1,U=u)\cdot P(U=u|X=1) - \\ &\quad \sum_u P(Y=1|X=0,U=u)\cdot P(U=u|X=0) \\ &= .42-.60. \;=\; -.18. \end{aligned} \tag{1.6}$$

Because the two $(X=1)$-conditional probabilities $P(U=Joe|X=1)=.02/.4=.05$ and $P(U=Ann|X=1)=.38/.4=.95$ are very different, the probability of success in treatment 1 is much closer to .4 than to .8 (see the dots above $X=1$ in Fig. 1.3). (The size of the area of the dotted circles is proportional to the conditional probabilities $P(U=u|X=x)$ that are used in the computation of the conditional expectation values $E(Y|X=x)$. This kind of graphics has been adopted from Agresti, 2007).

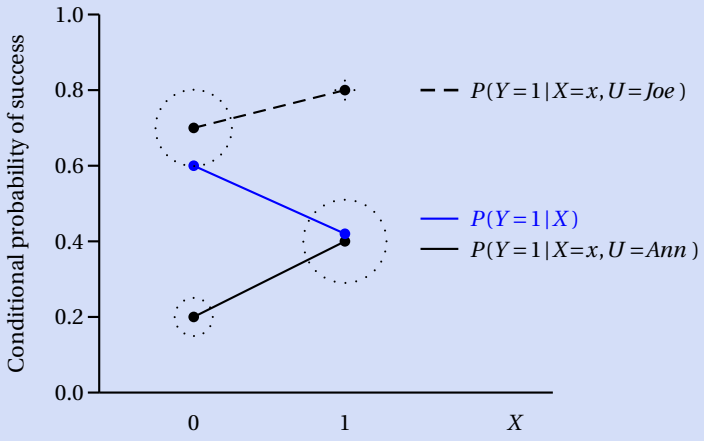

Figure 1.3. Conditional probabilities of success given treatment and person

Expectation of the individual total effects

The prima facie effect is not identical to the *expectation of the individual total effects*, which is the expectation of the function $ITE_{U;10}(U)$, the values of which are the two individual total effects $ITE_{U;10}(Joe)$ and $ITE_{U;10}(Ann)$ for Joe and Ann, respectively, that is,

$$\begin{aligned} E\big(ITE_{U;10}(U)\big) &= \sum_u ITE_{U;10}(u)\cdot P(U=u) \\ &= \sum_u P(Y=1\,|\,X=1,U=u)\cdot P(U=u) - \\ &\quad \sum_u P(Y=1\,|\,X=0,U=u)\cdot P(U=u). \end{aligned} \tag{1.7}$$

Because the two individual effects are $ITE_{U;10}(Joe) = .1$ and $ITE_{U;10}(Ann) = .2$,

$$E\big(ITE_{U;10}(U)\big) = .1\cdot P(U=Joe) + .2\cdot P(U=Ann) = .1\cdot\frac{1}{2} + .2\cdot\frac{1}{2} = .15.$$

Hence, whereas the *prima facie effect* $PFE_{10} = P(Y=1\,|\,X=1) - P(Y=1\,|\,X=0)$ is *negative*, namely −.18, the *expectation of the individual total-effect function* is *positive*, namely .15 (see Exercise 1-15).

1.1.6 How to evaluate the treatment?

The conclusions drawn from the prima facie effect

$$PFE_{10} = P(Y=1\,|\,X=1) - P(Y=1\,|\,X=0)$$

and from the individual effects

$$ITE_{U;10}(u) = P(Y=1\,|\,X=1,U=u) - P(Y=1\,|\,X=0,U=u)$$

Table 1.5. Random experiment of a two-factorial nonorthogonal design

U	Z	$P(U=u)$	$P(X=1\,\vert\,U=u)$	$P(X=2\,\vert\,U=u)$	$E(Y\,\vert\,U=u,X=0)$	$E(Y\,\vert\,U=u,X=1)$	$E(Y\,\vert\,U=u,X=2)$
Tom	*low*	1/8	10/60	3/60	120	100	80
Tim	*low*	1/8	18/60	9/60	120	100	80
Joe	*med*	1/8	26/60	17/60	90	90	70
Jim	*med*	1/8	26/60	17/60	100	100	80
Ann	*med*	1/8	26/60	17/60	120	100	100
Eva	*med*	1/8	26/60	17/60	130	110	110
Sue	*hi*	1/8	12/60	44/60	60	100	140
Mia	*hi*	1/8	16/60	36/60	60	100	140

are contradictory. Which of these comparisons should we trust? Is the treatment harmful as $P(Y=1\,|\,X=1)-P(Y=1\,|\,X=0)=-.18$ suggests? Or is it beneficial as suggested by the two positive differences $P(Y=1\,|\,X=1,U=u)-P(Y=1\,|\,X=0,U=u)$? Which of these comparisons are meaningful for evaluating the causal total effect of the treatment on the success variable Y? Before we come back to these questions, we consider another example.

1.2 Example 2 — Nonorthogonal Two-Factorial Experiment

In this section, we treat an example with three treatment conditions, representing two treatments and a control, for instance. Furthermore, there are a discrete covariate with three values, representing, for example, educational status, and a quantitative outcome variable, indicating the degree of success, for instance.[1] The relevant parameters of this random experiment are displayed in Table 1.5. For this example to be realistic, we have to assume that there is still variation of Y in each combination of person and treatment condition. This conditional variance may be due to (a) *measurement error*, but also to (b) *mediator effects*, that is, effects of variables and events that are in between X and the outcome variable Y in the process considered. Because Y is quantitative and is subject to measurement error and mediator effects, a full table similar to Table 1.2 with all possible

[1] In this example, we consider a (3×3)-factorial design with crossed, non-orthogonal factors. The analysis of such designs has been puzzling many statisticians (see, e. g., Aitkin, 1978; Appelbaum & Cramer, 1974; Carlson & Timm, 1974; Gosslee & Lucas, 1965; Jennings & Green, 1984; Keren & Lewis, 1976; Kramer, 1955; Overall & Spiegel, 1969, 1973b, 1973a; Overall, Spiegel, & Cohen, 1975; Williams, 1972), and it continues to do so (see, e. g., Langsrud, 2003; Nelder & Lane, 1995). In fact, none of the statistical packages such as SAS, SysStat, or SPSS with their Type I, II, III or IV sums of squares provide correct estimates and tests of the average effects (or main effects) for such a design unless the second factor has a uniform distribution, with equal probabilities for all values of the second factor. In this case Type III analysis yields correct results, at least, if the second factor is assumed to be fixed. However, in most applications in the social sciences, the second factor is not fixed but stochastic with varying sample means and sample variances, for instance. In chapter **??**, we will outline a correct analysis including average total effects.text

Table 1.6. Conditional expectation values of the outcome variable given treatments

Treatment condition	$E(Y \mid X{=}x)$	$P(X{=}x)$
$X=0$ (Control)	111.25	1/3
$X=1$ (Treatment 1)	100	1/3
$X=2$ (Treatment 2)	114.25	1/3
$E(Y)$	108.5	

outcomes, including those given person and treatment condition would be too large. It even may not exist if Y is actually continuous, which would be true if we would assume, for example, that Y has a normal distribution given the combination of a person u and a treatment condition x.

Nevertheless, it is still possible to present a compressed table that is analog to Table 1.1. Aside from the sampling probabilities for the eight persons and their person-specific probabilities of being assigned to treatment 1, this table contains the person-specific conditional expectation values $E(Y \mid X{=}x, U{=}u)$ of Y in both treatment conditions, $X=0$ and $X=1$. Such conditional expectation values have also been presented in Table 1.1 because the conditional probabilities $P(Y=1 \mid X{=}x, U{=}u)$ shown in that table are identical to the conditional expectation values $E(Y \mid X{=}x, U{=}u)$ if Y is binary with values 0 and 1 (see SN-Remark 9.8).

1.2.1 Prima facie effects

The conditional expectation values of the outcome variable Y given the three treatment conditions x are displayed in Table 1.6. The ratios in the last column are the treatment probabilities $P(X{=}x)$, which are $P(X{=}x) = 1/3$ for all three values x of X. Note that this is *not* a randomized design as will become obvious if we look at the second factor Z and the joint probabilities of X and Z (see Table 1.7). Furthermore, considering the conditional expectation values, and not the sample means, should make clear that we are not discussing *statistical inference* (i. e., inference from sample statistics to true parameters), but *causal inference*, that is, inference from the conditional expectation values such as $E(Y \mid X{=}x)$ or $E(Y \mid X{=}x, Z{=}z)$ to causal effects.

If our evaluation of the treatment effects were based on the differences between the conditional expectation values $E(Y \mid X{=}x)$ of Y in the three treatment conditions x, then we would conclude that there is a negative effect of treatment 1 compared to control, namely,

$$E(Y \mid X=1) \;-\; E(Y \mid X=0) \;=\; 100 - 111.25 \;=\; -11.25,$$

and a positive effect of treatment 2 compared to control, namely,

$$E(Y \mid X{=}2) \;-\; E(Y \mid X=0) \;=\; 114.25 - 111.25 \;=\; 3$$

(see Exercise 1-16).

Table 1.7. Conditional expectation values $E(Y \mid X{=}x, Z{=}z)$ given treatment and status

	Status						
Treatment	Low ($Z{=}0$)		Medium ($Z{=}1$)		High ($Z{=}2$)		
$X=0$	120	(20/120)	110	(17/120)	60	(3/120)	(40/120)
$X=1$	100	(7/120)	100	(26/120)	100	(7/120)	(40/120)
$X=2$	80	(3/120)	90	(17/120)	140	(20/120)	(40/120)
		(30/120)		(60/120)		(30/120)	

Note. Probabilities $P(X{=}x, Z{=}z)$, $P(Z{=}z)$, and $P(X{=}x)$ in parentheses.

1.2.2 Prima facie effects controlling for the qualitative covariate Z

A second attempt to evaluate the 'effects' of the *treatment* is to look at the differences between the conditional expectation values of Y in the three treatment conditions *given one of the three values of* Z: low, medium, and high. Note that these $(Z{=}z)$-conditional effects are also called *simple effects* in the literature on analysis of variance.

Table 1.7 displays the conditional expectation values of the outcome variable Y in the nine cells of the (3×3)-design. The ratios in parentheses are the probabilities that the pairs (x, z) of values of X and Z are observed. Hence, this table contains the conditional expectation values (true cell means) $E(Y \mid X{=}x, Z{=}z)$ of the outcome variable Y, and the joint probabilities $P(X{=}x, Z{=}z)$ determining the true joint distribution of X and Z.[2]

In the *low status condition* $(Z=0)$, there are large negative effects, both of treatment 1 and of treatment 2 compared to the control:

$$PFE_{Z;10}(0) = E(Y \mid X{=}1, Z{=}0) - E(Y \mid X{=}0, Z{=}0) = 100 - 120 = -20$$

and

$$PFE_{Z;20}(0) = E(Y \mid X{=}2, Z{=}0) - E(Y \mid X{=}0, Z{=}0) = 80 - 120 = -40.$$

In the *medium status condition* $(Z=1)$, there are also negative effects of treatment 1 and of treatment 2 compared to the control:

$$PFE_{Z;10}(1) = E(Y \mid X{=}1, Z{=}1) - E(Y \mid X{=}0, Z{=}1) = 100 - 110 = -10$$

and

$$PFE_{Z;20}(1) = E(Y \mid X{=}2, Z{=}1) - E(Y \mid X{=}0, Z{=}1) = 90 - 110 = -20.$$

Finally, in the *high status condition* $(Z=2)$, the effects of treatment 1 and treatment 2 are both positive:

$$PFE_{Z;10}(2) := E(Y \mid X{=}1, Z{=}2) - E(Y \mid X{=}0, Z{=}2) = 100 - 60 = 40$$

and

$$PFE_{Z;20}(2) := E(Y \mid X{=}2, Z{=}2) - E(Y \mid X{=}0, Z{=}2) = 140 - 60 = 80.$$

Based on these comparisons, we can conclude that the 'effects' of the treatments depend on the status of the subjects: the differences between the expectations of Y are negative for subjects with low and medium status, and they are positive for the subjects with high status.

[2] In this context, 'true' just indicates that we are not referring to sample means or relative frequencies in a sample. Instead these are the true means around which sample means would fluctuate.

1.2.3 Individual effects

In this fictive example we can also look at the individual effects of treatment 1 compared to control and treatment 2 compared to control. These both effects can be read from Table 1.5 for each person. For example, for Tom the individual effect of treatment 1 compared to control is

$$ITE_{U;10}(Tom) \;=\; E(Y \mid X=1, Tom) - E(Y \mid X=0, Tom) \;=\; 100-120 \;=\; -20,$$

and his individual effect of treatment 2 compared to control is

$$ITE_{U;20}(Tom) \;=\; E(Y \mid X=1, Tom) - E(Y \mid X=0, Tom) \;=\; 80-120 \;=\; -40.$$

Correspondingly, for Joe the individual effect of treatment 1 compared to control is

$$ITE_{U;10}(Joe) \;=\; E(Y \mid X=1, Joe) - E(Y \mid X=0, Joe) \;=\; 90-90 \;=\; 0,$$

and his individual effect of treatment 2 compared to control is

$$ITE_{U;20}(Joe) \;=\; E(Y \mid X=1, Joe) - E(Y \mid X=0, Joe) \;=\; 70-90 \;=\; -20.$$

For the reasons mentioned in section 1.1.1, unlike the $(Z=z)$-conditional effects treated in section 1.2.2, the individual effects usually cannot be estimated in empirical applications. Nevertheless, they will play a crucial role in the definition of causal effects.

1.2.4 Average of the individual effects

Of course, individual effects are more informative than average treatment effects if we want to know which treatment is the best for which individual. Nevertheless, we might ask: What are the total treatment effects on average? And we can ask: If the total individual effects cannot be estimated in empirical applications under realistic assumptions, is it possible to estimate at least the total average effects? And if yes, under which conditions?

Note that we have two average effects in this example, because we can compare treatment 1 to control and treatment 2 to control. Because we already looked at the corresponding individual effects, we just have to compute their averages, that is, the expectations of these conditional effects over the distribution of the person variable U, that is,

$$\begin{aligned} E\big(ITE_{U;10}(U)\big) \;&=\; \sum_u ITE_{U;10}(u) \cdot P(U=u) \\ &=\; (100-120)\cdot\frac{1}{8} + (100-120)\cdot\frac{1}{8} + (90-90)\cdot\frac{1}{8} + \ldots + (100-60)\cdot\frac{1}{8} \;=\; 0. \end{aligned}$$

Hence, the average total effect of treatment 1 compared to the control is 0. Comparing treatment 2 to control yields

$$\begin{aligned} E\big(ITE_{U;20}(U)\big) \;&=\; \sum_u ITE_{U;20}(u) \cdot P(U=u) \\ &=\; (80-120)\cdot\frac{1}{8} + (80-120)\cdot\frac{1}{8} + (70-90)\cdot\frac{1}{8} + \ldots + (140-60)\cdot\frac{1}{8} \;=\; 0. \end{aligned}$$

According to this result, the average total effect of treatment 2 compared to the control is 0 as well.

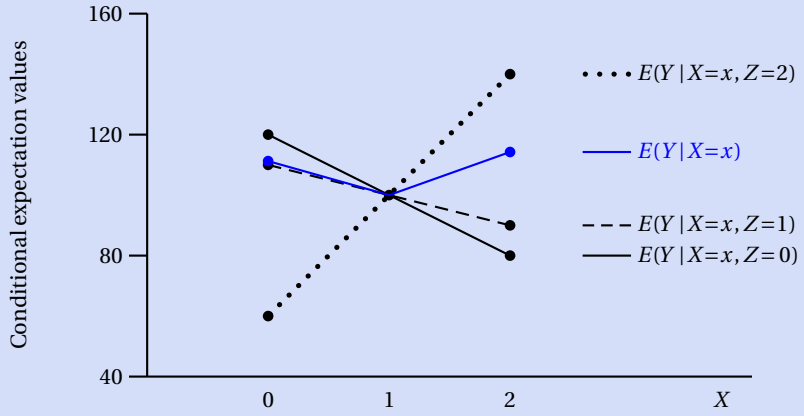

Figure 1.4. Conditional expectation values of Y given treatment and status

1.2.5 Average of the $(Z=z)$-conditional total effects

Now we consider the average of the $(Z=z)$-conditional total effects, where Z is the qualitative covariate *status*. Note that in analysis of variance with equal cell sizes, the averages of $(Z=z)$-conditional total effects are called the *main effects*.[3]

Because we already looked at the corresponding $(Z=z)$-conditional total effects (see section 1.2.2), we just have to compute their averages, more precisely, the expectations of these conditional effects over the distribution of status:

$$E\big(PFE_{Z;10}(Z)\big) = \sum_z PFE_{Z;10}(z)\cdot P(Z=z) = (-20)\cdot\frac{1}{4}+(-10)\cdot\frac{1}{2}+40\cdot\frac{1}{4} = 0.$$

Hence, the average effect of treatment 1 compared to the control is 0.

Comparing treatment 2 to control yields the average effect:

$$E\big(PFE_{Z;20}(Z)\big) = \sum_z PFE_{Z;20}(z)\cdot P(Z=z) = (-40)\cdot\frac{1}{4}+(-20)\cdot\frac{1}{2}+80\cdot\frac{1}{4} = 0.$$

According to this result, the average effect of treatment 2 compared to the control is 0 as well. Hence, in this example, these averages of the $(Z=z)$-conditional total effects are identical to the averages of the individual total effects. Is this just a coincidence? Or is this due to systematic conditions that hold in this example? If yes, which are these conditions? In chapters 6 to **??** we will provide the general answers.

[3] Note that we assume that Z is a random variable. In contrast, in analysis of variance it is assumed that Z is a fixed factor with a fixed number of observations for each value z of Z, that is, these numbers of observations are assumed to be invariant across different samples. In many empirical applications, this assumption is not realistic, but it does not invalidate the statistical conclusions as long as the parameters of interest do not involve the distribution of Z. However, a hypothesis about the average total effect *does* involve the distribution of Z, and this is the reason why programs on analysis of variance usually are not able to correctly estimate and test hypotheses about average total effects. For more details see section chapter **??** ??? RST.

1.2.6 How to evaluate the treatment?

To summarize, we discussed three ways that may, at first sight, be used to evaluate the treatment effects in empirical applications: *First*, we may compare the differences between the conditional expectation values $E(Y|X{=}x)$ of the outcome variable in the three treatment conditions $x{=}0$, $x{=}1$, and $x{=}2$. *Second*, we may consider the corresponding differences between the conditional expectation values $E(Y|X{=}x,Z{=}z)$ *given each of the three values* $z{=}0$, $z{=}1$, *and* $z{=}2$ *of status. Third*, we may compare the expectations of these differences between the $(X{=}x,Z{=}z)$-conditional expectation values over the distribution of Z (see Box 1.1 for a summary of these effects).[4] All these comparisons yield different results. Which of them are meaningful for the evaluation of the treatment effects? All three of them, or only two, just one, or none at all? And if one or more of these comparisons are meaningful, which assumptions are necessary allowing for causal interpretations?

1.3 Summary and Conclusion

In this chapter, we treated two examples. In the first one, a dichotomous treatment variable X has a negative (prima facie) effect $P(Y=1|X=1)-P(Y=1|X=0)$ on a *dichotomous outcome variable* Y ('success'), although the corresponding individual treatment effects

$$P(Y=1|X=1,U=Joe)-P(Y=1|X=1,U=Ann)$$

are positive. Taking the expectation of these two individual effects also yielded a positive effect. In the second example, there are nonzero differences $E(Y|X=1)-E(Y|X=0)$ and $E(Y|X{=}2)-E(Y|X=0)$, where Y is a *quantitative outcome variable*, and nonzero conditional 'effects' $E(Y|X=1,Z{=}z)-E(Y|X=0,Z{=}z)$ and $E(Y|X{=}2,Z{=}z)-E(Y|X=0,Z{=}z)$ for the different levels z of *status*. The expectations of these $(Z{=}z)$-conditional 'effects' (comparing treatment 1 to 0 and comparing treatment 2 to 0) over the three status conditions, that is, the two 'average total effects', are zero.

The problem

Because the conclusions drawn in each of these analyses are contradictory, which of these should we trust? In the first example: Is the treatment harmful — as the difference $P(Y=1|X=1)-P(Y=1|X=0)$ suggests? Or is it beneficial as suggested by the individual effects $P(Y=1|X=1,U=u)-P(Y=1|X=1,U=u)$? In the second example: Do the prima facie effects $E(Y|X=1)-E(Y|X=0)$ have a meaningful causal interpretation? Or do the $(Z{=}z)$-conditional prima facie effects $E(Y|X=1,Z{=}z)-E(Y|X=0,Z{=}z)$ have a meaningful causal interpretation? And, does this apply also to their expectation?

In the first example, we demonstrated that we can not expect that the difference

$$P(Y=1|X=1)-P(Y=1|X=0)$$

is the average (expectation) of the corresponding person-specific differences

[4] In fact, there are even more than three ways. Types II and III of computing the sums of squares in nonorthogonal ANOVA are not yet considered in our discussion. In chapter **??**, we show that all four types of computing sums of squares in such a design yield wrong results in our example (see also Exercise 1-17).

Box 1.1 Various total effects treated in this chapter

$PFE_{xx'}$ — *Prima facie effect* of treatment x compared to treatment x'. It is defined by

$$PFE_{xx'} := E(Y|X{=}x) - E(Y|X{=}x').$$

$PFE_{Z;xx'}(z)$ — $(Z{=}z)$-*conditional prima facie effect* of treatment x compared to treatment x'. It is defined by

$$PFE_{Z;xx'}(z) := E(Y|X{=}x, Z{=}z) - E(Y|X{=}x', Z{=}z).$$

$E\big(PFE_{Z;xx'}(Z)\big)$ — *Expectation of the* $(Z{=}z)$-*conditional prima facie effects* of treatment x compared to treatment x'. If Z is discrete, then it is defined by

$$E\big(PFE_{Z;xx'}(Z)\big) := \sum_z PFE_{Z;xx'}(z) \cdot P(Z{=}z).$$

$ITE_{U;xx'}(u)$ — *Individual effect* of treatment x compared to treatment x'. It is defined by

$$ITE_{U;xx'}(u) := E(Y|X{=}x, U{=}u) - E(Y|X{=}x', U{=}u).$$

$E\big(ITE_{U;xx'}(U)\big)$ — *Expectation of the individual effects* of treatment x compared to treatment x'. It is defined by

$$E\big(ITE_{U;xx'}(U)\big) := \sum_u ITE_{U;xx'}(u) \cdot P(U{=}u).$$

$$P(Y{=}1|X{=}1, U{=}u) - P(Y{=}1|X{=}0, U{=}u).$$

Similarly, in the second example, we showed that we can not expect that the difference

$$E(Y|X{=}1) - E(Y|X{=}0)$$

is the average (expectation) of the corresponding differences

$$E(Y|X{=}1, Z{=}z) - E(Y|X{=}0, Z{=}z)$$

given a value z of status. And, how do we know that these $(Z{=}z)$-conditional effects are meaningful for the evaluation of the treatment? As noted before, these questions are not related to *statistical* inference; they are not raised at the sample level, but on the level of true conditional expectation values!

Hence our examples show that conditional expectation values and their differences, the prima facie effects, can be totally misleading in evaluating the effects of a treatment variable X on an outcome variable Y. This conclusion can also be extended to conditional probabilities, to correlations and to all other parameters describing relationships and dependencies between random variables. They all are like the shadow in the metaphor of the invisible man (see the preface).

If this is true, is the whole idea of *learning from experience* — the core of empirical sciences — wrong? Our answer is 'No'. However, we have to be more explicit in what we

mean by terms like 'X affects Y', 'X has an effect on Y', 'X influences Y', 'X leads to Y', and so on used in our theories and hypotheses. How can these terms be uniquely defined in a language that is compatible with statistical analyses of empirical data? How to design an empirical study and how to look at the resulting data if we want to probe our theories and learn about the causal effects postulated in these theories and hypotheses?

In the chapters to come we will show that these parameters are meaningful *under certain conditions*, just like the shadow of the invisible man can be meaningful under certain conditions in order to measure his height. In the metaphor a crucial condition is the 45° angle of the sun. Do we also have such a crucial condition for causal inference? We know that a reversal of total effects does not occur in randomized experiments, that is, in experiments in which observational units (in the social and behavioral sciences, usually the subjects or individuals) are randomly assigned to one of at least two treatment conditions. In the randomized experiment comparing conditional expectation values *is* informative about total causal treatment effects. But why? What is so special in the randomized experiment? Which are the mathematical conditions that we create in a randomized experiment? Are there also conditions that can be utilized in quasi-experimental evaluation studies? How can we estimate causal effects in quasi-experiments? Obviously, conclusive answers to these questions can be hoped for only within a theory of causal effects.

Relevance of the problem

Obviously, these questions are of fundamental importance for the methodology of empirical sciences and for the empirical sciences themselves. The answers to these questions have consequences for the design and analysis of experiments, quasi-experiments, and other studies aiming at estimating the effects of *treatments*, *interventions*, or *expositions*. No *prevention study* can meaningfully be conducted and analyzed without knowing the concepts of causal effects and how they can be estimated from empirical data. Similarly, without a clear concept of causal effects we are not able to learn from our data about the effects of a certain (possibly harmful) environment on our health, or about the effects of certain behaviors such as smoking or drug abuse. Again, this is similar to the problem of measuring the invisible man's size via the length of his shadow: only with a clear concept of *size*, some basic knowledge in geometry, and the additional information such as the angle of the sun at the time of measurement are we able to determine his size from the length of his shadow.

Research traditions

Of course, raising these questions and attempting answers is not new. Immense knowledge and wisdom about experiments and quasi-experiments has been collected in the Campbellian tradition of experiments and quasi-experiments (see, e. g., Campbell & Stanley, 1963; T. D. Cook & Campbell, 1979; Shadish et al., 2002). In the last decades, a more formal approach has been developed supplementing the Campbellian theory and terminology in important aspects: the theory of causal effects in the Neyman-Rubin tradition (see, e. g., Splawa-Neyman, 1923/1990; Rubin, 1974, 2005). Many papers and books indicate the growing influence of this theory (see, e. g., Greenland, 2000, 2004; Höfler, 2005; Rosenbaum, 2002a; Rubin, 2006; Winship & Morgan, 1999; Morgan & Winship, 2007) and remarkable efforts have already been made to integrate it into the Campbellian framework

(West, Biesanz, & Pitts, 2000). Furthermore, these questions have also been dealt with in the graphical modeling tradition (see, e. g., Pearl, 2009; Spirtes et al., 2000) as well as in biometrics, econometrics, psychometrics, epidemiology, and other fields dealing with the methodology of empirical research.

Outlook

In this volume, we present the theory of causal total effects in terms of classical probability theory. We show that a number of questions that have been debated controversially and inconclusively can now be given a clear-cut answer. What kinds of causal effects can meaningfully be defined? Which design techniques allow for unbiased estimation of causal effects? How to analyze nonorthogonal ANOVA designs (cf., e. g., Aitkin, 1978; Appelbaum & Cramer, 1974; Gosslee & Lucas, 1965; Maxwell & Delaney, 2004; Overall et al., 1975)? How to analyze non-equivalent control-group designs (cf., e. g., Reichardt, 1979)? Should we compare pre-post differences between treatment groups (cf., e. g., Lord, 1967; Senn, 2006; van Breukelen, 2006; Wainer, 1991)? Should we use analysis of covariance to adjust for differences between treatment and control that already existed prior to treatment (cf., e. g., Maxwell & Delaney, 2004; Cohen, Cohen, West, & Aiken, 2003)? Should we use propensity score methods instead of the more traditional procedures mentioned above (cf., e. g., Rosenbaum & Rubin, 1983b)? How do we deal with non-compliance to treatment assignment (cf., e. g., Cheng & Small, 2006; Dunn et al., 2003; Jo, 2002a, 2002b, 2002c; Jo, Asparouhov, Muthén, Ialongo, & Brown, 2008; J. Robins & Rotnitzky, 2004; J. M. Robins, 1998)?

We do not treat the statistical sampling models with their distributional assumptions, their implications for parameter estimation, and the evaluation (or tests) of hypotheses about these parameters. However, we will discuss the virtues and problems of general strategies of data analysis such as the analysis of difference scores, analysis of covariance, its generalizations, and analysis based on propensity scores.

1.4 Exercises

▷ **Exercise 1-1** Why do we need the concept of a causal treatment effect?

▷ **Exercise 1-2** What is the relationship between the unconditional prima facie effect PFE_{10} and the expectations $E(Y \mid X = 0)$ and $E(Y \mid X = 1)$ of the outcome variable Y in the two treatment conditions?

▷ **Exercise 1-3** Verify that Table 1.3 is in fact obtained by collapsing the two corresponding tables for Joe and Ann (see Table 1.4). RST ???

▷ **Exercise 1-4** Which are the kinds of prima facie effects treated in this chapter?

▷ **Exercise 1-5** What is the difference between statistical inference and causal inference?

▷ **Exercise 1-6** Why are the conditional expectation values $E(Y \mid X{=}x)$ in treatment conditions x also the $(X{=}x)$-conditional probabilities for the event $\{Y{=}1\}$ in the first example treated in this chapter?

▷ **Exercise 1-7** Download *K-book table 1.1.sav* from *www.causal-effects.de*. This data set has been generated from Table 1.1 for a sample of size $N = 10{,}000$. Compute the contingency table corresponding to Table 1.3 and the associated estimates of the conditional probabilities $P(Y = 1 \mid X = 0)$ and $P(Y = 1 \mid X = 1)$.

▷ **Exercise 1-8** Use $P(Y=1\,|\,X=1) = .42$ and $P(Y=1\,|\,X=0) = .6$ computed in section 1.1.3 in order to compute the corresponding odds ratio, its logarithm, and the risk ratio, according to the definitions of these parameters presented in SN-Remarks 13.14 to 13.16.

▷ **Exercise 1-9** Compute the conditional probability $P(Y=1\,|\,X=1, U=Joe)$ from Table 1.4.

▷ **Exercise 1-10** Compute the probability $P(Y=1\,|\,X=0)$ from the corresponding conditional probabilities $P(Y=1\,|\,X=0, U=u)$.

▷ **Exercise 1-11** What (i.e., how big) are the unconditional prima facie effects of the treatments, that is, the prima facie effects $E(Y\,|\,X=1) - E(Y\,|\,X=0)$ and $E(Y\,|\,X=2) - E(Y\,|\,X=0)$ in the second example of this chapter?

▷ **Exercise 1-12** What are the conditional prima facie effects of the treatments, that is, the prima facie effects $E(Y\,|\,X=1, Z=z) - E(Y\,|\,X=0, Z=z)$ and $E(Y\,|\,X=2, Z=z) - E(Y\,|\,X=0, Z=z)$ in the second example of this chapter?

▷ **Exercise 1-13** What are the averages of the conditional prima facie effects

$$E(Y\,|\,X=1, Z=z) \;-\; E(Y\,|\,X=0, Z=z) \qquad \text{and} \qquad E(Y\,|\,X=2, Z=z) \;-\; E(Y\,|\,X=0, Z=z)$$

in the second example of this chapter?

▷ **Exercise 1-14** Compute the conditional probability $P(U=Tom \mid X=0)$ from the parameters presented in Tables 1.5 and 1.6.

▷ **Exercise 1-15** Open the Causal Effects Xplorer with table *K-book table 1.1.tab*. Change the conditional probabilities $P(X=1\,|\,U=u)$ of receiving treatment 1 for Joe and Ann to 2/5. Then compare the two individual treatment effects of Joe and Ann and their average to the prima facie effect $E(Y\,|\,X=1) - E(Y\,|\,X=0)$.

▷ **Exercise 1-16** Open the Causal Effects Xplorer with table *K-book table 1.1.tab* displaying the conditional probabilities $P(U=u\,|\,X=x)$. Then use SN-Box 9.2 (ii) in order to compute the three conditional expectation values $E(Y\,|\,X=x)$ displayed in Table 1.6 from the parameters presented in Table 1.5.

▷ **Exercise 1-17** Download *K-book table 1.5.sav* from *www.causal-effects.de*. This data set has been generated (with the Causal Effects Xplorer) from Table 1.5 for a sample of size $N = 10{,}000$ with error variance 10 given each person.

(a) Compute the cell means and the relative frequencies of observations in each of the nine cells of the (3×3)-table.

(b) Use each of the procedures offered by your statistical program package to analyze the data including a test of the main effects of the treatment factor (most programs offer Typ I, II and III sums of squares for such an analysis).

(c) Compare the results of these analyses to the parameters presented in Table 1.7.

Solutions

▷ **Solution 1-1** We need the concept of a causal treatment effect, because the two examples show that differences between conditional expectation values are meaningless for the evaluation of the effects of a treatment, unless we can show how the differences between these conditional expectation values are related to the causal treatment effects. Obviously, without a definition of causal treatment effects this is be possible. Estimating causal treatment effects is crucial for answering questions such as 'Does the treatment help our patients with respect to the outcome variable considered?'

▷ **Solution 1-2** The unconditional prima facie effect PFE_{10} is defined as the difference between the two conditional expectation values $E(Y|X=1)$ and $E(Y|X=0)$.

▷ **Solution 1-3** This can easily be verified by adding the probabilities for the observations of the pairs (x,z) of X and Z over males and females. This yields .144 + .096 = .240, .004 + .228 = .232, .336 + .024 = .360 and .016 + .152 = .168.

▷ **Solution 1-4** The two kinds of prima facie effects treated in this chapter are: the *unconditional prima facie effect*, and the *conditional prima facie effect* given the value z of a potential confounder Z. The unconditional prima facie effect of treatment 1 compared to treatment 0 is the difference $PFE_{10} = E(Y|X=1) - E(Y|X=0)$ between the conditional expectation values of an outcome variable Y given the two treatment conditions. The $(Z{=}z)$-conditional prima facie effect is the difference $PFE_{Z;10}(z) = E(Y|X{=}1, Z{=}z) - E(Y|X{=}0, Z{=}z)$ between the $(X=1, Z{=}z)$-conditional expectation value and the $(X=0, Z{=}z)$-conditional expectation value of the outcome variable Y.

▷ **Solution 1-5** In *statistical* inference we estimate and test hypotheses about parameters characterizing the (joint or marginal) distributions of random variables from sample data. In *causal* inference we interpret some of these parameters as causal effects, provided that certain conditions are satisfied that allow for such a causal interpretation.

▷ **Solution 1-6** $E(Y|X{=}x) = P(Y{=}1|X{=}x)$, because, in this example, Y is dichotomous with values 0 and 1. In this case, the term $P(Y{=}1|X{=}x)$ is defined by $E(Y|X{=}x)$ (see SN-Remark 9.8).

▷ **Solution 1-7** No solution provided. Just compare your results to the true parameters presented in Table 1.3 and to the conditional probabilities $P(Y=1|X=0)$ and $P(Y=1|X=1)$ presented in section 1.1.3.

▷ **Solution 1-8** The *odds ratio* is

$$\frac{P(Y=1|X=1)}{1-P(Y=1|X=1)} \Big/ \frac{P(Y=1|X=0)}{1-P(Y=1|X=0)} \approx .483.$$

Because this number is smaller than 1 it indicates that there is a negative effect of the treatment. The natural logarithm of the odds ratio is the *log odds ratio*, which is

$$\ln\left[\frac{P(Y=1|X=1)}{1-P(Y=1|X=1)} \Big/ \frac{P(Y=1|X=0)}{1-P(Y=1|X=0)}\right] \approx -0.728.$$

This number is smaller than 0 indicating that there is a negative effect of the treatment. The log odds ratio is identical to the *logistic regression coefficient* λ_1 in the equation

$$P(Y=1|X) = \frac{\exp(\lambda_0 + \lambda_1 \cdot X)}{1+\exp(\lambda_0 + \lambda_1 \cdot X)}.$$

Another closely related parameter is the *risk ratio*

$$\frac{P(Y=1|X=1)}{P(Y=1|X=0)} = .7.$$

Because this number is smaller than 1, it indicates that there is a negative effect of the treatment. Hence, no matter which of these parameters we use, we would always come to the same (wrong) conclusion that the treatment is detrimental for our patients.

▷ **Solution 1-9** Using the joint probabilities presented in Table 1.4, the definition of the conditional probability $P(Y=1|X=1, U=Joe)$

$$P(Y=1|X=1, U=Joe) = \frac{P(X=1, Y=1, U=Joe)}{P(X=1, U=Joe)} = \frac{.016}{.016+.004} = .8.$$

▷ **Solution 1-10** First of all, note that the theorem of total probability [see SN-Th. 4.25 (ii)], can also be applied to conditional probabilities (see SN-Th. 4.28), in this exercise, the $(X=0)$-conditional probabilities. Hence, according to this theorem,

$$\begin{aligned} P(Y=1\,|\,X=0) \;=\; & P(Y=1\,|\,X=0, U=\mathit{Joe})\cdot P(U=\mathit{Joe}\,|\,X=0) + \\ & P(Y=1\,|\,X=0, U=\mathit{Ann})\cdot P(U=\mathit{Ann}\,|\,X=0). \end{aligned}$$

The probabilities $P(Y=1\,|\,X=0, U=\mathit{Joe}) = .7$ and $P(Y=1\,|\,X=0, U=\mathit{Ann}) = .2$ are computed analogously to Exercise 1-9 and the other two probabilities occurring in this formula are $P(U=\mathit{Joe}\,|\,X=0) = .48/.6$ and $P(U=\mathit{Ann}\,|\,X=0) = .12/.6$ (see Table 1.4). Hence,

$$P(Y=1\,|\,X=0) \;=\; \frac{.7\cdot.48}{.6} \;+\; \frac{.2\cdot.12}{.6} \;=\; .6.$$

▷ **Solution 1-11** The prima facie effects $E(Y\,|\,X=1)-E(Y\,|\,X=0)$ and $E(Y\,|\,X=2)-E(Y\,|\,X=0)$ can be computed from Table 1.6. They are as follows:

$$\mathit{PFE}_{10} \;=\; E(Y\,|\,X=1)-E(Y\,|\,X=0) \;=\; 100.00-111.25 \;=\; -11.25$$

and

$$\mathit{PFE}_{20} \;=\; E(Y\,|\,X=2)-E(Y\,|\,X=0) \;=\; 114.25-111.25 \;=\; 3.00.$$

▷ **Solution 1-12** The conditional prima facie effects

$$E(Y\,|\,X=1, Z=z)-E(Y\,|\,X=0, Z=z) \quad \text{and} \quad E(Y\,|\,X=2, Z=z)-E(Y\,|\,X=0, Z=z)$$

can be computed from Table 1.7. For *low status* $(Z=0)$, they are:

$$\mathit{PFE}_{Z;10}(0) \;=\; E(Y\,|\,X=1, Z=0)-E(Y\,|\,X=0, Z=0) \;=\; 100-120 \;=\; -20$$

$$\mathit{PFE}_{Z;20}(0) \;=\; E(Y\,|\,X=2, Z=0)-E(Y\,|\,X=0, Z=0) \;=\; 80-120 \;=\; -40.$$

For *medium status* $(Z=1)$, they are:

$$\mathit{PFE}_{Z;10}(1) \;=\; E(Y\,|\,X=1, Z=1)-E(Y\,|\,X=0, Z=1) \;=\; 100-110 \;=\; -10$$

$$\mathit{PFE}_{Z;20}(1) \;=\; E(Y\,|\,X=2, Z=1)-E(Y\,|\,X=0, Z=1) \;=\; 90-110 \;=\; -20.$$

Finally, for *high status* $(Z=2)$, the conditional prima facie effects are:

$$\mathit{PFE}_{Z;10}(2) \;=\; E(Y\,|\,X=1, Z=2)-E(Y\,|\,X=0, Z=2) \;=\; 100-60 \;=\; 40$$

$$\mathit{PFE}_{Z;20}(2) \;=\; E(Y\,|\,X=2, Z=2)-E(Y\,|\,X=0, Z=2) \;=\; 140-60 \;=\; 80.$$

▷ **Solution 1-13** Using the results of the last exercise, the average of the $(Z=z)$-conditional prima facie effects can be computed from the conditional effects as follows:

$$\begin{aligned} E\big(\mathit{PFE}_{Z;10}(Z)\big) \;&=\; \mathit{PFE}_{Z;10}(0)\cdot P(Z=0) + \mathit{PFE}_{Z;10}(1)\cdot P(Z=1) + \mathit{PFE}_{Z;10}(2)\cdot P(Z=2) \\ &=\; (-20)\cdot\frac{1}{4} + (-10)\cdot\frac{1}{2} + 40\cdot\frac{1}{4} \;=\; 0. \end{aligned}$$

$$\begin{aligned} E\big(\mathit{PFE}_{Z;20}(Z)\big) \;&=\; \mathit{PFE}_{Z;20}(0)\cdot P(Z=0) + \mathit{PFE}_{Z;20}(1)\cdot P(Z=1) + \mathit{PFE}_{Z;20}(2)\cdot P(Z=2) \\ &=\; (-40)\cdot\frac{1}{4} + (-20)\cdot\frac{1}{2} + 80\cdot\frac{1}{4} \;=\; 0. \end{aligned}$$

▷ **Solution 1-14**

$$\begin{aligned} P(U=\mathit{Tom}\;|\;X=0) \;&=\; \frac{P(U=\mathit{Tom}, X=0)}{P(X=0)} \;=\; \frac{P(X=0\,|\,U=\mathit{Tom})\cdot P(U=\mathit{Tom})}{P(X=1)} \\ &=\; \frac{(47/60)\cdot(1/8)}{1/3} \;=\; \frac{47}{160}. \end{aligned}$$

▷ **Solution 1-15** With this change, the prima facie effect changes to $E(Y|X=1) - E(Y|X=0) = .6 - .45 = .15$, which is the average of the two individual total effects, which still are .10 for Joe and .20 for Ann. Note that identical treatment probabilities $P(X=1|U=u)$ for all persons u is what we create by randomly assigning a person to treatment 1 in a randomized experiment.

▷ **Solution 1-16** Rewriting SN-Box 9.2 (ii) for our example,

$$E(Y|X{=}x) = \sum_u E(Y|X{=}x, U{=}u) \cdot P(U{=}u|X{=}x).$$

Hence,

$$\begin{aligned} E(Y|X=0) &= \sum_u E(Y|X=0, U=u) \cdot P(U=u|X=0) \\ &= 120 \cdot \frac{47}{160} + 120 \cdot \frac{33}{160} + \ldots + 60 \cdot \frac{8}{160} = 111.25, \end{aligned}$$

$$\begin{aligned} E(Y|X=1) &= \sum_u E(Y|X=1, U=u) \cdot P(U=u|X=1) \\ &= 100 \cdot \frac{10}{160} + 100 \cdot \frac{18}{160} + \ldots + 100 \cdot \frac{16}{160} = 100, \end{aligned}$$

and

$$\begin{aligned} E(Y|X=2) &= \sum_u E(Y|X=2, U=u) \cdot P(U=u|X=2) \\ &= 80 \cdot \frac{3}{160} + 80 \cdot \frac{9}{160} + \ldots + 140 \cdot \frac{36}{160} = 114.25. \end{aligned}$$

▷ **Solution 1-17** No solution provided. Just compare your results to the parameters presented in Table 1.7.

Chapter 2
Some Typical Kinds of Random Experiments

RST: 28. Febr. 2019

Einige Stellen sind noch mit RST markiert. Hier fehlt noch eine Entscheidung, ob auch kausale Effekte einer pre-treatment variable auf X und Y betrachtet werden sollen. Nur wenn in spaeteren Kapiteln explizit solche Beispiele behandelt werden.

In chapter 1 we have shown that comparing conditional expectation values of an outcome variable between treatment groups can be completely misleading if used for the evaluation of treatment effects. In this chapter we continue preparing the stage for the theory of causal total effects, describing the kind of empirical phenomena it refers to: single-unit trials of experiments or quasi-experiments, but also single-unit trials of observational studies in which causal total effects and dependencies can be investigated. First examples of such a single-unit trial have already been treated in sections 1.1 and 1.2 of chapter 1.

A single-unit trial is a specific random experiment. Note the distinction between a *random experiment* and a *randomized experiment*. Stochastic dependencies between events and between random variables always refer to a random experiment, but not necessarily to a *randomized experiment* in which a subject is assigned to one of the treatment conditions by a randomization procedure. In the simplest case of such a randomization we assign the subject to treatment or control according to the outcome of flipping a coin. In contrast, a *random experiment* is the concrete empirical phenomenon to which stochastic dependencies between events and random variables (described by conditional distributions, probabilities, correlations, and conditional expectations) refer to.

The single-unit trial *is not the sample* dealt with in statistical models. In a sample, we consider repeating the single-unit trial many times in one way or another. This is necessary if we want to deal with estimation of parameters and tests of hypotheses about these parameters, some of which might be causal effects. The single-unit trial does *not allow* treating problems of parameter estimation or hypothesis testing. However, it is sufficient for defining causal effects and studying how to identify them, that is, investigating under which conditions and how they can be computed from empirically estimable parameters.

A single-unit trial is also what we refer to in hypotheses and theories of the empirical sciences. In many text books on applied statistics the dazzling term ‘population’ is used instead, obfuscating what we are actually talking about when we use probabilistic terms such as expectation, variance, covariance, regression, etc. Furthermore, single-unit trials are what is of interest in practical work. How does the treatment of a patient affect the outcome of this patient if compared to another possible treatment? What is the treatment effect for a male, and what is its effect for a female? Which variables explain inter-individual differences in individual causal effects? All these questions are raised using concepts referring to a single-unit trial.

Overview

We start with the single-unit trial of simple experiments and then treat increasingly more complex ones introducing additional design features. Specifically, we will introduce the single-unit trials of experiments and quasi-experiments with fallible covariates, a multifactorial design with more than one treatment, multilevel experiments and quasi-experiments, and experiments and quasi-experiments with latent covariates and/or outcome variables.

We also discuss different kinds of random variables that will play a crucial role in the chapters to come. Among these random variables are the observational-unit variable, manifest and latent covariates, treatment variables, as well as manifest and latent outcome variables. In this chapter, we confine ourselves to an informal description of single-unit trials and the random variables involved, preparing the stage for their mathematical representations in the following chapters.

2.1 Simple Experiments

As a first class of random experiments we consider the single-unit trials of *simple experiments and quasi-experiments*. Such single-unit trials are experiments and quasi-experiments in which *no fallible covariates* are assessed.

Such a single-unit trial consists of:

(a) sampling an observational unit u (e. g., a person) from a set of units,
(b) assigning the unit or observing its assignment to one of several experimental conditions (represented by the value x of the treatment variable X),
(c) recording the value y of the outcome (or response) variable Y.

Figure 2.1 displays a tree representation of the set of possible outcomes of this single-unit trial. Note that this is the kind of random experiment we considered in the Joe-Ann example presented in section 1.1 and in the two-factorial design example treated in section 1.2. The random variables X (treatment), Y (success), and Z (status), the conditional expectation values $E(Y|X=x)$ and $E(Y|X=x,Z=z)$, as well as the probabilities $P(X=x)$, $P(Z=z)$, $P(X=x,Z=z)$ all referred to such a single-unit trial. Of course, all these conditional expectation values and probabilities are unknown in empirical applications. Nevertheless, they are the parameters that determine the outcome of the single-unit trial, just in the same way as the probability of tossing *heads* determines the outcome of flipping a coin.

In order to illustrate this point, imagine flipping a deformed coin that has the shape of a Chinese wok, and suppose that in this case the probability of flipping *heads* is .80 instead of .50. Although this probability does not deterministically determine the outcome of flipping the coin, it stochastically determines the outcome.

In fact, we may consider the single-unit trial of (a) sampling a coin u from a set of coins, (b) forming ($X=1$) or not forming ($X=0$) a wok out of it, and (c) observing whether ($Y=1$) or not ($Y=0$) we flip *heads*. In this single-unit trial, the difference $.80-.50=.30$ would be the causal total effect of the treatment variable X on the outcome variable Y. Note that the probabilities .80 and .50 and their difference .30 refer to this single-unit trial, although these probabilities can only be estimated if we conduct many of these single-unit trials, that is, if we draw a data sample. However, if these probabilities were known, we could dis-

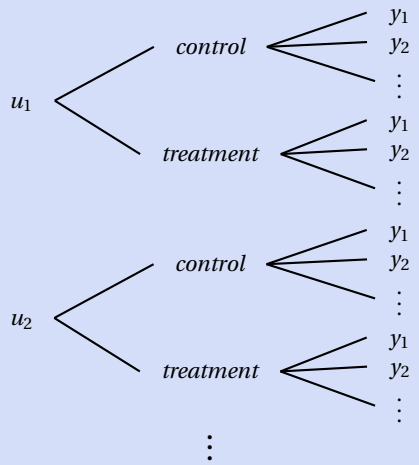

Figure 2.1. A simple experiment or quasi-experiment

pense with a sample (including the data that would result from drawing it), and still have a perfect theory and prediction for the outcome of such a single-unit trial (see Exercise 2-1).

Sampling a unit

The first part of this single-unit trial consists of sampling an observational unit. In the social sciences, units often are persons, but they might be groups, school classes, schools and even countries. Usually such units change over time. Therefore, it should be emphasized that, in simple experiments and quasi-experiments, we are talking about the *units at the onset of treatment*. Later we will see that we have to distinguish between *units at the onset of treatment* and *units at the time of assessment of the outcome variable*, which might be months or even years later. In a single-unit trial of simple experiments and quasi-experiments, the units can be represented by the observational-unit variable U, whose possible values u are the *units at the onset of treatment*.

Note that the unit at the onset of treatment also comprises his or her experiences a year and/or the day before treatment, as well as the psycho-bio-social situation in which he or she is at the onset of treatment. Both, the experiences and the situation, already happen *before* the onset of treatment (see Steyer, Mayer, Geiser, & Cole, 2015 for more details). Therefore, they are attributes of the observational units u. They can be treated in the same way as other attributes such as sex and educational status. However, if these attributes are actually assessed and if this assessment is fallible, then we have to distinguish between these attributes and their fallible assessments (see section 2.2).

Treatment variable

In an experiment or quasi-experiment, there is always a treatment variable, which we focus as a cause[1] and usually denote it by X. In a *true experiment*, a unit drawn is assigned — for example, by the experimenter or by some other person (such as a physician, a psychologist, or a social worker) — to one of the possible treatments. In contrast, in a *quasi-experiment* we just observe self-selection to one of the treatment conditions. In the simplest case there are at least two treatment conditions, for example, *treatment* and *control*. These treatment conditions are the possible values of the treatment variable X. For simplicity, we use the values $0, 1, \ldots, J$ to represent $J+1$ treatment conditions. Furthermore, unless stated otherwise, we presume that treatment *assignment* and actual *exposure* to treatment are equivalent, that is, we assume that there is perfect compliance.

Selection of a unit into one of the treatment conditions x may happen with unknown probabilities, for example, when there is self-selection or assignment by an unknown physician. This is often the case in quasi-experiments. However, assignment can also be done with known probabilities that are equal for different units (such as in the simple randomized experiment) or with known probabilities that may be unequal for different units (such as in the conditionally randomized experiment). In this case, these treatment probabilities may also depend on a covariate Z representing pre-treatment attributes of the units. As mentioned above, *conditional and unconditional randomized assignment, distinguish the true experiment from the quasi-experiment*, in which the assignment probabilities are unknown. (See Remarks 8.30 and 8.31 for more details on randomization and conditional randomization.)

Potential confounders and covariates

In simple experiments and quasi-experiments, the focus is usually on total treatment effects on an outcome variable. Hence, if we are interested in the treatment variable as a cause, then each attribute of the observational units is a potential confounder. Examples are *sex*, *race*, *educational status*, and *socio-economic status*. Once the unit is drawn, its *sex*, *race*, *educational status*, and *socio-economic status* are fixed. This means that there is no additional sampling process associated with assessing these potential confounders. This is also the reason why they do not appear in points (a) to (c) describing the single-unit trial.

A potential confounder is also called a *covariate* if it is actually assessed and used together with X in a conditional expectation or a conditional distribution. Note that a potential confounder can also be unobserved, and in this case we usually do not call it a covariate.

Because potential confounders represent attributes of the unit *at the onset of treatment* they can never be affected by the treatment. However, there can be (stochastic) dependencies between the treatment variable and potential confounders. In the Joe-Ann example treated in section 1.1, for instance, there is a stochastic dependence of the treatment variable X and the person variable U. Similarly, in the second example presented in section 1.2 there is a stochastic dependence of Z (status) and the treatment variable X.

[1] We use the term (putative) cause for a random variable if we consider its causal effect on an outcome variable. Note that a causal effect can also be 0.

Multidimensional potential confounders and covariates

Potential confounders – and therefore also covariates – may be uni- or multi-dimensional, qualitative (such as Z_1 :=*sex* and Z_2 := *educational status*) or quantitative (such as Z_3 := *height* and Z_4 :=*body mass index*) or, if it is a multivariate variable made up of several uni-dimensional variables, it may consist of qualitative *and* quantitative potential confounders such as $Z_5 = (Z_1, Z_4)$.

Specific potential confounders

Note that the observational-unit variable U and the U-conditional treatment probability $P(X{=}x \mid U)$ are potential confounders as well. (The values of $P(X{=}x \mid U)$ are the conditional probabilities $P(X{=}x \mid U{=}u)$), which are attributes of the persons u [see Eq. (1.1)]. Similarly, the Z-conditional treatment probability $P(X{=}x \mid Z)$ is also a potential confounder provided that Z is a covariate [see Def. 4.4 (i) and SN-chapters 2 and 10]. Furthermore, the *assignment* to treatment x with values '*yes*' and '*no*' is also a potential confounder if *assignment to treatment* and *exposure to treatment* (again with values '*yes*' and '*no*') are not identical and *exposure to treatment* is focused as a (putative) cause. This distinction is useful in experiments with non-compliance (see, e. g., Jo, 2002a, 2002b, 2002c; Jo et al., 2008).

Unobserved potential confounders

Even if we consider a multivariate potential confounder Z consisting of several univariate potential confounders, there are always unobserved variables that are prior or simultaneous to treatment. Such variables are called *unobserved potential confounders*. Sometimes they are also called *hidden confounders* (cf., e. g., Rosenbaum, 2002a). Of course such an unobserved potential confounder may bias the conditional expectation values of the outcome variable just in the same way as an observed covariate. Whether or not the conditional expectation values of the outcome variable in the treatment conditions are unbiased such that their differences represent causal total effects does not only depend on the relationship between the observed variables such as X, Y, and the observed (possible multivariate) covariate, say Z, but also on the relationship of these variables to the unobserved potential confounders. In other words, potential confounders exert their maleficent effects irrespective of whether or not we observe them.

Outcome variable

Of course, the outcome variable Y refers to a time at which the treatment might have had its impact. Hence, treatment variables are always prior to the outcome variable. In principle, we may also observe several outcome variables, for example, in order to study how the effects of a treatment grow or decline over time or to study treatment effects that are not confined to a single outcome variable. All random variables mentioned above refer to a concrete single-unit trial and they have a joint distribution. Each combination of unit, treatment condition, and score of the outcome variable may be an observed result of such a single-unit trial. This implies that the variables U, Z, X, and Y, as well as unobserved potential confounders, say W, have a joint distribution (see, e. g., SN-section 5.3). Once

we specified the random experiment to be studied, this joint distribution is fixed, even though it might be known only in parts or even be unknown altogether.

Causal effects and causal dependencies

There is already a plenitude of different kinds of causal effects that can be considered in the single-unit trial of a simple experiment or quasi-experiment. For simplicity, suppose the treatment has just two values, say *treatment* and *control*. *First*, there is the *causal average total effect* of treatment (compared to control) on the outcome variable Y. *Second*, there are the *causal conditional total treatment effects* on Y, where we may condition on any function of the observational-unit variable U. If, for example, $Z :=$ *sex* with values m for *male* and f for *female*, then we may consider the causal $(Z = m)$-conditional total treatment effect on Y, that is, the causal average total treatment effect for males, and the causal $(Z = f)$-conditional total treatment effect on Y, that is, the causal average total treatment effect for females. Similarly, if $Z :=$ *socio-economical status*, we may consider the causal conditional total treatment effects on Y for each status group, etc. *Third*, although difficult and often impossible to estimate, we may also consider the causal *individual total effect* of *treatment* compared to *control* on Y.

By definition, within a *simple* experiment and quasi-experiment we cannot consider any *direct* treatment effects with respect to a one or more specified intermediate variables, that is, the effects of the treatment on the outcome variable that *are not* transmitted through specified intermediate variables. However, the causal total treatment effects discussed above are, of course, transmitted through intermediate variables, irrespective of whether or not we observe (or are aware of) these intermediate variables.

2.2 Experiments With Fallible Covariates

Another class of random experiments are single-unit trials of experiments and quasi-experiments in which we assess a fallible covariate. In this case, the fallible covariate does *not* represent a (deterministic) attribute of the observational units. The single-unit trial of such an experiment of quasi-experiment consists of:

(a) sampling an observational unit u (e. g., a person) from a set of units,
(b) assessing the values $z_1, \ldots, z_k$ of the covariates (pre-treatment variables) $Z_1, \ldots, Z_k$, $k \geq 1$.
(c) assigning the unit or observing its assignment to one of several experimental conditions (represented by the value x of the treatment variable X),
(d) recording the value y of the outcome variable Y.

The crucial distinction between a simple (quasi-) experiment and a (quasi-) experiment *with fallible covariates* is that there is variability of at least one of the covariates *given the observational unit* u (see Fig. 2.2). In this case, we may distinguish between the *latent covariate*, say ξ, representing the attribute to be assessed and its *fallible measures*, some manifest variables that can actually be observed. (For a theory of latent variables see Steyer et al., 2015). Also note that sometimes it is crucial to adjust the effect of X on Y by conditioning on the *latent variable* ξ in order to fully adjust for the bias of the prima-facie effect

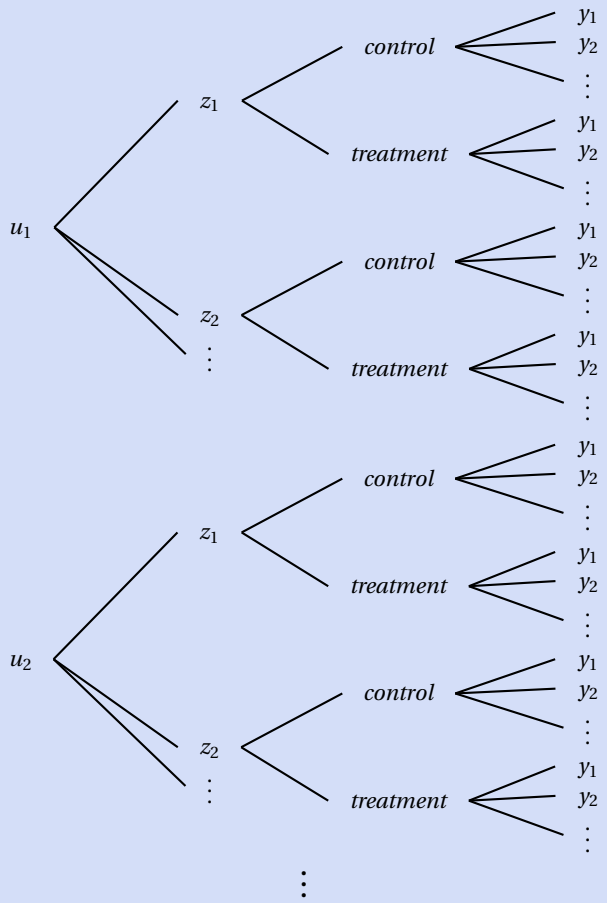

Figure 2.2. Experiment or quasi-experiment with a fallible covariate

of X on Y. In these cases, only adjusting for the manifest variables that measure the latent covariate ξ does not completely remove bias (see, e. g., ?, ? RST).

Furthermore, this distinction also implies that the *unit* whose attributes are measured *at the time when the potential confounder is assessed* is not identical any more with the *unit at the onset of treatment* (see section 2.1). The covariate might be assessed some months before the treatment is given — enough time and plenty of possibilities for the unit to change in various ways, for example, due to maturation, learning, critical life events, and other experiences that are not fixed yet at the time of assessing the covariate. As a consequence, a variable, say W, representing such intermediate events or experiences may also affect the outcome variable Y and the treatment variable. Hence, such intermediate variables are also potential confounders. This is one of the reasons why we need to define causal effects in a more general way than in the Neyman-Rubin tradition (see chs. 4 and 5).

Note that assessing a fallible covariate does not only change the interpretation of the observational-unit variable U (now its values are the units of the time of assessment of the manifest covariates), but it also changes the random experiment, and with it, the empirical phenomenon we are considering. Assessing a fallible covariate often involves that the sampled person fills in a questionnaire or takes a test. Assessing, prior to treatment, a fallible covariate such as a test of an ability, an attitude, or a personality trait, may change the observational units and their attributes, as well as the effects of the treatment on a specified outcome variable, which usually is related to such pre-treatment variables. This has already been discussed by Campbell and Stanley (1963), who also recommended designs for studying how pre-treatment assessment modifies the effects of the treatment variable on the outcome variable.

Potential confounders and covariates

Which are the potential confounders of the treatment variable X in such a single-unit trial? First of all, it is each attribute of the units at the time of the assessment of the observed covariates. This does not only include variables such as *sex*, *race*, and *educational status*, but also a latent covariate, say ξ, (which might be multi-dimensional). Furthermore, aside from the manifest covariates, each variable W representing an intermediate event or experience of the unit (occurring in between the assessment of the observed covariates and the onset of the treatment), as well as any attribute of the *unit at the onset of treatment* is a potential confounder as well, irrespective of whether or not these potential confounders are observed.

RST drin lassen nur wenn spaeter Beispiele kommen

Note that a latent covariate ξ may be considered a cause of its fallible measures $Z_1, \ldots, Z_k$ but also of the outcome variable Y. This is not in conflict with the theory that the treatment variable X is a cause of Y as well. In this kind of single-unit trial, we have several causes and several outcome variables, and a cause itself can be considered as an outcome variable. For example, it would be possible to consider the treatment variable X to be causally dependent on the manifest or latent covariates. In other words, we may also raise the question if the treatment probabilities $P(X=1|Z_1,\ldots,Z_k)$ or $P(X=1|\xi)$ describe causal dependencies. This makes clear that the terms potential confounder and 'covariate' can only be defined with respect to a focused cause.

RST Ende

2.3 Two-Factorial Experiments

As a third class of random experiments we consider two-factorial experiments. The single-unit trial of such a two-factorial experiment or quasi-experiment consists of:

(a) sampling an observational unit u (e. g., a person) from a set of units,
(b) assigning the unit or observing its assignment to one of several experimental conditions that are defined by the pair (x,z) of levels of two treatment variables X and Z, respectively.
(c) recording the value y of the outcome variable Y.

Sampling a unit

Because we presume that no fallible potential confounders such as 'severity of symptoms', 'motivation for treatment', etc. are assessed before treatment, sampling an observational unit means that we are sampling a *unit at the onset of treatment*.

Treatment variables

As a simple example, let us consider an experiment in which we study the effects — including the joint effects — of two treatment factors, say *individual therapy* represented by X (with values '*yes*' and '*no*') and *group therapy* represented by Z (with values '*yes*' and '*no*').

In such a two-factorial experiment, we may consider *group therapy* as a covariate and *individual therapy* to be the cause in order to ask for the conditional and average total effects of individual therapy given group therapy. In contrast, we may also consider *individual therapy* to be a covariate and *group therapy* to be the focused treatment variable. Finally, we may also consider the two-dimensional variable (X,Z) as the cause. Which option is chosen depends on the causal effects we are interested in (see below).

Outcome variable

Again, the outcome variable Y refers to a time at which the treatment might have exerted the effect to be estimated. Hence, both treatment variables are prior to the outcome variable considered. And again, we may also observe several outcome variables, for example, in order to study how effects of a treatment grow or decline over time or to study effects that are not confined to a single outcome variable.

Causal effects

There are several causal effects we might look at. If X and Z have only two values, then we may be interested in the following effects on the outcome variable Y:

- (a_1) the conditional total effect of 'individual therapy' as compared to 'no individual therapy' given that the unit treated also receives 'group therapy',
- (b_1) the corresponding conditional total effect given that the unit does *not* receive 'group therapy', and
- (c_1) the average of these conditional total effects of 'individual therapy' as compared to 'no individual therapy', averaging over the two values of Z.

Vice versa, we might also be interested in the following effects on the outcome variable Y:

- (a_2) the conditional total effect of 'group therapy' as compared to 'no group therapy' given that the unit treated also receives 'individual therapy',
- (b_2) the corresponding conditional total effect given that the unit does *not* receive 'individual therapy', and
- (c_2) the average of these conditional total effects of 'group therapy' as compared to 'no group therapy', averaging over the two values of X.

Furthermore, there are other causal effects on Y we might study, namely

(a_3) the total effect of receiving 'individual therapy' *and* 'group therapy' as compared to receiving none of the two treatments.

(b_3) the total effect of receiving 'individual therapy' *and* 'no group therapy' as compared to receiving 'group therapy' *and* 'no individual therapy'.

All these effects may answer meaningful causal questions. In fact there are even more causal effects than those listed above. For example, we could compare each of the four combinations of the two treatments to an average of the other treatments. Furthermore, many additional causal effects can be considered if we condition on other covariates such as *sex* or *educational status*.

Potential confounders and covariates

If we focus on the effect of X (individual therapy), then we consider Z (group therapy) as a covariate of X. In contrast, we treat X as a covariate of Z if we study the effects of Z (group therapy). Furthermore, in both cases, each attribute of the unit at the onset of treatment (such as *sex* or *educational status*) could be considered as covariates as well. Assessing these covariates does not appear in points (a) to (c) of the random experiment, because these covariates are (deterministic) functions of the observational-unit variable. Therefore, there is no additional sampling process associated with their assessment.

This is also true for other potential confounders, for example, variables characterizing the situation in which the unit is at the onset of treatment, the number of *hours slept* last night, or *day time* at which the unit receives its treatment. Even variables that characterize early experiences in the childhood of the unit such as a *broken home* or *mother's child care behavior* are potential confounders in this single-unit trial. They are there and exert their effects even if they are not assessed.

Note again that assessment of these potential confounders in a questionnaire filled in by the person constitutes a new random experiment that may differ in important ways from a random experiment in which the unit has no such task (see section 2.2). In psychology, an assessment often is a treatment of its own.

2.4 Multilevel Experiments

In multilevel experiments and quasi-experiments we also study the effect of a treatment on an outcome variable. However, in such a design the observational units are nested within higher hierarchical units referred to as *clusters*. Examples include experiments, in which students are nested within classrooms, patients are nested within clinics, and inhabitants are nested in cities and neighborhoods. Multilevel designs can be classified as designs with treatment assignment at the unit-level or at the cluster-level. Furthermore, multilevel designs differ with respect to the assignment of units to clusters. There are designs with pre-existing clusters and there are designs with assignment of units to clusters. All these designs involve different single-unit trials.

A single-unit trial with *pre-existing clusters* consists of:

(a) sampling a cluster c (e. g., a school class, a neighborhood, or a hospital) from a set of clusters,

(b) sampling an observational unit u (e. g., a person) from a set of units within the cluster,

(c) assigning the unit or the cluster (depending on the design) or observing their assignment to one of several experimental conditions (represented by the value x of the treatment variable X),

(d) recording the value y of the outcome variable Y.

In contrast, a single-unit trial *with assignment of units to clusters* consists of:

(a) sampling an observational unit u (e. g., a person) from a set of units,

(b) assigning the unit or observing its assignment to one of several clusters (represented by the value c of the cluster variable C),

(c) assigning the unit or the cluster (depending on the design) or observing their assignment to one of several experimental conditions (represented by the value x of the treatment variable X),

(d) recording the value y of the outcome variable Y.

In the experiment with pre-existing clusters, each unit can only appear in one cluster, whereas in the experiment with assignment of units to a cluster, each unit can appear in more than one cluster. Note again, that we are considering single-units trials from the pre-factual perspective, not from the post-factual or 'counter-factual' perspective (see the remarks following the description of the random experiment presented in Table 1.1). Hence, in experiments with assignment of units to a cluster, the cluster variable can bias the dependency of the outcome variable on the treatment variable *on the level of the observational unit*. In this aspect this design resembles the multifactorial design described in section 2.3.

Potential confounders and covariates

Which are the potential confounders in multilevel designs if the treatment variable X is considered as the cause? The answer depends on the type of design considered: In designs with treatment assignment of units to clusters, attributes of the observational unit such as *sex*, *race*, or *educational status*, are potential confounders of X. Other potential confounders are attributes of the cluster such as *school type*, *hospital ownership*, or *school-level socio-economic status* or *school-level intelligence*. The last two kinds of potential confounders would be defined as conditional expectations of the corresponding potential confounders at the unit-level given the cluster variable.

In these designs, clusters may not only be considered as potential confounders, but also as treatments, because some of the effects observed later on may depend on the composition of the group to which a particular unit, say Joe, is assigned. Receiving group therapy together with beautiful Ann in the same group might make a great difference as compared to getting it together with awful Joe. In designs in which clusters as a whole are assigned to treatment conditions, only attributes of the cluster can influence the assignment. Hence, in data analysis we would focus on controlling for the potential confounders on the cluster level (see, e. g., Nagengast, 2009, for more details).

2.5 Experiments With Latent Outcome Variables

We may also consider single-unit trials of experiments with a *latent outcome variable*. The basic goal of such experiments is to investigate the effect of the treatment variable X on a *latent* outcome variable, say η. This is of interest, for example, where a quantitative outcome variable can only be measured by qualitative observations such as solving or not solving certain items indicating the (latent) ability. However, it can also be of interest if the manifest measures are *linearly* related to the latent variable such as in models of classical test theory (see, e. g., Steyer, 2001) or in models of latent state-trait theory (see, e. g., Steyer et al., 2015). If, for example, there are three manifest variables Y_1, Y_2, and Y_3 measuring a single latent variable η, we may ask if there is just one single effect of the treatment on the latent outcome variable η – which transmits these effects to the manifest variables Y_1, Y_2, and Y_3 – instead of three separate effects of X on each variable Y_i. Hence, the latent variable may also be considered to be a mediator variable. Showing that all effects of X on the variables Y_i are indirect, that is, mediated by η is one of the research efforts that aims at establishing construct validity of the latent variable η.

In the simplest case with a single latent variable, we consider the following single-unit trial:

(a) Sampling a person u out of a set of persons,
(b) assigning the unit or observing its assignment to one of several experimental conditions (represented by the value x of the treatment variable X),
(c) recording the values $y_1, \ldots, y_m$ of the manifest outcome variables $Y_1, \ldots, Y_m$.

In this single-unit trial, the values u of the observational-unit variable U again represent the observational *unit at the onset of treatment*, while the latent outcome variable η represents some attribute of the unit at the time point at which the outcome of the treatment is assessed. Clearly, this time point is *after* treatment and *prior* to the observation of the manifest outcome variables Y_i, at least as long as we preclude change in the latent variable during the process of assessing the manifest outcome variables. If this cannot be precluded, we would have to consider the time sequence in assessing the manifest outcome variables (e. g., of the items to be solved or answered) as well.

Potential confounderes and covariates

Which are the potential confounders in such a single-unit trial? Again, the answer depends on the cause considered. If it is the treatment variable X, then each attribute of the *unit at the onset of treatment* is a potential confounder (with respect to X). Obviously, this again includes variables such as *sex*, *race*, and *educational status*. Note that in this kind of experiments, the set of potential confounders of X is the same irrespective of the choice of the outcome variable. Remember, we may not only consider the *latent* outcome variable η but also the *manifest* outcome variables Y_i, for example, in order to study whether or not the effects of X on these manifest outcome variables are perfectly transmitted (or mediated) through the latent variable η.

Choosing the latent outcome variable η as a cause of the manifest outcomes variables Y_i brings additional potential confounders into play, for instance, all those variables that are *in between* treatment and the assessment of η. If, for example, we consider an experiment studying the effects of different teaching methods, these additional potential con-

Box 2.1 Glossary of new concepts

Random experiment	The kind of empirical phenomenon to which events, random variables, and their dependencies refer.
Single-unit trial	A particular kind of random experiment that consists of sampling a single unit from a set of observational units and observing the values of one or more random variables related to this unit.
Cause	A random variable. Its effect on an outcome variable is considered.
Outcome variable	A random variable. Its dependency on a cause is considered.
Potential confounder	If we confine the discussion to total causal effects, then it is a random variable that is prior or simultaneous to the cause considered. It might be correlated with the cause and the outcome variable.
Covariate	A potential confounder that is considered together with X in a conditional expectation or a conditional distribution.
Fallible covariate	A covariate that is assessed with measurement error.
Latent covariate	A covariate that is not directly observed. Instead it is defined using some parameters of the joint distribution of a set of manifest random variables.
Intermediate variable	A variable that might mediate (transmit) the effect of the cause on the outcome variable. The cause is always prior to an intermediate variable and an intermediate variable is always prior to the outcome variable. An intermediate variable is not *necessarily* affected by the cause and it does not *necessarily* have an effect on the outcome variable.
Mediator	An intermediate variable on which X has a causal effect and which itself has a causal effect on the outcome variable Y.

Note that all these terms are still of an informal nature. Their mathematical treatment starts in chapter 4.

founders are critical life events (such as father or mother leaving the family), or additional lessons taken after treatment and before outcome assessment, for instance.

2.6 Summary and Conclusion

In this chapter we described a number of random experiments in informal terms. The purpose was to get a first idea which kind of empirical phenomena causal theories and hypotheses refer to. We focused on single-unit trials, which are the kinds of empirical phenomena we are interested in, both in theory and practice. We emphasized that a single-unit trial is a random experiment and discussed several kinds of random variables playing

a crucial role in the theory of causal effects. We also mentioned that there is a certain *time order* among these random variables, for example, saying that the potential confounders are 'prior' or 'simultaneous' to the treatment variable, which itself is 'prior' to the outcome variable. Furthermore, for each single-unit trial and each cause in such a single-unit trial, we discussed the *potential confounders* involved. We emphasized that each cause considered in such a single-unit trial has its own set of potential confounders.

Other single-unit trials

The single-unit trials discussed in this chapter are just a small selection of single-unit trials in which causal effects and causality of stochastic dependencies are of interest. We might also consider single-unit trials with latent covariates *and* latent outcome variables *and* manifest and/or latent intermediate variables, but also single-unit trials with multiple mediation. Furthermore, we could also consider single-unit trials of growth curve models (see, e. g., Biesanz, Deeb-Sossa, Aubrecht, Bollen, & Curran, 2004; Bollen & Curran, 2006; Meredith & Tisak, 1990; Singer & Willett, 2003; Tisak & Tisak, 2000), latent change models (see, e. g., McArdle, 2001; Steyer, Eid, & Schwenkmezger, 1997; Steyer, 2005), or cross-lagged panel models (see, e. g., Kenny, 1975; Rogosa, 1980b; Watkins, Lei, & Canivez, 2007; Wolf, Chandler, & Spies, 1981). Causality is also an issue in uni- and multivariate time-series analysis as well as in stochastic processes with continuous time. However, in this book our examples will usually deal with experiments and quasi-experiments, including latent covariates and outcome variables.

Outlook

Steyer and Nagel (2017) treat the fundamental concepts related to probability, conditional expectations, conditional independence, and conditional distributions. In chapter 4 we introduce additional mathematic concepts that allow us to meaningfully talk about time order between events and random variables. There, we also define concepts such as *potential confounders*. This will provide the mathematical framework and the language in which causal effects can meaningfully be discussed.

2.7 Exercises

▷ **Exercise 2-1** Imagine that the probabilities of a crash for a flight with Airline A is ten times smaller than with Airline B. Which airline would you choose?

▷ **Exercise 2-2** Why does the theory of causal effects refer to single-unit trials?

▷ **Exercise 2-3** Why is it important to know which random experiment we are talking about?

▷ **Exercise 2-4** Which type of random experiment did we refer to in the two examples described in chapter 1?

▷ **Exercise 2-5** Why is it important to emphasize that, in simple experiments and quasi-experiments (see section 2.1), the observational-unit variable U represents the observational units *at the onset of treatment*?

▷ **Exercise 2-6** What is the basic idea of a potential confounder of a cause?

▷ **Exercise 2-7** Which kinds of causal effects can be considered in the simple experiment or quasi-experiment in which no *fallible* potential confounder and no intermediate variable is assessed?

Solutions

▷ **Solution 2-1** If your answer is A, then you implicitly apply these probabilities to the random experiment of flying *once* with A or B, even if these probabilities have been estimated in a sample. This example serves to emphasize that, not only in theory but also in practice, we are mainly interested in a single-unit trial, not in a sample consisting of many such single-unit trials, and in particular not in what applies to sample size going to infinity. (This is how many applied statisticians try to specify the term 'population'.)

▷ **Solution 2-2** Within such a single-unit trial, the various concepts of causal effects can be defined and we can study how to identify these causal effects from the parameters describing the joint distribution of the random variables considered. In such a single-unit trial, there usually is a clear time order which helps (but is not sufficient) to disentangle the possible causal relationships between the random variables considered.

▷ **Solution 2-3** Different random experiments are different empirical phenomena. Although the names of the variables in different random experiments might be the same, the variables themselves are different entities, implying that the dependencies and effects between these variables might be different in different random experiments.

▷ **Solution 2-4** The type of random experiment we refer to in these examples is the single-unit trial of simple experiments and quasi-experiments described in section 2.1.

▷ **Solution 2-5** In the social sciences, units are often persons, and persons can change over time. If, in a simple experiment or quasi-experiment, a value u of U represents the observational unit sampled *at the onset of treatment*, each potential confounder is a function of U. If, in contrast, U represents the *observational unit at the assessment of a fallible covariate* (see section 2.2), which is some time prior to the onset of treatment, then there can be other potential confounders in between assessment of the fallible potential confounder and the onset of treatment. We have to consider these additional potential confounders both in the definition of causal effects and in data analysis.

▷ **Solution 2-6** A potential confounder of a cause is a random variable that is prior or simultaneous to the cause, at least as long as we only consider total effects. (If we also consider direct effects, then a potential confounder can also be posterior to a cause.)

▷ **Solution 2-7** If the treatment has just two values, say *treatment* and *control*, there are different kinds of causal effects of the treatment variable on the outcome variable Y, such as the average total treatment effect, the conditional total treatment effects given a value of a covariate Z, and the *individual total effect* of X on Y given an observational unit u. Aside from these treatment effects, we may also consider the causal effects of a potential confounder Z on the treatment variable X, but also on the outcome variable Y.

Chapter 3

Probability and Conditional Expectation

In chapter 1 it has been shown that the conditional expectation values $E(Y|X=x)$ of an outcome variable Y and their differences $E(Y|X=x) - E(Y|X=x')$, the *prima facie effects*, can be misleading in evaluating the causal total effect of a (treatment) variable X on an (outcome) variable Y. There we treated a number of other effects that might be useful for the evaluation of a treatment (see Box 1.1). In chapter 2, we described random experiments of various research designs in which a causal total effect is of interest. These chapters hopefully motivated the reader to make himself familiar with the most important concepts of probability theory. Aside from the fundamental concepts such as σ-algebra , probability measure, probability space, random variable, expectation, variance, covariance, etc., Steyer and Nagel (2017) extensively treat the concepts that can be used to describe stochastic dependencies, including conditional expectation values, conditional expectations, and conditional distributions. This book includes many theorems as well as other propositions and their proofs. It will often be used and referred to by SN followed by Definition, Equation, Theorem, Remark, etc.

In this chapter we review the most important of these concepts of probability theory, for two purposes. The first is to refresh them so that the reader has them at hand when they are used in this book. The second purpose is to sensitize the reader to the fact that the traditional probabilistic concepts alone cannot be used offhandedly for describing the causal effects in which we are interested when we evaluate a treatment, an intervention, or an exposition. We start with the components of a probability space that can be used for a mathematical representation of a concrete random experiment. Then we define the concepts of a random variable and its distribution. Finally, we turn to concepts that can be used to describe stochastic dependencies among random variables, including conditional expectation values and conditional expectations. Hence, now, among others, we belatedly introduce the concepts already used in chapter 1. All concepts presented in this chapter are treated in more detail in Steyer and Nagel (2017). They are just selected for the purpose of this book and illustrated by some new examples.

3.1 Probability Space

In this section we introduce the three components of a probability space, a *set of possible outcomes* (of the random experiment considered), a *set of possible events* to be considered, and a *probability measure* that assigns a probability to each of these possible events. All three components have a certain mathematical structure.

Table 3.1. Joe and Ann with randomized assignment

Outcomes ω_i						Observables			Conditional expectations				
Unit	Treatment	Success	$P(\{\omega_i\})$	$P^{X=0}(\{\omega_i\})$	$P^{X=1}(\{\omega_i\}) = P^B(\{\omega_i\})$	Person variable U	Treatment variable X	Outcome variable Y	$P(Y=1\|X,U) = E(Y\|X,U)$	$P(Y=1\|X) = E(Y\|X)$	$P(X=1\|U) = E(X\|U)$	$E^{X=0}(Y\|U)$	$E^{X=1}(Y\|U)$
ω_1 = (*Joe*, *no*, −)			.09	.15	0	*Joe*	0	0	.7	.45	.4	.7	.8
ω_2 = (*Joe*, *no*, +)			.21	.35	0	*Joe*	0	1	.7	.45	.4	.7	.8
ω_3 = (*Joe*, *yes*, −)			.04	0	.1	*Joe*	1	0	.8	.6	.4	.7	.8
ω_4 = (*Joe*, *yes*, +)			.16	0	.4	*Joe*	1	1	.8	.6	.4	.7	.8
ω_5 = (*Ann*, *no*, −)			.24	.4	0	*Ann*	0	0	.2	.45	.4	.2	.4
ω_6 = (*Ann*, *no*, +)			.06	.1	0	*Ann*	0	1	.2	.45	.4	.2	.4
ω_7 = (*Ann*, *yes*, −)			.12	0	.3	*Ann*	1	0	.4	.6	.4	.2	.4
ω_8 = (*Ann*, *yes*, +)			.08	0	.2	*Ann*	1	1	.4	.6	.4	.2	.4

Note. The probabilities of the elementary events are fictive. The terms illustrated in this table are introduced in the course of this chapter. The event B occurring in the third column denotes the event that the drawn person is treated (see Example 3.18).

3.1.1 Set of Possible Outcomes

The *set of possible outcomes* (of a random experiment), also called *sample space*, is the first component of a probability space. If we think about a specific random experiment, then we should at least know its set of possible outcomes. The mathematical structure of this component is simply the structure of a set. This means that we have to know its elements, the possible outcomes. We use the term *possible* outcomes in order to communicate and keep in mind that we are always talking about a random experiment from the *pre-factual perspective*, that is, from the perspective *before* it is actually conducted. Even if a random experiment is already executed, then, talking about the probabilities of certain events, we do *as if* it has not yet been conducted. Only in this way does it make sense to talk about the *probability* of an event.

Example 3.1 (Joe and Ann With Randomized Assignment) A first example is presented in Table 3.1. The first column in this table contains the eight elements of the set Ω of possible outcomes of the random experiment considered. Hence, in this example,

$$\Omega = \{\omega_1, \ldots, \omega_8\} = \big\{ (\textit{Joe}, \textit{no}, -),\ (\textit{Joe}, \textit{no}, +), \ldots,\ (\textit{Ann}, \textit{yes}, +) \big\}, \tag{3.1}$$

that is, Ω consists of the eight triples (*Joe*, *no*, −), (*Joe*, *no*, +), …, (*Ann*, *yes*, +). If we define $\Omega_U = \{\textit{Joe}, \textit{Ann}\}$, $\Omega_X = \{\textit{no}, \textit{yes}\}$, and $\Omega_Y = \{-, +\}$, then Ω can also be written as the Cartesian product

$$\Omega = \Omega_U \times \Omega_X \times \Omega_Y. \tag{3.2}$$

In this example, the set Ω of possible outcomes has eight elements. $\triangleleft$

Example 3.2 (Nonorthogonal Two-Factorial Experiment) Another example has already been presented in Table 1.5. If we define the sets

$$\Omega_U = \{Tom, Tim, Joe, Jim, Ann, Eva, Sue, Mia\},$$

$$\Omega_X = \{control, treatment\,1, treatment\,2\},$$

and

$$\Omega_Y = \mathbb{R},$$

then Ω is the Cartesian product of these three sets, that is,

$$\Omega = \Omega_U \times \Omega_X \times \Omega_Y.$$

In contrast to Example 3.1, now the set Ω has an uncountable number of elements, which would be necessary if we want to consider a response variable Y that can take on as a value any real number, representing, for example, the degree of success. This is also the reason why we choose the compressed form in Table 1.5, which does not show the possible values of Y but only its $(X{=}x, U{=}u)$-conditional expectation values [see Eq. (3.27) and Def. 3.66]. Note that in many empirical applications in psychology or the social sciences, Ω_Y may just be a subset of $\mathbb{R}$. $\triangleleft$

3.1.2 Set of Possible Events

The second component of a probability space is a *set of possible events*. It is a set of subsets of the set Ω of possible outcomes. This set of subsets of Ω has the properties of a σ-*algebra* . An element of a σ-algebra is called a *measurable set*, or an *event*, if there is a probability measure on this σ-algebra. We use A^c to denote the *complement* of A, that is, $A^c := \Omega \setminus A$, where $\Omega \setminus A$ consists of all elements of Ω that are not elements of A.

Definition 3.3 (σ-Algebra and Measurable Space)
A set $\mathscr{A}$ of subsets of a nonempty set Ω is called a σ*-algebra (or* σ*-field) on Ω, if the following three conditions hold:*

(a) $\Omega \in \mathscr{A}$.
(b) If $A \in \mathscr{A}$, *then* $A^c \in \mathscr{A}$.
(c) If $A_1, A_2, \ldots \in \mathscr{A}$, *then* $\bigcup_{i=1}^{\infty} A_i \in \mathscr{A}$.

An element A of a σ-algebra $\mathscr{A}$ is called a measurable set. The pair $(\Omega, \mathscr{A})$ is called a measurable space.

Remark 3.4 (Closure With Respect to Set Operations) Condition (c) means that a σ-algebra is closed with respect to *countable* unions of sets $A_1, A_2, \ldots \in \mathscr{A}$. However, in conjunction with (a) and (b) this implies that a σ-algebra is also closed with respect to *finite* unions of sets $A_1, \ldots, A_n \in \mathscr{A}$. That is, if $A_1, \ldots, A_n \in \mathscr{A}$, then $A_1 \cup \ldots \cup A_n \in \mathscr{A}$. Furthermore, although condition (c) explicitly requires only that σ-algebras are closed with respect to countable unions, Definition 3.3 implies that a σ-algebra is closed also with respect to intersections such as $A_1 \cap A_2$ and set differences $A_1 \setminus A_2$. In other words, if A_1 and A_2 are

elements of $\mathscr{A}$, then $A_1 \cup A_2$, $A_1 \cap A_2$, and $A_1 \setminus A_2$ are elements of $\mathscr{A}$ as well, provided that $\mathscr{A}$ is a σ-algebra. The same is true for countable intersections $A_1 \cap A_2 \cap \ldots$ of elements of $\mathscr{A}$. (For more details see SN-section 1.2.) ◁

Example 3.5 (Joe and Ann With Randomized Assignment) We may define $\mathscr{A} := \mathscr{P}(\Omega)$ to be the power set, that is, the set of all subsets of the set Ω that has been specified in Equation (3.1). The power set of the set Ω is always a σ-algebra on Ω. In this example, the power set contains $2^8 = 256$ elements. Other σ-algebras on Ω are the trivial σ-algebra $\{\Omega, \emptyset\}$ and $\{A, A^c, \Omega, \emptyset\}$, provided that $A \subset \Omega$ (see Exercises 3-1 and 3-2). ◁

Example 3.6 (Nonorthogonal Two-Factorial Experiment) In this example we cannot use the power set of Ω as a σ-algebra, because this would lead to contradictions (see SN-Rem. 1.8). Instead we use the *product of the* σ*-algebras* $\mathscr{A}_U = \mathscr{P}(\Omega_U)$, $\mathscr{A}_X = \mathscr{P}(\Omega_X)$, and the *Borel* σ*-algebra* $\mathscr{B}$ on $\mathbb{R}$, which contains as elements all singletons $\{\alpha\}$, $\alpha \in \mathbb{R}$, as well as all (open, half open, and closed) intervals, their countable unions and intersections (for the definition of these concepts and more details see SN-section 1.2.2 and SN-Def. 1.31). ◁

3.1.3 Probability Measure

The last component of a probability space is a *probability measure*, which assigns a probability to each element of a σ-algebra. This concept has beed introduced by Kolmogorov (1933/1977) (for the English version of this book see Kolmogorov, 1956). In the following definition we use $[0,1]$ to denote the *closed interval* of the real numbers between 0 and 1, inclusively.

Definition 3.7 (Probability Measure)
Let $(\Omega, \mathscr{A})$ *be a measurable space. Then the function* $P\colon \mathscr{A} \to [0,1]$ *is called a probability measure on* $(\Omega, \mathscr{A})$*, if the following conditions hold:*

(a) $P(\Omega) = 1$ *(standardization).*
(b) $P(A) \geq 0,\ \forall\, A \in \mathscr{A}$ *(nonnegativity).*
(c) $A_1, A_2, \ldots \in \mathscr{A}$ *are pairwise disjoint* $\Rightarrow P\Big(\bigcup_{i=1}^{\infty} A_i\Big) = \sum_{i=1}^{\infty} P(A_i)$ *(*σ*-additivity).*

Remark 3.8 (Probability of an Event and Probability Space) Let P be a probability measure on $(\Omega, \mathscr{A})$. Then the triple $(\Omega, \mathscr{A}, P)$ is called a *probability space* and a value $P(A)$ of P is called the *probability* of (the event) A. If a probability space is used to represent a random experiment, then it contains all information about this random experiment, that is, everything we can ever learn about this random experiment can be computed from the probabilities $P(A)$, $A \in \mathscr{A}$. The most important properties of a probability measure are gathered in SN-Box 4.1. (For the general concept of a measure and its properties as well as for other measures than probability measures see SN-chapter 1.) ◁

Remark 3.9 (Elementary Event and Event) Let $(\Omega, \mathscr{A}, P)$ be a probability space. Then A is called a (possible) *event*, if $A \in \mathscr{A}$ and a set $\{\omega\}$ is called a (possible) *elementary event* if $\omega \in \Omega$ and $\{\omega\} \in \mathscr{A}$. Note the distinction between an *outcome* $\omega \in \Omega$ and an *elementary event*

$\{\omega\} \in \mathscr{A}$. Also note that the term *event* is only used in the context of a probability space $(\Omega, \mathscr{A}, P)$. Otherwise $A \in \mathscr{A}$ is called a *measurable set*. For simplicity we often drop the long-winded term 'possible' if we talk about outcomes and events. Nevertheless, in applications we continue considering a random experiment from the pre-factual perspective. ◁

Remark 3.10 (Probability Space and Random Experiment) Note the disctinction between a *probability space* and a *random experiment*. The term probability space is a mathematical concept. It does not have any empirical meaning unless we interpret Ω by saying that it represents the set of possible outcomes of a concrete random experiment. In contrast, 'random experiment' is a term of our colloquial language referring to an empirical phenomenom that we might be interested in. Giving Ω a concrete empirical meaning, the σ-algebra $\mathscr{A}$ also obtains a concrete empirical meaning: the set of possible events A to be considered *in this concrete random experiment*. Correspondingly, the probability measure P assigns to each of these possible events A their probabilities $P(A)$ that refer *to this specific random experiment*. Hence, propositions about probabilities refer to a concrete random experiment and cannot readily be generalized to other random experiments. ◁

Example 3.11 (Joe and Ann With Randomized Assignment) The second column of Table 3.1 displays the assignment of the probabilities to all eight elementary events. In this example, each nonempty element $A \in \mathscr{A}$ is a union of the elementary events $\{\omega_i\}$, $\omega_i \in \Omega$, and because a measure is additive, the probability measure $P\colon \mathscr{A} \to [0,1]$ is uniquely defined by the second column of Table 3.1 [see SN-Box 4.1 (x)]. The probabilities of all other events that can be considered in this random experiment can be computed from the probabilities of the eight elementary events. Hence, because Ω has been fixed in Example 3.1 and $\mathscr{A}$ in Example 3.5, now the probability space $(\Omega, \mathscr{A}, P)$ is completely specified. Note that the probabilities of the elementary events differ from the example presented in Table 1.2. Now the probabilities may be interpreted to describe an experiment with randomized assignment of persons to treatment conditions. ◁

Example 3.12 (Nonorthogonal Two-Factorial Experiment) In this example, specifying the probability measure is more difficult than in Example 3.11. What has been fixed in the third column of Table 1.5 are the probabilities of the events $\{U=u\} := \{\omega \in \Omega\colon U(\omega) = u\}$. In the fourth column of that table we also specified the conditional probabilities $P(X=1 \mid U=u)$ of the event $\{X=1\} := \{\omega \in \Omega\colon X(\omega) = 1\}$ given the event $\{U=u\}$. Furthermore, we also fixed the conditional expectation values $E(Y \mid X=x, U=u)$ for all pairs (x,u) of values X and U. What has not been specified in that table are the $(X=x, U=u)$-conditional distributions of Y. If we assume that Y is continuous and has a normal distribution with identical conditional variances $Var(Y \mid X=x, U=u) = \sigma^2$ given each pair (x,u) of values X and U, then the probability measure P on $(\Omega, \mathscr{A})$ is completely specified. (The crucial arguments are SN-Theorem 17.37 and the fact that the normal distribution is completely determined by the expectation and the variance of the random variable considered.) Instead of assuming identical conditional variances, we may also specify different conditional variances and, in principle, also different (non-normal) conditional distributions of Y given each pair (x,u) of values X and U. ◁

3.1.4 Conditional Probability

The conditional probability of an event A given an event B can be used to describe the *dependency* of A on B with respect to a probability measure P on $\mathscr{A}$. We will also use this concept in order to introduce the concept of a *conditional probability measure*.

Definition 3.13 (Conditional Probability)
Let $(\Omega, \mathscr{A}, P)$ be a probability space, $A, B \in \mathscr{A}$, and $P(B) > 0$. Then

$$P(A \mid B) := \frac{P(A \cap B)}{P(B)} \tag{3.3}$$

is called the conditional probability of A given B (with respect to P).

Example 3.14 (Joe and Ann With Randomized Assignment) Consider again Table 3.1, define $\Omega_U = \{Joe, Ann\}$ and $\Omega_X = \{yes, no\}$, and let

$$C = \Omega_U \times \Omega_X \times \{+\} = \big\{(Joe, no, +), (Joe, yes, +), (Ann, no, +), (Ann, yes, +)\big\}$$

be the event that the *drawn person is successful* (no matter who is drawn and whether or not he or she is treated). Furthermore, let

$$B = \Omega_U \times \{yes\} \times \Omega_Y = \big\{(Joe, yes, -), (Joe, yes, +), (Ann, yes, -), (Ann, yes, +)\big\}$$

denote the event that the *drawn person is treated*. Then Equation (3.3) yields:

$$P(C \mid B) = \frac{P(C \cap B)}{P(B)} = \frac{P(\Omega_U \times \{yes\} \times \{+\})}{P(\Omega_U \times \{yes\} \times \Omega_Y)} = \frac{.16 + .08}{.04 + .16 + .12 + .08} = .6.$$

Conditioning on the event B^c that the *drawn person is not treated* yields

$$P(C \mid B^c) = \frac{P(C \cap B^c)}{P(B^c)} = \frac{P(\Omega_U \times \{no\} \times \{+\})}{P(\Omega_U \times \{no\} \times \Omega_Y)} = \frac{.21 + .06}{.09 + .21 + .24 + .06} = .45.$$

In this example, the difference $P(C \mid B) - P(C \mid B^c) = .6 - .45 = .15$ *can* be used to evaluate the average effect of the treatment. In fact, if we additionally take into account the event

$$A = \{Joe\} \times \Omega_X \times \Omega_Y = \big\{(Joe, no, -), (Joe, no, +), (Joe, yes, -), (Joe, yes, +)\big\},$$

then it is easy to see that .15 is the average of the two individual treatment effects

$$P(C \mid B \cap A) - P(C \mid B^c \cap A) = .8 - .7 = .1$$

and

$$P(C \mid B \cap A^c) - P(C \mid B^c \cap A^c) = .4 - .2 = .2$$

of Joe and Ann, respectively. In contrast, if we compute the corresponding conditional probabilities for the random experiment presented in Table 1.2, then we receive $P(C \mid B) - P(C \mid B^c) = .42 - .6 = -.18$, while the two individual treatment effects remain unchanged. Hence, in the example of Table 1.2, the difference $P(C \mid B) - P(C \mid B^c) = -.18$ is completely misleading if interpreted as the average effect of the treatment. ◁

Example 3.15 (Nonorthogonal Two-Factorial Experiment) In the fourth column of Table 1.5 we displayed the conditional treatment probabilities $P(X=1\,|\,U=u)$. These conditional probabilities are the conditional probabilities of an event A given an event B as introduced in Definition 3.13. Consider the event

$$A \;=\; \Omega_U \times \{\mathit{treatment}\,1\} \times \Omega_Y$$

and the event

$$B \;=\; \{\mathit{Tom}\} \times \Omega_X \times \Omega_Y$$

(see Example 1.5 for the definition of the sets Ω_U, Ω_X, and Ω_Y). Then

$$P(A\,|\,B) \;=\; P(X=1\,|\,U=\mathit{Tom}) \;=\; 10/60$$

(see the first row in the fourth column of Table 1.5). The notation $P(X=1\,|\,U=u)$ will be introduced in Example 3.43. ◁

3.1.5 Conditional Probability Measure

Just like (unconditional) probabilities, conditional probabilities of events $A \in \mathscr{A}$ given an event $B \in \mathscr{A}$ are values of a probability measure. This is stated in the following theorem. (For a proof see SN-Th. 4.28.)

Theorem 3.16 (Conditional Probability Measure)
Let $(\Omega, \mathscr{A}, P)$ be a probability space, let $B \in \mathscr{A}$ and assume $P(B) > 0$. Then the function $P^B\colon \mathscr{A} \to [0,1]$ defined by

$$P^B(A) \;=\; P(A\,|\,B), \quad \forall\, A \in \mathscr{A}, \tag{3.4}$$

is a probability measure on $(\Omega, \mathscr{A})$.

According to this theorem, for each $B \in \mathscr{A}$ with $P(B) > 0$, the function P^B is a probability measure on $(\Omega, \mathscr{A})$. Hence, according to Remark 3.8, the triple $(\Omega, \mathscr{A}, P^B)$ is a probability space.

Definition 3.17 (Conditional Probability Measure)
Let $(\Omega, \mathscr{A}, P)$ be a probability space, $B \in \mathscr{A}$, and $P(B) > 0$. Then the function P^B defined by Equation (3.4) is called the B-conditional probability measure on $(\Omega, \mathscr{A})$.

The function P^B assigns to each event $A \in \mathscr{A}$ its conditional probability given B. Of course, if B and C are different events, the conditional probabilities $P^B(A) = P(A\,|\,B)$ and $P^C(A) = P(A\,|\,C)$ can differ from each other.

Example 3.18 (Joe and Ann With Randomized Assignment) Consider the example presented in Table 3.1. We specify the B-conditional probability measure $P^B\colon \mathscr{A} \to [0,1]$ for the event

$$B = \{(Joe, yes, -),\ (Joe, yes, +),\ (Ann, yes, -),\ (Ann, yes, +)\}$$

that the *drawn person is treated*. Using the probabilities of the elementary events displayed in the second column of Table 3.1 and Rule (ii) of SN-Box 4.1, we compute

$$\begin{aligned} P(B) &= P\big(\{(Joe, yes, -),\ (Joe, yes, +),\ (Ann, yes, -),\ (Ann, yes, +)\}\big) \\ &= P\big(\{(Joe, yes, -)\}\big) + P\big(\{(Joe, yes, +)\}\big) + P\big(\{(Ann, yes, -)\}\big) + P\big(\{(Ann, yes, +)\}\big) \\ &= .04 + .16 + .12 + .08 = .4. \end{aligned}$$

For the first two elementary events,

$$P^B(\{\omega_1\}) = P^B(\{(Joe, no, +)\}) = P^B(\{\omega_2\}) = P^B(\{(Joe, no, -)\}) = 0,$$

because the intersections $\{(Joe, no, -)\} \cap B$ and $\{(Joe, no, +)\} \cap B$ are the empty set. For the next two elementary events, the B-conditional probabilities are

$$P^B(\{\omega_3\}) = P^B\big(\{(Joe, yes, -)\}\big) = \frac{P\big(\{(Joe, yes, -)\} \cap B\big)}{P(B)} = \frac{.04}{.4} = .1$$

and

$$P^B(\{\omega_4\}) = P^B\big(\{(Joe, yes, +)\}\big) = \frac{P\big(\{(Joe, yes, +)\} \cap B\big)}{P(B)} = \frac{.16}{.4} = .4.$$

For the next two elementary events,

$$P^B(\{\omega_5\}) = P^B\big(\{(Ann, no, -)\}\big) = P^B(\{\omega_6\}) = P^B\big(\{(Ann, no, +)\}\big) = 0,$$

because again $\{(Ann, no, -)\} \cap B = \{(Ann, no, +)\} \cap B = \emptyset$. Finally, for the last two elementary events, the B-conditional probabilities are

$$P^B(\{\omega_7\}) = P^B\big(\{(Ann, yes, -)\}\big) = \frac{P\big(\{(Ann, yes, -)\} \cap B\big)}{P(B)} = \frac{.12}{.4} = .3$$

and

$$P^B(\{\omega_8\}) = P^B\big(\{(Ann, yes, +)\}\big) = \frac{P\big(\{(Ann, yes, +)\} \cap B\big)}{P(B)} = \frac{.08}{.4} = .2.$$

These eight probabilities are summarized in the last column of Table 3.1.

Except for $\emptyset$, all other events $A \in \mathscr{A}$ are unions of these elementary events. Because the elementary events are *disjoint,* the probabilities of their unions can easily be computed using finite additivity of a probability measure [see Rule (ii) of SN-Box 4.1]. For example, the B-conditional probability of the event

$$C = \{(Joe, no, +),\ (Joe, yes, +),\ (Ann, no, +),\ (Ann, yes, +)\}$$

that the sampled person has success is

$$\begin{aligned} P^B(C) &= P^B\big(\{(Joe, no, +),\ (Joe, yes, +),\ (Ann, no, +),\ (Ann, yes, +)\}\big) \\ &= P^B\big(\{(Joe, no, +)\}\big) + P^B\big(\{(Joe, yes, +)\}\big) + P^B\big(\{(Ann, no, +)\}\big) + P^B\big(\{(Ann, yes, +)\}\big) \\ &= 0 + .4 + 0 + .2 = .6. \end{aligned}$$

◁

Example 3.19 (Nonorthogonal Two-Factorial Experiment) Consider again the example presented in Table 1.5. For the event

$$A = \{Tom, Tim\} \times \Omega_X \times \Omega_Y$$

that the sampled person has status *low*, we specify its B-conditional probability, where

$$B = \Omega_U \times \{treatment\,1\} \times \Omega_Y$$

is the event that the *drawn person receives treatment* 1. According to the second row of the last column of Table 1.6, $P(B) = 1/3$. Then, using the probabilities displayed in Table 1.5,

$$\begin{aligned} P^B(A) &= \frac{P\big(\{Tom, Tim\} \times \Omega_X \times \Omega_Y \cap \Omega_U \times \{treatment\,1\} \times \Omega_Y\big)}{P(B)} && \text{[Eqs. (3.3), (3.4)]} \\ &= \frac{(10/60 + 18/60) \cdot 1/8}{1/3} = \frac{7/120}{1/3} = \frac{7}{40}. && [P(A \cap B) = P(A \mid B) \cdot P(B)] \end{aligned}$$

This result is consistent with the probabilities displayed in Table 1.7. Again note the distinction between the two measures P and P^B. While $P^B(A) = 7/40$, the probability of A with respect to the measure P is $P(A) = 2/8$ (see again Table 1.5.) ◁

In the following lemma we consider the relationship between conditional probabilities with respect to the measures P^B and P. (For a proof see SN-Lemma 4.30.)

Lemma 3.20 (Conditional Probabilities With Respect to P^B)
Let $(\Omega, \mathscr{A}, P)$ be a probability space. If $A, B, C \in \mathscr{A}$ and $P(B \cap C) > 0$, then

$$P^B(A \mid C) = P(A \mid B \cap C). \tag{3.5}$$

According to this lemma, the C-conditional probability of the event A with respect to the B-conditional probability measure P^B is identical to the $(B \cap C)$-conditional probability of A with respect to the probability measure P.

According to the following lemma, the measure P^B is *absolutely continuous with respect to P*, that is,

$$\forall\, A \in \mathscr{A}\colon\ P(A) = 0 \quad \Rightarrow \quad P^B(A) = 0. \tag{3.6}$$

Hence, if A is an event such that $P(A) = 0$, then the B-conditional probability of A is null as well. Proposition (3.6) is denoted by $P^B \ll P$ [see SN-Def. 3.70 (i)]. (For a proof of Lemma 3.21 see SN-Lemma 4.32.)

Lemma 3.21 (Absolute Continuity of a Conditional Probability Measure)
Let $(\Omega, \mathscr{A}, P)$ be a probability space, $B \in \mathscr{A}$, and $P(B) > 0$. Then P^B is absolutely continuous with respect to P, that is, $P^B \ll P$.

Remark 3.22 (P is not Necessarily Absolute Continuous With Respect to $P^{X=x}$) In contrast to $P^B \ll P$, which always holds if $P(B) > 0$, the proposition $P \ll P^B$ does *not* always hold (see Example 3.23). ◁

Table 3.2. No treatment for Joe

Outcomes ω_i				Observables			Conditional expectations				
Unit Treatment Success	$P(\{\omega_i\})$	$P^{X=0}(\{\omega_i\})$	$P^{X=1}(\{\omega_i\}) = P^B(\{\omega_i\})$	Person variable U	Treatment variable X	Outcome variable Y	$E(Y\|X,U)$	$E(Y\|X)$	$E(X\|U)$	$E^{X=0}(Y\|U)$	$E^{X=1}(Y\|U)$
$\omega_1 = (Joe,\ no,\ -)$	.15	.1875	0	*Joe*	0	0	.7	.512	0	.7	99
$\omega_2 = (Joe,\ no,\ +)$	.35	.4375	0	*Joe*	0	1	.7	.512	0	.7	99
$\omega_3 = (Joe,\ yes,\ -)$	0	0	0	*Joe*	1	0	99	.4	0	.7	99
$\omega_4 = (Joe,\ yes,\ +)$	0	0	0	*Joe*	1	1	99	.4	0	.7	99
$\omega_5 = (Ann,\ no,\ -)$	.24	.3	0	*Ann*	0	0	.2	.512	.5	.2	.4
$\omega_6 = (Ann,\ no,\ +)$	.06	.075	0	*Ann*	0	1	.2	.512	.5	.2	.4
$\omega_7 = (Ann,\ yes,\ -)$	.12	0	.6	*Ann*	1	0	.4	.4	.5	.2	.4
$\omega_8 = (Ann,\ yes,\ +)$	.08	0	.4	*Ann*	1	1	.4	.4	.5	.2	.4

Example 3.23 (No Treatment for Joe) In the example displayed in Table 3.2, the event

$$B = \{(Joe, yes, -),\ (Joe, yes, +),\ (Ann, yes, -),\ (Ann, yes, +)\}$$

that the drawn person is treated, has the probability $P(B) = .2$ and the B-conditional probability of the event

$$A = \{(Joe, no, -),\ (Joe, no, +),\ (Joe, yes, -),\ (Joe, yes, +)\}$$

that Joe is drawn is $P^B(A) = 0$, whereas $P(A) = .5$. ◁

3.2 Random Variable

Although a probability space contains all information about the random experiment considered, this information is not yet processed well enough to be easily grasped. Random variables and their expectations, variances, covariances, and distributions are important concepts for processing the information contained in a probability space and, in particular, in a probability measure.

Remark 3.24 (Inverse Images) In Definition 3.25 we will use the concept of the *inverse image* of a set A' under a mapping Y with domain Ω and co-domain Ω'_Y, which will be abbreviated by $Y\colon \Omega \to \Omega'_Y$. We use the notation $Y^{-1}(A')$

$$Y^{-1}(A') := \{\omega \in \Omega\colon Y(\omega) \in A'\}, \quad A' \in \mathscr{A}'_X. \tag{3.7}$$

Other notations for inverse images under Y are $\{Y \in A'\} = Y^{-1}(A')$ and $\{Y = y\} = Y^{-1}(\{y\})$. If $(\Omega, \mathscr{A}, P)$ is a probability space, then according to this definition, the inverse image $Y^{-1}(A')$ is the event that Y takes on a value in the set A'.

In Definition 3.25 we require that all inverse images $Y^{-1}(A')$ are elements of the σ-algebra $\mathscr{A}$ on Ω. Because the measure $P : \mathscr{A} \to [0,1]$ assigns a probability to *all* elements of $\mathscr{A}$, the probabilities $P[Y^{-1}(A')]$ of these inverse images are determined by P. The measurable space $(\overline{\mathbb{R}}, \overline{\mathscr{B}})$ occurring in Definition 3.25 refers to the set $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$ and the Borel σ-algebra $\overline{\mathscr{B}}$ on this set (for more details see SN-section 1.2.2). ◁

Definition 3.25 (Random Variable)
Let $(\Omega, \mathscr{A}, P)$ be a probability space and $(\Omega'_Y, \mathscr{A}'_Y)$ a measurable space. Then the mapping $Y : \Omega \to \Omega'_Y$ is called a random variable on $(\Omega, \mathscr{A}, P)$ with value space $(\Omega'_Y, \mathscr{A}'_Y)$, if

$$Y^{-1}(A') \in \mathscr{A}, \quad \forall\, A' \in \mathscr{A}'_Y. \tag{3.8}$$

If $(\Omega'_Y, \mathscr{A}'_Y) = (\mathbb{R}, \mathscr{B})$, then Y is called real-valued, and if $(\Omega'_Y, \mathscr{A}'_Y) = (\overline{\mathbb{R}}, \overline{\mathscr{B}})$, then Y is called numerical.

Example 3.26 (Indicator of an Event) Let $(\Omega, \mathscr{A}, P)$ be a probability space, $A \in \mathscr{A}$, and consider the measurable space $(\mathbb{R}, \mathscr{B})$. Then $1_A : \Omega \to \mathbb{R}$ defined by

$$1_A(\omega) = \begin{cases} 1, \text{ if } \omega \in A \\ 0, \text{ otherwise} \end{cases} \tag{3.9}$$

is a random variable on $(\Omega, \mathscr{A}, P)$ with value space $(\mathbb{R}, \mathscr{B})$. It is called the *indicator* of (the event) A. If we consider the event $\{X{=}x\} = \{\omega \in \Omega : X(\omega) = x\}$, then we also use the notation $1_{X=x}$.

There are four different inverse images of sets $B \in \mathscr{B}$ under 1_A:

$$\forall B \in \mathscr{B}: \quad 1_A^{-1}(B) = \begin{cases} A, & \text{if } 0 \notin B \text{ and } 1 \in B \\ A^c, & \text{if } 0 \in B \text{ and } 1 \notin B \\ \Omega, & \text{if } 0 \in B \text{ and } 1 \in B \\ \emptyset, & \text{if } 0 \notin B \text{ and } 1 \notin B. \end{cases}$$

If $A \in \mathscr{A}$, then all four inverse images are elements of $\mathscr{A}$, which follows from the definition of a σ-algebra (see Def. 3.3). Note that the set of these four inverse images is a σ-algebra on Ω. ◁

Example 3.27 (Joe and Ann With Randomized Assignment) In Equation (3.2) we specified the set $\Omega = \Omega_U \times \Omega_X \times \Omega_Y$ of possible outcomes of this random experiment. The third column of Table 3.1 displays the person variable U assigning an element of the set $\Omega_U = \{Joe, Ann\}$ to each possible outcome $\omega_i \in \Omega$. Hence, U is a mapping with domain Ω and co-domain Ω_U, that is, $U : \Omega \to \Omega_U$. If we choose $\mathscr{A} = \mathscr{P}(\Omega)$ to be the power set of Ω and the measurable space $(\Omega_U, \mathscr{A}_U)$ with $\mathscr{A}_U = \mathscr{P}(\Omega_U) = \big\{\{Joe\}, \{Ann\}, \Omega_U, \emptyset\big\}$, then the mapping $U : \Omega \to \Omega_U$ is a random variable on $(\Omega, \mathscr{A}, P)$ with value space $(\Omega_U, \mathscr{A}_U)$. The four inverse images of elements $A' \in \mathscr{A}_U$ are

$$\begin{aligned} U^{-1}(\{Joe\}) &= \{Joe\} \times \Omega_X \times \Omega_Y = \{\omega_1, \ldots, \omega_4\} \\ U^{-1}(\{Ann\}) &= \{Ann\} \times \Omega_X \times \Omega_Y = \{\omega_5, \ldots, \omega_8\} \\ U^{-1}(\Omega_U) &= \Omega = \{\omega_1, \ldots, \omega_8\} \end{aligned}$$

$$U^{-1}(\emptyset) = \emptyset.$$

Because we choose $\mathscr{A}$ to be the power set of Ω, all four inverse images $U^{-1}(A')$, $A' \in \mathscr{A}_U$, are necessarily elements of $\mathscr{A}$ [see Eq. (3.8)]. Note again that the set of these four inverse images is a σ-algebra on Ω.

The fourth column of Table 3.1 displays the treatment variable X assigning an element of the set $\mathbb{R}$, the numbers 0 and 1, to each possible outcome $\omega_i \in \Omega$. If we choose $\mathscr{A} = \mathscr{P}(\Omega)$ to be the power set of Ω and the measurable space $(\mathbb{R}, \mathscr{B})$, then the mapping $X\colon \Omega \to \mathbb{R}$ is a random variable on $(\Omega, \mathscr{A}, P)$ with value space $(\mathbb{R}, \mathscr{B})$. ◁

Example 3.28 (Nonorthogonal Two-Factorial Experiment) In Example 3.2 we specified the set $\Omega = \Omega_U \times \Omega_X \times \Omega_Y$ of possible outcomes of this random experiment. The person variable U appearing in Table 1.5 is defined by

$$\forall\, \omega \in \Omega\colon \quad U(\omega) = \begin{cases} \mathit{Tom}\,, \text{ if } \omega \in \{\mathit{Tom}\} \times \Omega_X \times \Omega_Y \\ \mathit{Tim}\,, \text{ if } \omega \in \{\mathit{Tim}\} \times \Omega_X \times \Omega_Y \\ \quad \vdots \\ \mathit{Mia}\,, \text{ if } \omega \in \{\mathit{Mia}\} \times \Omega_X \times \Omega_Y. \end{cases} \tag{3.10}$$

Because we chose $\mathscr{A}$ to be the product of the σ-algebras $\mathscr{A}_U = \mathscr{P}(\Omega_U)$, $\mathscr{A}_X = \mathscr{P}(\Omega_X)$, and the Borel σ-algebra $\mathscr{B}$ (see Example 3.6), the definition of a product of σ-algebras (see SN-Def. 1.31) implies that all inverse images $U^{-1}(A')$, $A' \in \mathscr{A}_U$, are elements of $\mathscr{A}$ [see again Eq. (3.8)]. Hence, U is a random variable on $(\Omega, \mathscr{A}, P)$ with value space $(\Omega_U, \mathscr{A}_U)$.

The treatment variable $X\colon \Omega \to \mathbb{R}$ appearing in Table 1.5 is defined by

$$\forall\, \omega \in \Omega\colon \quad X(\omega) = \begin{cases} 0, & \text{if } \omega \in \Omega_U \times \{\mathit{control}\} \times \Omega_Y \\ 1, & \text{if } \omega \in \Omega_U \times \{\mathit{treatment}\,1\} \times \Omega_Y \\ 2, & \text{if } \omega \in \Omega_U \times \{\mathit{treatment}\,2\} \times \Omega_Y. \end{cases} \tag{3.11}$$

Again, because we chose $\mathscr{A}$ to be the product of the σ-algebras $\mathscr{A}_U = \mathscr{P}(\Omega_U)$, $\mathscr{A}_X = \mathscr{P}(\Omega_X)$, and the Borel σ-algebra $\mathscr{B}$, the definition of a product of σ-algebras implies that all inverse images $X^{-1}(B)$, $B \in \mathscr{B}$, are elements of $\mathscr{A}$ [see again Eq. (3.8)]. Hence, X is a random variable on $(\Omega, \mathscr{A}, P)$ with value space $(\mathbb{R}, \mathscr{B})$ (see also Exercise 3-3). The inverse images $X^{-1}(B)$, $B \in \mathscr{B}$, are

$$X^{-1}(B) = \begin{cases} \Omega_U \times \{\mathit{control}\} \times \Omega_Y, & \text{if } 0 \in B, 1, 2 \notin B \\ \Omega_U \times \{\mathit{treatment}\,1\} \times \Omega_Y, & \text{if } 0, 2 \notin B, 1 \in B \\ \Omega_U \times \{\mathit{treatment}\,2\} \times \Omega_Y, & \text{if } 0, 1 \notin B, 2 \in B \\ \Omega_U \times \{\mathit{treatment}\,0,\ \mathit{treatment}\,1\} \times \Omega_Y, & \text{if } 0, 1 \in B, 2 \notin B \\ \Omega_U \times \{\mathit{treatment}\,0,\ \mathit{treatment}\,2\} \times \Omega_Y, & \text{if } 0, 2 \in B, 1 \notin B \\ \Omega_U \times \{\mathit{treatment}\,1,\ \mathit{treatment}\,2\} \times \Omega_Y, & \text{if } 0 \notin B, 1, 2 \in B \\ \Omega, & \text{if } 0, 1, 2 \in B \\ \emptyset, & \text{if } 0, 1, 2 \notin B. \end{cases} \tag{3.12}$$

Note that the set of these eight inverse images is a σ-algebra on Ω.

The (qualitative) covariate $Z\colon \Omega \to \Omega'_Z$ appearing in Table 1.5 is defined by

$$\forall\,\omega\in\Omega\colon\quad Z(\omega)\;=\;\begin{cases} \mathit{low}, & \text{if } \omega\in\{\mathit{Tim},\mathit{Tom}\}\times\Omega_X\times\Omega_Y\\ \mathit{med}, & \text{if } \omega\in\{\mathit{Joe},\ldots,\mathit{Eva}\}\times\Omega_X\times\Omega_Y\\ \mathit{hi}, & \text{if } \omega\in\{\mathit{Sue},\mathit{Mia}\}\times\Omega_X\times\Omega_Y. \end{cases} \tag{3.13}$$

For this random variable we choose the value space $\big(\Omega'_Z,\mathscr{P}(\Omega'_Z)\big)$. Again, because we chose $\mathscr{A}$ to be the product of the σ-algebras $\mathscr{A}_U=\mathscr{P}(\Omega_U)$, $\mathscr{A}_X=\mathscr{P}(\Omega_X)$, and the Borel σ-algebra $\mathscr{B}$, the definition of a product of σ-algebras implies that all inverse images $Z^{-1}(B)$, $B\in\mathscr{P}(\Omega'_Z)$, are elements of $\mathscr{A}$ [see again Eq. (3.8)].

The inverse images $Z^{-1}(B)$, $B\in\mathscr{P}(\Omega'_Z)$, are

$$\begin{aligned} Z^{-1}(\{\mathit{low}\}) &= \{\mathit{Tim},\mathit{Tom}\}\times\Omega_X\times\Omega_Y\\ Z^{-1}(\{\mathit{med}\}) &= \{\mathit{Joe},\mathit{Jim},\mathit{Ann},\mathit{Eva}\}\times\Omega_X\times\Omega_Y\\ Z^{-1}(\{\mathit{hi}\}) &= \{\mathit{Sue},\mathit{Mia}\}\times\Omega_X\times\Omega_Y\\ Z^{-1}(\{\mathit{low},\mathit{med}\}) &= \{\mathit{Tim},\mathit{Tom},\mathit{Joe},\mathit{Jim},\mathit{Ann},\mathit{Eva}\}\times\Omega_X\times\Omega_Y\\ Z^{-1}(\{\mathit{low},\mathit{hi}\}) &= \{\mathit{Tim},\mathit{Tom},\mathit{Sue},\mathit{Mia}\}\times\Omega_X\times\Omega_Y\\ Z^{-1}(\{\mathit{med},\mathit{hi}\}) &= \{\mathit{Joe},\mathit{Jim},\mathit{Ann},\mathit{Eva},\mathit{Sue},\mathit{Mia}\}\times\Omega_X\times\Omega_Y\\ Z^{-1}(\Omega'_Z) &= \Omega\\ Z^{-1}(\emptyset) &= \emptyset. \end{aligned} \tag{3.14}$$

Note again that the set of these eight inverse images is a σ-algebra on Ω. As shown in Theorem 3.29 this is not a coincidence. ◁

3.2.1 σ-Algebra Generated by a Random Variable

In Examples 3.26, 3.27, and 3.28 we already noted that the set of all inverse images under the random variable considered is a σ-algebra on Ω. In a sense, such a σ-algebra carries the information associated with the random variable considered; it contains all events that can be represented by that random variable. In the following theorem we formulate the general proposition. This theorem and Definition 3.31 are of fundamental importance for probability theory.

Theorem 3.29 (σ-Algebra Generated by a Random Variable)
Let Y be a random variable on $(\Omega,\mathscr{A},P)$ with value space $(\Omega'_Y,\mathscr{A}'_Y)$. Then

$$Y^{-1}(\mathscr{A}'_Y)\;:=\;\big\{\,Y^{-1}(A')\colon A'\in\mathscr{A}'_Y\,\big\} \tag{3.15}$$

is a σ-algebra on Ω.

For a proof see Klenke (2013, Theorem 1.81, p. 33).

Remark 3.30 (Smallest σ-Algebra) Note that $Y^{-1}(\mathscr{A}'_Y)$ is the smallest σ-algebra $\mathscr{C}$ on Ω such that $Y^{-1}(\mathscr{A}'_Y)\subset\mathscr{C}$. ◁

The set $Y^{-1}(\mathscr{A}'_Y)$ contains all sets in $\mathscr{A}$ that can be represented by Y and elements of $\mathscr{A}'_Y$. Because $Y^{-1}(\mathscr{A}'_Y)$ is important, it has an own name and an alternative notation, which is sometimes more convenient.

Definition 3.31 (σ-Algebra Generated by a Random Variable)
The set $Y^{-1}(\mathscr{A}'_Y)$ defined by Equation (3.15) is called the σ-algebra generated by Y and $\mathscr{A}'_Y$. If there is no ambiguity about $\mathscr{A}'_Y$, then we also say that $Y^{-1}(\mathscr{A}'_Y)$ is generated by Y and use the notation

$$\sigma(Y) := Y^{-1}(\mathscr{A}'_Y). \tag{3.16}$$

Remark 3.32 (Measurability of a Random Variable) If Y is a random variable on a probability space $(\Omega, \mathscr{A}, P)$, $\mathscr{C} \subset \mathscr{A}$ is a σ-algebra, and

$$\sigma(Y) \subset \mathscr{C}, \tag{3.17}$$

then we say that Y is *$\mathscr{C}$-measurable* or *measurable with respect to $\mathscr{C}$*. If Z is a random variable Z on $(\Omega, \mathscr{A}, P)$ and

$$\sigma(Y) \subset \sigma(Z), \tag{3.18}$$

then we say that Y is *Z-measurable* or *measurable with respect to Z*. (See SN-chapter 2 for more details on this and other concepts related to measurability of mappings.) ◁

3.2.2 Distribution of a Random Variable

In Definition 3.25 we require that the inverse images $Y^{-1}(A')$ of all sets $A' \in \mathscr{A}'_Y$ under Y are elements of the σ-algebra $\mathscr{A}$ on Ω. Using the notation introduced in Definition 3.31, this is equivalent to requiring $\sigma(Y) \subset \mathscr{A}$. This property allows us to define the *distribution* of a random variable as follows:

Definition 3.33 (Distribution of a Random Variable)
Let Y be a random variable on $(\Omega, \mathscr{A}, P)$ with value space $(\Omega'_Y, \mathscr{A}'_Y)$. Then the function $P_Y\colon \mathscr{A}'_Y \to [0,1]$ defined by

$$P_Y(A') = P[Y^{-1}(A')], \quad \forall\, A' \in \mathscr{A}'_Y, \tag{3.19}$$

is called the distribution of Y (with respect to P).

Remark 3.34 (A New Probability Space) Definition 3.33 implies that *every* random variable Y on a probability space $(\Omega, \mathscr{A}, P)$ has a distribution P_Y. Furthermore, $P_Y\colon \mathscr{A}'_Y \to [0,1]$ is also a measure, the *image measure of P under X* (see SN-Th. 2.80 and SN-Def. **??**). Because $P_Y(\Omega'_Y) = P(\Omega) = 1$, we can conclude that P_Y is a probability measure, and $(\Omega'_Y, \mathscr{A}'_Y, P_Y)$ is also a probability space. ◁

Example 3.35 (Joe and Ann With Randomized Assignment) In Example 3.27 we showed that U is random variable on $(\Omega, \mathscr{A}, P)$ with value space $(\Omega_U, \mathscr{A}_U)$, where $\Omega_U = \{Joe, Ann\}$ and $\mathscr{A}_U = \big\{\{Joe\}, \{Ann\}, \Omega_U, \emptyset\big\}$. The distribution of U is

$$\begin{aligned} P_U(\{Joe\}) &= P[U^{-1}(\{Joe\})] = P(\{\omega_1, \ldots, \omega_4\}) = P(\{\omega_1\}) + \ldots + P(\{\omega_4\}) = .5, \\ P_U(\{Ann\}) &= P[U^{-1}(\{Ann\})] = P(\{\omega_5, \ldots, \omega_8\}) = P(\{\omega_5\}) + \ldots + P(\{\omega_8\}) = .5, \end{aligned}$$

$$P_U(\{\Omega_U\}) = P[U^{-1}(\Omega_U)] = P(\{\omega_1,\ldots,\omega_8\}) = P(\{\omega_1\})+\ldots+P(\{\omega_8\}) = 1,$$
$$P_U(\{\emptyset\}) = P[U^{-1}(\emptyset)] = P(\emptyset) = 0.$$

In Example 3.27 we also specified the random variable X on $(\Omega,\mathscr{A},P)$ with value space $(\mathbb{R},\mathscr{B})$. The distribution of X is as follows:

$$\forall B\in\mathscr{B}: \quad P_X(B) = P[X^{-1}(B)] = \begin{cases} P(\{\omega_1,\omega_2,\omega_5,\omega_6\}) = .6, & \text{if } 0\in B \text{ and } 1\notin B \\ P(\{\omega_3,\omega_4,\omega_7,\omega_8\}) = .4, & \text{if } 0\notin B \text{ and } 1\in B \\ P(\Omega) = 1, & \text{if } 0\in B \text{ and } 1\in B \\ P(\emptyset) = 0, & \text{if } 0\notin B \text{ and } 1\notin B. \end{cases}$$

Note that in this equation we assign a probability to *all* elements $B\in\mathscr{B}$, not only to four elements of $\mathscr{B}$. ◁

Example 3.36 (Nonorthogonal Two-Factorial Experiment) In Example 3.28 we defined the random variable U with value space $(\Omega_U,\mathscr{A}_U)$, where $\mathscr{A}_U=\mathscr{P}(\Omega_U)$. The distribution of U can be specified as follows:

$$\forall\{u\}\in\mathscr{A}_U: \quad P_U(\{u\}) = P[U^{-1}(\{u\})] = \frac{1}{8}. \tag{3.20}$$

Because, except for the empty set, all $A'\in\mathscr{A}_U$ are unions of the singletons $\{u\}$ and the singletons are disjunct, the probabilities of all elements $A'\in\mathscr{A}_U$ can be computed as a sum of the probabilities of these singletons. For example, the probability of sampling a male person is

$$\begin{aligned} P_U(\{Tom, Tim, Joe, Jim\}) &= P[U^{-1}(\{Tom, Tim, Joe, Jim\})] \\ &= P_U(\{Tom\})+P_U(\{Tim\})+P_U(\{Joe\})+P_U(\{Jim\}) \\ &= \frac{1}{8}+\frac{1}{8}+\frac{1}{8}+\frac{1}{8} = \frac{1}{2} \end{aligned}$$

[see SN-Box 4.1 (ii)]. Note that the distribution of U, that is, the function $P_U\colon\mathscr{A}_U\to[0,1]$, assigns a probability $P_U(A')$ to $2^8=256$ sets $A'\in\mathscr{A}_U$.

In Example 3.28 we also specified the random variable X with value space $(\mathbb{R},\mathscr{B})$. The distribution of X can be specified as follows:

$$\forall B\in\mathscr{B}: \quad P_X(B) = P[X^{-1}(B)]$$
$$= \begin{cases} P(\Omega_U\times\{control\}\times\Omega_Y) = 1/3, & \text{if } 0\in B, 1,2\notin B \\ P(\Omega_U\times\{treatment\,1\}\times\Omega_Y) = 1/3, & \text{if } 0,2\notin B, 1\in B \\ P(\Omega_U\times\{treatment\,2\}\times\Omega_Y) = 1/3, & \text{if } 0,1\notin B, 2\in B \\ P(\Omega_U\times\{treatment\,0,\ treatment\,1\}\times\Omega_Y) = 2/3, & \text{if } 0,1\in B, 2\notin B \\ P(\Omega_U\times\{treatment\,0,\ treatment\,2\}\times\Omega_Y) = 2/3, & \text{if } 0,2\in B, 1\notin B \\ P(\Omega_U\times\{treatment\,1,\ treatment\,2\}\times\Omega_Y) = 2/3, & \text{if } 0\notin B, 1,2\in B \\ P(\Omega) = 1, & \text{if } 0,1,2\in B \\ P(\emptyset) = 0, & \text{if } 0,1,2\notin B. \end{cases}$$

Although there are only eight elements in $\sigma(X)$, this equation assigns a probability to *all* elements $B\in\mathscr{B}$. The number of such elements is uncountable. ◁

3.2.3 Expectation of a Numerical Random Variable

In this section we introduce the concept of the expectation of a numerical random variable. Later we will see that *variances*, *covariances* and *correlations* are expectations of special random variables (see section 3.2.4). All these quantities describe important properties of numerical random variables, although, in general, they do not determine the complete distribution.

In the following definition we use the concept of a measure integral that is well-known in measure theory (for an introduction see SN-chapter 3.) Also note that a random variable Y is called *quasi-integrable* with respect to the measure P if $\int Y^+ dP$ or $\int Y^- dP$ are finite, where Y^+ and Y^- denote the positive and negative parts of Y, respectively (see SN-Rem. 2.62 and SN-Def. 3.28).

Definition 3.37 (Expectation of a Numerical Random Variable)
Let Y be a numerical random variable on $(\Omega, \mathscr{A}, P)$ that is quasi-integrable with respect to P. Then we define

$$E(Y) := \int Y \, dP, \tag{3.21}$$

call it the expectation of Y (with respect to P), and say that it exists.

Remark 3.38 (Existence of the Expectation) Note that $E(Y)$ can be infinite. Furthermore, if $E(Y)$ *exists* then we also say that Y is a random variable *with expectation* $E(Y)$. If Y is not quasi-integrable with respect to P and therefore also not P-integrable, then we say that the expectation of Y with respect to P does *not exist*. ◁

Remark 3.39 (Notation and Synonymous Terms) A synonym for expectation is *expectation value*. The reference to the measure P is usually omitted if the context is unambiguous. If we consider the expectation with respect to another probability measure on $(\Omega, \mathscr{A})$, for example, the conditional probability measure P^B (see Def. 3.17), then we adapt the notation as follows:

$$E^B(Y) := \int Y \, dP^B. \tag{3.22}$$

◁

Remark 3.40 (Conditional Expectation Given the Event B) The expectation $E^B(Y)$ of Y with respect to the B-conditional probability measure P^B is also called the *B-conditional expectation value of Y* or the *conditional expectation value of Y given the event B*. We also use the notation

$$E(Y|B) := E^B(Y). \tag{3.23}$$

Note that this definition presumes that $B \in \mathscr{A}$ such that $P(B) > 0$. ◁

Remark 3.41 (Numerical Random Variable With a Finite Number of Values) Assume that Y has only a finite number of different values $y_1, \ldots, y_n \in \mathbb{R}$, that is, assume that the image $Y(\Omega)$ of Ω under Y is $\{y_1, \ldots, y_n\} \subset \mathbb{R}$. Then the expectation $E(Y)$ exists and

$$E(Y) = \sum_{i=1}^{n} y_i \cdot P(Y=y_i), \tag{3.24}$$

using the notation $P(Y=y_i) := P[Y^{-1}(\{y_i\})]$ [see SN-Rem. 6.5]. If P_Y denotes the distribution of Y, then we may also write

$$E(Y) = \sum_{i=1}^{n} y_i \cdot P_Y(\{y_i\}). \tag{3.25}$$

Hence, in this case the expectation of Y is simply the weighted sum of its values, each one weighted by its probability $P(Y=y_i) = P_Y(\{y_i\})$. Correspondingly, if $Y(\Omega) = \{y_1,\ldots,y_n\} \subset \mathbb{R}$, then

$$E(Y|B) = \sum_{i=1}^{n} y_i \cdot P(Y=y_i|B). \tag{3.26}$$

If, additionally, $B = \{X=x\}$, then we also use the notation

$$E(Y|X=x) := E(Y|B) \tag{3.27}$$

and call it the $(X=x)$-*conditional expectation value of* Y. For $B = \{X=x\}$, Equation (3.23) yields

$$E(Y|X=x) = E^{X=x}(Y), \tag{3.28}$$

and Equation (3.26) can also be written

$$E(Y|X=x) = \sum_{i=1}^{n} y_i \cdot P(Y=y_i|X=x). \tag{3.29}$$

Note again that in this definition of $E(Y|X=x)$ we presume $P(X=x) > 0$ (see Def. 3.17), which implies that $E(Y|X=x)$ is a uniquely defined number (see Exercise 3-4). In Definition 3.64 we will introduce a more general definition that applies without assuming $P(X=x) > 0$. ◁

Example 3.42 (Expectation of an Indicator) If $(\Omega,\mathscr{A},P)$ is a probability space and 1_A is the indicator of $A \in \mathscr{A}$, then Equations (3.24) and (3.9) yield

$$E(1_A) = 0 \cdot P(1_A=0) + 1 \cdot P(1_A=1) = P(1_A=1) = P(A). \tag{3.30}$$

Considering the event $\{Y=y\}$ and using the notation $1_{Y=y} := 1_{\{Y=y\}}$, this yields

$$E(1_{Y=y}) = P(Y=y). \tag{3.31}$$

◁

Example 3.43 (Conditional Expectation Value of an Indicator) Correspondingly, considering the indicator 1_A of an event $A \in \mathscr{A}$, Equations (3.26) and (3.9) yield

$$E(1_A|B) = 0 \cdot P(1_A=0|B) + 1 \cdot P(1_A=1|B) = P(1_A=1|B) = P(A|B). \tag{3.32}$$

Finally, if $B = \{X=x\}$ and $1_{Y=y} = 1_{\{Y=y\}}$, we introduce the notation

$$P(Y=y|X=x) := E(1_{Y=y}|B) \tag{3.33}$$

and call this number the $(X=x)$-*conditional probability of (the event)* $\{Y=y\}$. ◁

Example 3.44 (Joe and Ann With Randomized Assignment) Consider the random experiment displayed in Table 3.1 and define the event

$$B = \{(Joe, yes, -), (Joe, yes, +), (Ann, yes, -), (Ann, yes, +)\} = \Omega_U \times \{yes\} \times \Omega_Y,$$

that the drawn person is treated (irrespective of whether or not the person is treated and success occurs) and the event

$$C = \{(Joe, no, +), (Joe, yes, +), (Ann, no, +), (Ann, yes, +)\} = \Omega_U \times \Omega_X \times \{+\}.$$

that success occurs (irrespective of which person is drawn and whether or not the person is treated). In Table 3.1 we assigned probabilities to each elementary event $\{\omega_i\}$, $\omega_i \in \Omega$ and defined $X := \mathbb{1}_B$, the treatment variable, as well as $Y := \mathbb{1}_C$ the outcome variable. Applying Equation (3.30) to the indicator $\mathbb{1}_B$ yields:

$$\begin{aligned} E(X) &= E(\mathbb{1}_B) = P(B) \\ &= P[\{(Joe, yes, -)\}] + P[\{(Joe, yes, +)\}] + P[\{(Ann, yes, -)\}] + P[\{(Ann, yes, +)\}] \\ &= .04 + .16 + .12 + .08 = .4. \end{aligned}$$

Similarly, for the indicator $\mathbb{1}_C$, we obtain

$$\begin{aligned} E(Y) &= E(\mathbb{1}_C) = P(C) \\ &= P[\{(Joe, no, +)\}] + P[\{(Joe, yes, +)\}] + P[\{(Ann, no, +)\}] + P[\{(Ann, yes, +)\}] \\ &= .21 + .16 + .06 + .08 = .51. \end{aligned}$$

Correspondingly, Equations (3.26) and (3.9) yield

$$\begin{aligned} E(Y|X=1) &= E(\mathbb{1}_C|B) = P(C|B) = \frac{P(B \cap C)}{P(B)} = \frac{P\big(\{(Joe, yes, +), (Ann, yes, +)\}\big)}{P(B)} \\ &= \frac{.16 + .08}{.4} = .6. \end{aligned}$$

Hence, the conditional probability of success given treatment is .6. In contrast,

$$\begin{aligned} E(Y|X=0) &= E(\mathbb{1}_C|B^c) = P(C|B^c) = \frac{P(B^c \cap C)}{P(B^c)} = \frac{P\big(\{(Joe, no, +), (Ann, no, +)\}\big)}{P(B^c)} \\ &= \frac{.21 + .06}{.6} = .45. \end{aligned}$$

Hence, in this example, in which X and U are independent,

$$E(Y|X=1) - E(Y|X=0) = .15,$$

which is identical to the average of the two individual treatment effects. In contrast, if we compute the corresponding conditional expectation values for the random experiment presented in Table 1.2, then we receive

$$E(Y|X=1) - E(Y|X=0) = .42 - .6 = -.18,$$

while the two individual treatment effects remain unchanged. They are

$$E(Y \mid X=1, U=Joe) - E(Y \mid X=0, U=Joe) \;=\; .8 - .7 \;=\; .1$$

for Joe and

$$E(Y \mid X=1, U=Ann) - E(Y \mid X=0, U=Ann) \;=\; .4 - .2 \;=\; .2$$

for Ann (see Example 3.14 and Table 1.2). Hence, in the example of Table 1.2, the difference $E(Y \mid X=1) - E(Y \mid X=0) = -.18$ is completely misleading if interpreted as the average effect of the treatment. ◁

3.2.4 Variance, Covariance, and Correlation

Variance and standard deviation are the most important parameters describing the *variability* of a random variable. They are defined as follows:

Definition 3.45 (Variance and Standard Deviation)
Let Y be a numerical random variable on a probability space $(\Omega, \mathscr{A}, P)$ and assume that $E(Y^2)$ is finite. Then the variance of Y is defined by

$$Var(Y) \;:=\; E\big([Y - E(Y)]^2\big), \tag{3.34}$$

and the standard deviation of Y is the positive square root of the variance, that is,

$$SD(Y) \;:=\; \sqrt{Var(Y)}. \tag{3.35}$$

According to this definition, $Var(Y)$ is the expectation of the squared *mean centered* random variable $Y - E(Y)$. Note that variances and standard deviations are nonnegative. The variance of Y is also denoted by σ_Y^2 and its standard deviation by σ_Y. Important properties of variances are summarized in SN-Box 6.2.

While the variance quantifies the variability of a numerical random variable, the covariance quantifies the degree of co-variation of two numerical random variables, that is, the degree to which the two variables vary together in the following sense: If one variable takes on a large value (i. e., large positive deviation from its expectation), then the other one tends to take on a large value as well. Furthermore, if one variable takes on a small value (i. e., large negative deviation from its expectation), then the other one tends to take on a small value, too. In this case the covariance will be positive. However, the covariance may also be a negative real number. In this case, the two random variables co-vary in the following sense: If one variable takes on a large value, then the other one tends to take on a small value. Furthermore, if one variable takes on a small value, then the other one tends to take on a large value.

Definition 3.46 (Covariance)
Let X, Y be two numerical random variables on the probability space $(\Omega, \mathscr{A}, P)$ such that $E(X^2)$ and $E(Y^2)$ are finite. Then the covariance of X and Y is defined by

$$Cov(X,Y) \;:=\; E\big([X - E(X)] \cdot [Y - E(Y)]\big). \tag{3.36}$$

Comparing Equations (3.34) and (3.36) to each other shows that the variance is the covariance of a numerical random variable with itself.

Remark 3.47 (Correlated Numerical Random Variables) According to this definition, the *covariance* of X and Y is the expectation of the product of the mean centered variables $X - E(X)$ and $Y - E(Y)$. Hence, a covariance can be negative, zero, or positive. If the covariance is different from zero, then we say that X and Y are *correlated*; otherwise, we say that they are *uncorrelated*. The most important rules of computation for covariances are summarized in SN-Box 7.1. ◁

Note that the covariance is not invariant under multiplication with constants [scale transformations; see SN-Box 7.1 (iii)] of the random variables involved. In contrast, the correlation, which quantifies the strength of the same kind of dependence *is invariant* under scale transformations (see SN-Rem. 7.22).

Definition 3.48 (Correlation)
Let X, Y be two numerical random variables on the probability space $(\Omega, \mathscr{A}, P)$ such that $E(X^2)$ and $E(Y^2)$ are finite. Then the correlation of X and Y is defined by

$$Corr(X,Y) := \begin{cases} \dfrac{Cov(X,Y)}{SD(X) \cdot SD(Y)}, & \text{if } SD(X), SD(Y) > 0, \\ 0, & \text{otherwise.} \end{cases} \tag{3.37}$$

Remark 3.49 (Correlation of a Random Variable With Itself) Assume that $Var(X) > 0$. Because $Cov(X,X) = Var(X) = SD(X) \cdot SD(X)$, Equation (3.37) implies that $Corr(X,X) = 1$. Similarly, because $Cov(X,-X) = -Var(X) = -SD(X) \cdot SD(X)$, Equation (3.37) implies that $Corr(X,-X) = -1$. ◁

The covariance between two numerical random variables and their correlation quantify the strength of their dependence that can be described by a *linear quasi-regression* sometimes also referred to as the *ordinary least-squares regression*.

Definition 3.50 (Linear Quasi-Regression)
Let X, Y be real-valued random variables on the probability space $(\Omega, \mathscr{A}, P)$ such that $E(X^2)$ and $E(Y^2)$ are finite and $Var(X) > 0$, and define the function $MSE : \mathbb{R}^2 \to \mathbb{R}$ by

$$MSE(a_0, a_1) = E\left([Y - (a_0 + a_1 X)]^2\right), \quad \forall (a_0, a_1) \in \mathbb{R}^2. \tag{3.38}$$

Let $(\alpha_0, \alpha_1) \in \mathbb{R}^2$ minimize MSE. Then the function $f : \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = \alpha_0 + \alpha_1 x, \quad \forall x \in \mathbb{R}, \tag{3.39}$$

is called the linear quasi-regression of Y on X. The composition of X and f is denoted by $Q_{lin}(Y|X)$, that is,

$$Q_{lin}(Y|X) = f(X) = \alpha_0 + \alpha_1 X. \tag{3.40}$$

Remark 3.51 (Distinguishing Between f and $f(X)$) According to Equation (3.39) the linear quasi-regression f assigns a real number *to all real numbers*. This applies even if X only takes on two different real values. In contrast, the number of different values of the composition $f(X) = Q_{lin}(Y|X)$ is smaller than or equal to the number of values of X, provided that X takes on a finite number of values only. Also note that $Q_{lin}(Y|X)$ is a random variable on $(\Omega, \mathscr{A}, P)$, whereas the linear quasi-regression f is a random variable on $(\Omega'_X, \mathscr{A}'_X, P_X)$. ◁

3.3 Conditional Expectation and Related Concepts

In this book we build on the concept of a *conditional expectation* and some related concepts such as a *conditional expectation with respect to a conditional probability measure* and a *partial conditional expectation*. A conditional expectation $E(Y|X)$ is used to describe how the conditional expectation values $E(Y|X{=}x)$ of a numerical random variable Y depend on the values x of a (not necessarily numerical) random variable X. In many cases conditional expectations are what we try to estimate in statistical modeling. As shown in Example 3.44, in some applications a conditional expectation $E(Y|X)$ can be used to describe a certain kind of causal dependency, in other cases such causal interpretations of $E(Y|X)$ would lead us astray. In a sense, this book is devoted to spell out the conditions under which causal interpretations of a conditional expectation are warranted.

3.3.1 Conditional Expectation $E(Y|X)$

The concept of a conditional expectation has been introduced by Kolmogorov (1933/1977), together with the axioms of probability. (For an English translation see Kolmogorov, 1956). A detailed introduction to conditional expectations is presented in SN-chapters 9 to 11. In this section we just present the definition, some crucial properties, and simple examples. We start with the a special case and then turn to the general definition.

Definition 3.52 (Conditional Expectation if X is Discrete)
Let Y be a numerical random variable on $(\Omega, \mathscr{A}, P)$ that is nonnegative or has a finite expectation and let the random variable X be a random variable on $(\Omega, \mathscr{A}, P)$ with value space $(\Omega'_X, \mathscr{A}'_X)$. If $X(\Omega) = \{x_1, \ldots, x_m\}$ such that, for all $i = 1, \ldots, m$, $\{x_i\} \in \mathscr{A}'_X$ and $P(X{=}x_i) > 0$, then the X-conditional expectation of Y is defined by

$$E(Y|X) := \sum_{i=1}^{m} E(Y|X{=}x_i) \cdot 1_{X=x_i}\,. \tag{3.41}$$

Remark 3.53 (Values of $E(Y|X)$) Because the indicators $1_{X=x_i}$ are random variables on $(\Omega, \mathscr{A}, P)$ (see Example 3.26), Equation (3.41) shows that $E(Y|X)$ is a random variable on $(\Omega, \mathscr{A}, P)$ (see SN-Example 2.61) as well and that the values of $E(Y|X)$ are the conditional expectation values $E(Y|X{=}x_i)$. In more formal terms,

$$\forall\, \omega \in \Omega\colon \quad E(Y|X)(\omega) \;=\; E(Y|X{=}x_i), \quad \text{if } \omega \in \{X{=}x_i\}. \tag{3.42}$$

Also note that, under the assumptions of Definition 3.52, the conditional expectation is uniquely defined. Note that this does not apply to the general concept of a conditional expectation, which is introduced in the sequel. ◁

In the following general definition of $E(Y|X)$ we use the concept of a σ-algebra $\sigma(X)$ generated by random variable X (see Def. 3.31) and the concept of measurability of a random variable with respect to another one (see Rem. 3.32).

Definition 3.54 (Conditional Expectation)
Let X and Y be random variables on $(\Omega, \mathcal{A}, P)$ with value spaces $(\Omega'_X, \mathcal{A}'_X)$ and $(\overline{\mathbb{R}}, \overline{\mathcal{B}})$, respectively. Assume that Y is nonnegative or with finite expectation $E(Y)$. Then a numerical random variable V on $(\Omega, \mathcal{A}, P)$ is called a version of the X-conditional expectation of Y with respect to P, if the following two conditions hold:

(a) $\sigma(V) \subset \sigma(X)$.
(b) $E(1_C \cdot V) = E(1_C \cdot Y), \quad \forall\, C \in \sigma(X)$.

If V satisfies (a) and (b), then we also use the notation $E(Y|X) := V$.

This definition also applies if X is continuous, which is the case, for example, if X has a normal distribution. A version $E(Y|X)$ of the X-conditional expectation of Y is a random variable on $(\Omega, \mathcal{A}, P)$, and according to condition (a) of Definition 3.54, $E(Y|X)$ is measurable with respect to X (cf. Rem. 3.32).

Remark 3.55 (The Set $\mathcal{E}(Y|X)$) Note that there can be several random variables satisfying conditions (a) and (b). Therefore, we define $\mathcal{E}(Y|X)$ to be the set of all random variables satisfying conditions (a) and (b) of Definition 3.54. Hence, $\mathcal{E}(Y|X)$ denotes the set of all versions of the X-conditional expectation of Y with respect to the measure P. ◁

Remark 3.56 (P-Uniqueness) According to SN-Remark 10.15,

$$V, V^* \in \mathcal{E}(Y|X) \quad \Rightarrow \quad V \underset{P}{=} V^*, \tag{3.43}$$

where $V \underset{P}{=} V^*$ is a shortcut for

$$P\big(\{\omega \in \Omega\colon V(\omega) = V^*(\omega)\}\big) = 1. \tag{3.44}$$

If Equation (3.44) holds, then we say that V and V^* are *P-equivalent* or *identical almost surely with respect to P*. If (3.43) holds, that is, if *$V \underset{P}{=} V^*$ for all pairs $V, V^* \in \mathcal{E}(Y|X)$*, then we say that *$E(Y|X)$ is P-unique*. ◁

Remark 3.57 (Implications of P-Almost Sure Identity) If two random variables are P-almost surely identical, then they have identical distributions, expectations, variances, and covariances with other random variables provided that these expectations, variances, and covariances exist [see SN-Box 6.1 (ix), SN-Box 6.2 (v), and SN-Box 7.1 (x)]. ◁

Remark 3.58 (Conditional Probability Given a Random Variable) Let 1_A denote the indicator of the event $A \in \mathcal{A}$. We introduce the notation

$$P(A|X) := E(1_A|X) \tag{3.45}$$

and call it the *X-conditional probability of (the event) A* (with respect to P) (see SN-Remark 10.4). Furthermore, considering the event $\{Y=y\} = \{\omega \in \Omega\colon Y(\omega) = y\}$, we also use the notation

$$P(Y=y|X) \;:=\; P(\{Y=y\}|X) \;=\; E(\mathbb{1}_{Y=y}|X). \tag{3.46}$$

Note again that $P(A|X)$ and $P(Y=y|X)$ are X-measurable random variables on $(\Omega,\mathscr{A},P)$. ◁

Remark 3.59 (Multivariate X) If $X=(X_1,\ldots,X_n)$ is an n-variate random variable on the probability space $(\Omega,\mathscr{A},P)$ (see SN-section 5.3), then a version of $E(Y|X)$ is also denoted by $E(Y|X_1,\ldots,X_n)$. In the same vein, if $Y=\mathbb{1}_A$ or $Y=\mathbb{1}_{X=x}$, then we use the notation $P(A|X_1,\ldots,X_n)$ and $P(Y=y|X_1,\ldots,X_n)$, respectively. ◁

Many important properties of a conditional expectation $E(Y|X)$ are gathered in SN-Box 10.2. Some monotonicity properties of a conditional expectation $E(Y|X)$ are gathered in SN-Box 10.3. A proposition on strict monotonicity is stated in the following lemma. Reading this lemma, remember that $Y \gtrdot_P \alpha$ is defined by

$$P\big(\{\omega\in\Omega\colon Y(\omega)\le\alpha\}\big) \;=\; 0. \tag{3.47}$$

Lemma 3.60 (Strict Monotonicity of a Conditional Expectation)
Let $Y\colon\Omega\to\overline{\mathbb{R}}$ be a numerical random variable on $(\Omega,\mathscr{A},P)$ that is nonnegative or has a finite expectation, let X be a random variable on $(\Omega,\mathscr{A},P)$ with value space $(\Omega'_X,\mathscr{A}'_X)$, and let $\alpha\in\mathbb{R}$. Then:

$$Y \gtrdot_P \alpha \quad\Rightarrow\quad E(Y|X) \gtrdot_P \alpha. \tag{3.48}$$

(Proof p. 74)

3.3.2 First Examples

Example 3.61 (Joe and Ann With Randomized Assignment) Consider again the example presented in Table 3.1. In this table we already displayed the values .45 and .6 of the conditional expectation $E(Y|X)=P(Y=1|X)$. This random variable satisfies conditions (a) and (b) of Definition 3.54 (see Exercise 3-5). In this example, there is only one single version of the X-conditional expectation of Y. That is, the set $\mathscr{E}(Y|X)$ has only one single element. Hence, in this example, $E(Y|X)$ is uniquely defined. This also holds for the conditional expectation $E(Y|X,U)$ specified in the same table. ◁

Example 3.62 (No Treatment for Joe) Table 3.2 displays another example illustrating the concept of a conditional expectation. In this example there are uncountably many versions of the (X,U)-conditional expectation of Y. The column headed by $E(Y|X,U)$ displays *one* such version. Another one is obtained if, in this column, we replace the value 99 by any other number. Suppose $V=E(Y|X,U)$ is the version displayed in the table and $V^*\in\mathscr{E}(Y|X,U)$ is another version obtained by assigning to ω_3 and ω_4 the (arbitrarily chosen) number .8, leaving the assignments to the other $\omega_i\in\Omega$ untouched. Obviously, $V \underset{P}{=} V^*$. This illustrates that, in this example, the (X,U)-conditional expectation of Y is not uniquely defined. However, it is P-unique, because $V \underset{P}{=} V^*$, for all $V,V^*\in\mathscr{E}(Y|X,U)$. In contrast to $E(Y|X,U)$, the conditional expectations $E(Y|X)$ and $E(X|U)$ are uniquely

defined in this example. In other words, the sets $\mathscr{E}(Y|X)$ and $\mathscr{E}(X|U)$ consist of only one single element, the random variables $E(Y|X)$ and $E(X|U)$, respectively, which are specified in Table 3.2. ◁

3.3.3 Regression and Factorization of a Conditional Expectation

According to the following corollary, a version $E(Y|X) \in \mathscr{E}(Y|X)$ can always be written as a composition $g(X)$ (sometimes also denoted by $g \circ X$) of X and a numerical function g (see SN-Cor. 10.23).

Corollary 3.63 (Existence of the Factorization)
If the assumptions of Definition 3.54 hold and $E(Y|X) \in \mathscr{E}(Y|X)$, then there is a function $g\colon \Omega'_X \to \overline{\mathbb{R}}$ such that $g^{-1}(\overline{\mathscr{B}}) \subset \mathscr{A}'_X$ and

$$E(Y|X) \;=\; g(X). \tag{3.49}$$

The function g occurring in Corollary 3.63 plays in important role. Among other things, it is used for a general definition of the *regression of Y on X*.

Definition 3.64 (Factorization and Regression)
A function $g\colon \Omega'_X \to \overline{\mathbb{R}}$ that satisfies $g^{-1}(\overline{\mathscr{B}}) \subset \mathscr{A}'_X$ and Equation (3.49) is called a factorization of $E(Y|X)$. If $(\Omega'_X, \mathscr{A}'_X) = (\overline{\mathbb{R}}^n, \overline{\mathscr{B}}_n)$, $n \in \mathbb{N}$, then g is also called an n-variate regression of Y on X.

Remark 3.65 (Conditional Expectation vs. Regression) Hence, while a conditional expectation $E(Y|X) = g(X)$ is a random variable on $(\Omega, \mathscr{A}, P)$ with domain Ω, a factorization g is a function with domain Ω'_X. In fact, g is a numerical random variable on the probability space $(\Omega'_X, \mathscr{A}'_X, P_X)$. If $\Omega'_X = \overline{\mathbb{R}}^n$, then a factorization g is also called a *regression of Y on X*, or a *regression of Y on $X_1, \ldots, X_n$*. As is true for a conditional expectation, there can be many versions of a factorization and a regression. Note that the concept of a regression is defined without any reference to a specific parameterization. (In contrast, cf. SN-Def. 12.32 for the concept of a linear regression). ◁

3.3.4 Conditional Expectation Value $E(Y|X{=}x)$

A factorization of $E(Y|X)$ is also used for general definition of the concept of an $(X{=}x)$-conditional expectation value. In contrast to the elementary definition in Equation (3.27), in this definition we do *not* have to assume $P(X{=}x) > 0$.

Definition 3.66 ($(X{=}x)$-Conditional Expectation Value)
Let $g\colon \Omega'_X \to \overline{\mathbb{R}}$ be a function satisfying $g^{-1}(\overline{\mathscr{B}}) \subset \mathscr{A}'_X$ and Equation (3.49). Then, for all $x \in \Omega'_X$, we define an $(X{=}x)$-conditional expectation value of Y by

$$E(Y\,|\,X{=}x) \;:=\; g(x). \tag{3.50}$$

Remark 3.67 (Values of a Conditional Expectation) Let $E(Y|X) = g(X) \in \mathscr{E}(Y|X)$. Then

$$\forall\, \omega \in \Omega\colon \quad E(Y|X)(\omega) \;=\; g(x) \;=\; E(Y|X{=}x), \quad \text{if } \omega \in \{X{=}x\} \tag{3.51}$$

(see SN-Rem. 10.37). This equation also implies that the value of the random variable $E(Y|X)$ is constant on all sets $\{X{=}x\} = \{\omega \in \Omega\colon X(\omega) = x\}$. ◁

Remark 3.68 (Uniqueness of a Factorization) Note that $E(Y|X{=}x)$ is uniquely defined only if $P(X{=}x) > 0$. Similary, the concept of a regression of Y on X is not uniquely defined. For two elements $V, V^* \in \mathscr{E}(Y|X)$ there can be different factorizations g and g^* with $V = g(X)$ and $V^* = g^*(X)$. This is true even if $V = V^*$. Hence, there can be different factorizations of a single element $V \in \mathscr{E}(Y|X)$ (see SN-Example 10.32). In other words, $V = g(X) = g^*(X)$ with $g \neq g^*$ is not necessarily contradictory. If $g(X) = g^*(X)$ with $g \neq g^*$, then $g(x) = g^*(x)$ for all $x \in X(\Omega)$, whereas $g(x) = g^*(x)$ does *not* hold for all $x \in \Omega'_X$. However, SN-Theorem 10.9 (ii) and SN-Corollary 5.25 (i) imply the following corollary: ◁

Corollary 3.69 (P_X-Equivalence of Factorizations)
Let the assumptions of Definition 3.54 hold and let $g, g^*\colon \Omega'_X \to \overline{\mathbb{R}}$ be functions such that $g^{-1}(\overline{\mathscr{B}}), g^{*-1}(\overline{\mathscr{B}}) \subset \mathscr{A}'_X$. Furthermore, let $g(X), g^*(X) \in \mathscr{E}(Y|X)$. Then

$$g \underset{P_X}{=} g^*. \tag{3.52}$$

Hence, if $g(X), g^*(X) \in \mathscr{E}(Y|X)$, then the factorizations g and g^* are identical with probability 1 with respect to the distribution P_X of X. Note that P_X is a probability measure on $(\Omega'_X, \mathscr{A}'_X)$ (see Rem. 3.34).

Remark 3.70 (Properties of a Conditional Expectation) Many properties of conditional expectations are presented, proved, and illustrated in SN-chapters 10 and 11. Some special cases are treated in SN-chapters 12 and 13. The most important of these properties are gathered in SN-Box 10.2 and SN-Box 11.1. ◁

Remark 3.71 (Mean Independence) Under the assumptions of Definition 3.54 we define *mean independence of Y from X* by

$$E(Y|X) \underset{P}{=} E(Y), \tag{3.53}$$

and use $Y \underset{P}{\vdash} X$ as a short cut. Similarly, if Z is another random variable on the same probability space as X and Y, then we define *Z-conditional mean independence of Y from X*, by

$$E(Y|X,Z) \underset{P}{=} E(Y|Z), \tag{3.54}$$

and use the symbol $Y \underset{P}{\vdash} X\,|\,Z$ as a short cut for this equation. If there is no ambiguity with respect to the probability measure, then we also may omit the reference to the measure P and simply write $Y \vdash X$ and $Y \vdash X\,|\,Z$, respectively. Hence,

$$Y \vdash X \quad :\Leftrightarrow \quad E(Y|X) \underset{P}{=} E(Y) \tag{3.55}$$

and

$$Y \vdash X\,|\,Z \quad :\Leftrightarrow \quad E(Y|X,Z) \underset{P}{=} E(Y|Z)\,. \tag{3.56}$$

For more details on conditional mean independence see SN-section 10.6. ◁

3.3.5 Partial Conditional Expectation $E(Y|X{=}x,Z)$

According to Remark 3.59, the term $E(Y|X,Z)$ denotes the (X,Z)-conditional expectation of Y with respect to P. The concept of a *partial* $(X{=}x,Z)$*-conditional expectation* of Y builds on Corollary 3.63, according to which, for each version $E(Y|X,Z) \in \mathscr{E}(Y|X,Z)$, there is a function $g\colon \Omega'_X \times \Omega'_Z \to \overline{\mathbb{R}}$ such that

$$E(Y|X,Z) \;=\; g(X,Z), \tag{3.57}$$

where $g(X,Z)$ denotes the composition of the multivariate random variable (X,Z) and g. According to Equation (3.50), for $(x,z) \in \Omega'_X \times \Omega'_Z$,

$$E(Y\,|\,X{=}x,Z{=}z) \;=\; g(x,z), \tag{3.58}$$

is an $(X{=}x,Z{=}z)$-conditional expectation value of Y.

In Definition 3.73, we will refer to the function $g_x\colon \Omega'_Z \to \overline{\mathbb{R}}$ that, for $x \in \Omega'_X$, is defined by

$$g_x(z) \;=\; g(x,z), \quad \forall\, z \in \Omega'_Z. \tag{3.59}$$

Hence, a value $g_x(z)$ is identical to $E(Y\,|\,X{=}x,Z{=}z)$, that is,

$$g_x(z) \;=\; g(x,z) \;=\; E(Y\,|\,X{=}x,Z{=}z), \quad \forall\, z \in \Omega'_Z. \tag{3.60}$$

Remark 3.72 (Uniqueness of $E(Y\,|\,X{=}x,Z{=}z)$) Note that in Equations (3.59) and (3.60) we do not assume $P(X{=}x,Z{=}z) > 0$. However, there can be several versions $E(Y|X,Z) \in \mathscr{E}(Y|X,Z)$, and even for a given version $E(Y|X,Z) \in \mathscr{E}(Y|X,Z)$, there can be several factorizations satisfying Equation (3.57) (see Rem. 3.68). This implies that $E(Y\,|\,X{=}x,Z{=}z)$ is not uniquely defined if $P(X{=}x,Z{=}z) = 0$. (For more details see SN-section 10.4.4). ◁

Definition 3.73 (Partial Conditional Expectation)
Let X, Y, and Z be random variables on $(\Omega,\mathscr{A},P)$ with value spaces $(\Omega'_X,\mathscr{A}'_X)$, $(\overline{\mathbb{R}},\overline{\mathscr{B}})$, and $(\Omega'_Z,\mathscr{A}'_Z)$, respectively. Furthermore, assume that Y is nonnegative or with finite expectation $E(Y)$, let $E(Y|X,Z) \in \mathscr{E}(Y|X,Z)$ and let g be a factorization of $E(Y|X,Z)$ such that $g(X,Z) = E(Y|X,Z)$. Finally, for $x \in \Omega'_X$, let the function g_x be defined by Equation (3.59). *Then we call the function $E(Y|X{=}x,Z)\colon \Omega \to \overline{\mathbb{R}}$ defined by*

$$E(Y|X{=}x,Z) \;:=\; g_x(Z) \tag{3.61}$$

a version of the partial $(X{=}x,Z)$-conditional expectation of Y (with respect to P).

Remark 3.74 (A Partial Conditional Expectation is a Random Variable) For all $x \in \Omega'_X$, the function $E(Y|X{=}x,Z) = g_x(Z)$ denotes the composition of Z and g_x. Hence, for all $x \in \Omega'_X$, $E(Y|X{=}x,Z)$ is a Z-measurable random variable on $(\Omega,\mathscr{A},P)$ (see Rem. 3.32 and SN-Lemma 2.52). ◁

Remark 3.75 (Partial Conditional Probability) If $A \in \mathscr{A}$, then we also use the notation $P(A|\,X{=}x,Z) := E(1_A|\,X{=}x,Z)$ and call it a *partial* $(X{=}x,Z)$*-conditional probability* of (the event) A (with respect to P). Furthermore, if Y is binary with values 0 and 1, then we use the notation $P(Y{=}1|X{=}x,Z) := E(Y|X{=}x,Z)$ and call it a *partial* $(X{=}x,Z)$*-conditional probability* of (the event) $\{Y{=}1\}$ (with respect to P). ◁

Remark 3.76 (The Set $\mathscr{E}(Y|X{=}x,Z)$) Note that $E(Y|X{=}x,Z)$ is not uniquely defined for two reasons. The first is that $E(Y|X,Z)$ is not uniquely defined (see Rem. 3.55). The second reason is that even for a given version $E(Y|X,Z)$, the factorization g of $E(Y|X,Z) = g(X,Z)$ is not uniquely defined (see Rem. 3.68 and Cor. 3.69). Therefore, we use

$$\mathscr{E}(Y|X{=}x,Z) := \big\{g_x(Z)\colon g_x \text{ satisfies (3.60), where } g(X,Z) \in \mathscr{E}(Y|X,Z)\big\} \tag{3.62}$$

to denote the set of all versions of the partial $(X{=}x,Z)$-conditional expectation of Y. ◁

Remark 3.77 (Discrete X) Under the assumptions of Definition 3.73, suppose that X is a random variable on $(\Omega,\mathscr{A},P)$ with value space $(\Omega'_X,\mathscr{A}'_X)$ and the image $X(\Omega)$ of Ω under X is finite or countable with $\{x\} \in \mathscr{A}'_X$ for all $x \in X(\Omega)$. Then

$$E(Y|X,Z) = \sum_{x\in X(\Omega)} E(Y|X{=}x,Z)\cdot 1_{X=x} \tag{3.63}$$

holds *for the specific version* $E(Y|X,Z) \in \mathscr{E}(Y|X,Z)$ that is used in Definition 3.73 (for a proof see SN-Exercise 14-6). Furthermore, *for all versions* $V \in \mathscr{E}(Y|X,Z)$,

$$V \underset{P}{=} \sum_{x\in X(\Omega)} E(Y|X{=}x,Z)\cdot 1_{X=x}. \tag{3.64}$$

◁

Remark 3.78 (A Partial Conditional Expectation is not a Conditional Expectation) Note again, a partial conditional expectation is defined even if $P(X{=}x) = 0$. However, this definition is not unique (see Rem. 3.76). Also note that, in general, a partial conditional expectation is not a conditional expectation (with all its well-known properties) unless $P(X{=}x) > 0$. In the latter case a partial conditional expectation is in fact a version of a conditional expectation (see Th. 3.91). This is detailed in section 3.3.6. ◁

3.3.6 Conditional Expectation $E^{X=x}(Y|Z)$ With Respect to $P^{X=x}$

Presuming $P(X{=}x) > 0$, now we introduce the conditional expectation of Y given Z with respect to the $(X{=}x)$-conditional probability measure $P^{X=x}$. Note that $P^{X=x}$ is just a special case of the measure P^B introduced in Definition 3.17 for $B = \{X{=}x\} = \{\omega \in \Omega\colon X(\omega) = x\}$. Therefore, the expectation $E^{X=x}(Y)$ of a numerical random variable Y with respect to the measure $P^{X=x}$ has already been introduced in Equation (3.22).

Definition 3.79 (Z-Conditional Expectation With Respect to $P^{X=x}$)
Let X, Y, and Z be random variables on $(\Omega,\mathscr{A},P)$ with value spaces $(\Omega'_X,\mathscr{A}'_X)$, $(\overline{\mathbb{R}},\overline{\mathscr{B}})$, and $(\Omega'_Z,\mathscr{A}'_Z)$ respectively. Assume that Y is nonnegative or with finite expectation $E^{X=x}(Y)$ and $x \in \Omega'_X$ is a value of X such that $\{x\} \in \mathscr{A}'_X$ and $P(X{=}x) > 0$. Then a random variable $V_x\colon \Omega \to \overline{\mathbb{R}}$ on $(\Omega,\mathscr{A},P)$ is called a (version of the) Z-conditional expectation of Y with respect to $P^{X=x}$, if the following two conditions hold:

(a) $\sigma(V_x) \subset \sigma(Z)$.
(b) $E^{X=x}(1_C \cdot V_x) = E^{X=x}(1_C \cdot Y), \quad \forall\, C \in \sigma(Z)$.

If V_x satisfies (a) and (b), then we also use the notation $E^{X=x}(Y|Z) := V_x$.

Remark 3.80 ($E^{X=x}(Y|Z)$ is a Conditional Expectation) Comparing Definitions 3.54 and 3.79 to each other shows that we only replaced the measure P, which is used to define an expectation and a conditional expectation with respect to P by the measure $P^{X=x}$, which is used to define an expectation and a conditional expectation with respect to the measure $P^{X=x}$. Hence, $E^{X=x}(Y|Z)$ has all properties of a Z-conditional expectation of Y, provided that we replace the measure P by $P^{X=x}$. For example, the property $E\big(E(Y|X)\big) = E(Y)$ has to be translated to

$$E^{X=x}\big(E^{X=x}(Y|Z)\big) \;=\; E^{X=x}(Y)$$

[see SN-Box 10.2 (iv)]. Similarly, the property $E\big(E(Y|X)\,\big|\,f(X)\big) \underset{P}{=} E\big(Y\,\big|\,f(X)\big)$ has to be translated to

$$E^{X=x}\big(E^{X=x}(Y|Z)\,\big|\,f(Z)\big) \underset{P^{X=x}}{=} E^{X=x}\big(Y\,\big|\,f(Z)\big)$$

[see SN-Box 10.2 (v)]. ◁

Remark 3.81 (The Set $\mathscr{E}^{X=x}(Y|Z)$) As is true for any conditional expectation, there can be more than one single version of the Z-conditional expectation of Y with respect to the conditional probability measure $P^{X=x}$. Therefore, we use $\mathscr{E}^{X=x}(Y|Z)$ to denote the set of all versions of the Z-conditional expectation of Y with respect to $P^{X=x}$. ◁

Remark 3.82 ($P^{X=x}$-Uniqueness of $E^{X=x}(Y|Z)$) However,

$$\forall\, V_x, V_x^* \in \mathscr{E}^{X=x}(Y|Z)\colon \quad P^{X=x}\big(\{\omega\in\Omega\colon V_x(\omega) = V_x^*(\omega)\}\big) \;=\; 1 \tag{3.65}$$

(cf. Rem. 3.56). This is what we mean saying that $E^{X=x}(Y|Z)$ is $P^{X=x}$-*unique*. ◁

Remark 3.83 ($E^{X=x}(Y|Z)$ is not Necessarily P-Unique) Note that

$$\forall\, V_x, V_x^* \in \mathscr{E}^{X=x}(Y|Z)\colon \quad P\big(\{\omega\in\Omega\colon V_x(\omega) = V_x^*(\omega)\}\big) \;=\; 1 \tag{3.66}$$

does *not necessarily hold*. In other words, although $E^{X=x}(Y|Z)$ is always $P^{X=x}$-unique, it is not necessarily P-unique. ◁

In the following theorem, which is an adaptation of SN-Corollary 14.48, we present conditions that are equivalent to P-uniqueness of $E^{X=x}(Y|Z)$. In this theorem, we use the notation

$$P(X{=}x|Z) \;\underset{P}{\geqslant}\; 0 \quad :\Leftrightarrow \quad P\big(\{\omega\in\Omega\colon P(X{=}x|Z)(\omega) > 0\}\big) \;=\; 1$$

and

$$P \underset{\sigma(Z)}{\ll} P^{X=x} \quad :\Leftrightarrow \quad \forall A\in\sigma(Z)\colon \big(P^{X=x}(A) = 0 \;\Rightarrow\; P(A) = 0\big).$$

Hence, $P \underset{\sigma(Z)}{\ll} P^{X=x}$ means that the measure P is absolutely continuous on the σ-algebra $\sigma(Z)$ with respect to the measure $P^{X=x}$ (cf. Rem. 3.22).

Theorem 3.84 (P-Uniqueness of $E^{X=x}(Y|Z)$)
Let the assumptions of Definition 3.79 hold. Then the following propositions are equivalent to each other.

(a) $E^{X=x}(Y|Z)$ *is P-unique.*

(b) $P(X=x|Z) \gneq_P 0$.

(c) $P \underset{\sigma(Z)}{\ll} P^{X=x}$.

(d) $P_Z \ll P_{Z|X=x}$.

Furthermore, if there is a version $V_x \in \mathscr{E}^{X=x}(Y|Z)$ such that $E(V_x)$ is finite, then each of (a) to (d) is also equivalent to

(e) $\forall\, V_x, V_x^* \in \mathscr{E}^{X=x}(Y|Z)\colon\ E(V_x) = E(V_x^*)$.

Example 3.85 (No Treatment for Joe) The last two columns of Table 3.2 display versions of the conditional expectations $E^{X=0}(Y|U)$ and $E^{X=1}(Y|U)$, respectively. In this example, $E^{X=0}(Y|U)$ is $P^{X=0}$-unique and P-unique. It is even uniquely defined. In contrast, $E^{X=1}(Y|U)$ is not uniquely defined and it also not P-unique. However, it is $P^{X=1}$-unique, because $P(X=1\,|\,U)(\omega_3) = P(X=1\,|\,U)(\omega_4) = 0$. Hence, in this example, condition (b) of Theorem 3.84 does not hold. Exchanging in the last column of this table the number 99 by any other real number, say .8, yields another version of the U-conditional expectation of Y with respect to $P^{X=1}$. If the two versions are denoted V_1 and V_1^*, then $P^{X=1}(\{\omega \in \Omega\colon V_1(\omega) = V_1^*(\omega)\}) = 1$, that is, V_1 and V_1^* are identical almost surely with respect to the measure $P^{X=1}$ (see Exercise 3-6). ◁

Remark 3.86 (Factorization of $E^{X=x}(Y|Z)$) If the assumptions of Definition 3.79 hold and $E^{X=x}(Y|Z) \in \mathscr{E}^{X=x}(Y|Z)$, then there is a function $g_x\colon \Omega_Z' \to \overline{\mathbb{R}}$ such that $g_x^{-1}(\overline{\mathscr{B}}) \subset \mathscr{A}_Z'$ and

$$E^{X=x}(Y|Z) \;=\; g_x(Z) \tag{3.67}$$

(see Cor. 3.63). The function g_x is a *factorization of* $E^{X=x}(Y|Z)$ (see Def. 3.64). Note that a factorization is a random variable on $(\Omega_Z', \mathscr{A}_Z', P_Z)$, whereas $E^{X=x}(Y|Z)$ and the composition $g_x(Z)$ are random variables on $(\Omega, \mathscr{A}, P)$. ◁

A factorization of $E^{X=x}(Y|Z)$ can also be used for the definition of a $(Z=z)$-conditional expectation value of Y with respect to the conditional probability measure $P^{X=x}$. In this definition we assume $P(X=x) > 0$, but not $P(Z=z) > 0$.

Definition 3.87 ($(X=x)$-Conditional Expectation Value)
Let $g_x\colon \Omega_Z' \to \overline{\mathbb{R}}$ be a function satisfying $g_x^{-1}(\overline{\mathscr{B}}) \subset \mathscr{A}_Z'$ and Equation (3.67). Then, for all $z \in \Omega_Z'$, we define a $(Z=z)$-conditional expectation value of Y with respect to the probability measure $P^{X=x}$ by

$$E^{X=x}(Y\,|\,Z=z) \;:=\; g_x(z). \tag{3.68}$$

Remark 3.88 (Values of $E^{X=x}(Y|Z)$) The values of $E^{X=x}(Y|Z)$ are the conditional expectation values $E^{X=x}(Y|Z=z)$. In more formal terms,

$$\forall\,\omega\in\Omega:\quad E^{X=x}(Y|Z)(\omega)\;=\;E^{X=x}(Y|Z=z),\quad \text{if } \omega\in\{Z=z\}. \tag{3.69}$$

Furthermore, if $P(X=x,Z=z)>0$, then

$$E^{X=x}(Y|Z=z)\;=\;E(Y|X=x,Z=z) \tag{3.70}$$

(see SN-Rem. 14.37). ◁

Example 3.89 (No Treatment for Joe) Now we compute the values of the conditional expectation $E^{X=0}(Y|U)$ in the example presented in Table 3.2. According to Remark 3.88, the values of the conditional expectation $E^{X=0}(Y|U)$ are the two conditional expectation values $E(Y|X=0,U=Joe)$ and $E(Y|X=0,U=Ann)$. Because $E(Y|X=0,U=u)=P(Y=1|X=0,U=u)$, they can be computed from the probabilities of the elementary events presented in Table 3.2 as follows:

$$P(Y=1|X=0,U=Joe)\;=\;\frac{P(Y=1,X=0,U=Joe)}{P(X=0,U=Joe)}\;=\;\frac{.35}{.15+.35}\;=\;.7$$

and

$$P(Y=1|X=0,U=Ann)\;=\;\frac{P(Y=1,X=0,U=Ann)}{P(X=0,U=Ann)}\;=\;\frac{.06}{.06+.24}\;=\;.2.$$

◁

Important properties related to P-uniqueness of a conditional expectation $E^{X=x}(Y|Z)$ with respect to $P^{X=x}$ are found in SN-sections 14.6.3 and 14.6.4. One of these properties is related to absolutely continuity of P with respect to the measure $P^{X=x}$ (see the remark to Lemma 3.21).

Remark 3.90 (Additional Properties) Aside from the properties of a conditional expectation (see SN-chapters 10 and 11), $E^{X=x}(Y|Z)$ has a number of properties due to the fact that $E^{X=x}(Y|Z)$ is a random variable on the probability space $(\Omega,\mathscr{A},P)$ and also a random variable on $(\Omega,\mathscr{A},P^{X=x})$. This follows from the fact that $(\Omega,\mathscr{A},P)$ and $(\Omega,\mathscr{A},P^{X=x})$ share the same measurable space $(\Omega,\mathscr{A})$ (see SN-Def. 5.1). Some of these properties are addressed in the sequel. ◁

According to the following theorem, the partial conditional expectation $E(Y|X=x,Z)$ is also a version of the Z-conditional expectation of Y with respect to $P^{X=x}$, provided that $P(X=x)>0$ (for a proof see SN-Theorem 14.33).

Theorem 3.91 (Relationship Between $E(Y|X=x,Z)$ and $E^{X=x}(Y|Z)$)
Let the assumptions of Definition 3.73 hold and suppose that $x\in\Omega'_X$ such that $\{x\}\in\mathscr{A}'_X$ and $P(X=x)>0$. Then

$$E(Y|X=x,Z)\;\in\;\mathscr{E}^{X=x}(Y|Z). \tag{3.71}$$

This implies

$$E(Y|X=x,Z)\underset{P^{X=x}}{=}E^{X=x}(Y|Z),\quad \forall\,E^{X=x}(Y|Z)\in\mathscr{E}^{X=x}(Y|Z). \tag{3.72}$$

Box 3.1 Glossary of new concepts

Ω	*Set of possible outcomes.*
$\mathscr{A}$	*Set of possible events.* A set of subsets of Ω satisfying the requirements of a σ-*algebra* on Ω.
P	*Probability measure on* $\mathscr{A}$. A function on $\mathscr{A}$ with values in the closed interval $[0,1]$ satisfying the Kolmogorov axioms of probability.
$(\Omega, \mathscr{A}, P)$	*Probability space.* It consists of the three components listed above. In empirical applications it already contains all information about the random experiment considered.
$P(A\,\vert\,B)$	Conditional probability of the event A given the event B. If $A, B \in \mathscr{A}$ and $P(B) > 0$, then $P(A\,\vert\,B) := P(A \cap B)/P(B)$.
P^B	*B-conditional probability measure on* $\mathscr{A}$. If $B \in \mathscr{A}$ and $P(B) > 0$, then $P^B(A) := P(A\,\vert\,B)$ for all $A \in \mathscr{A}$.
Y	*Random variable with value space* $(\Omega'_Y, \mathscr{A}'_Y)$. It is a mapping on Ω with values in the set Ω'_Y such that $Y^{-1}(A') \in \mathscr{A}$, for all $A' \in \mathscr{A}'_Y$, where $Y^{-1}(A') := \{\omega \in \Omega\colon Y(\omega) \in A'\}$ denotes the inverse image of A' under Y.
P_Y	*Distribution of a random variable* Y. It is a probability measure on $(\Omega'_Y, \mathscr{A}'_Y)$ defined by $P_Y(A') = P[Y^{-1}(A')]$, for all $A' \in \mathscr{A}'_Y$.
$E(Y\,\vert\,X)$	A (version of the) *X-conditional expectation of Y* with respect to the probability measure P. It is a random variable on $(\Omega, \mathscr{A}, P)$ whose values are the conditional expectation values $E(Y\,\vert\,X{=}x)$.
$E^{X=x}(Y\,\vert\,Z)$	A (version of the) *Z-conditional expectation of Y with respect to the conditional probability measure* $P^{X=x}$. It is defined by $P^{X=x} = P^{\{X=x\}}$ if X is a random variable on $(\Omega, \mathscr{A}, P)$ and $P(\{X{=}x\}) > 0$, where $\{X{=}x\} := \{\omega \in \Omega\colon X(\omega) = x\}$.

Remark 3.92 (Relationship Between $E^{X=x}(Y\,\vert\,Z)$ and $E(Y\,\vert\,X,Z)$) Let the assumptions of Definition 3.79 hold and assume that the image $X(\Omega)$ of Ω under X is finite or countable with $\{x\} \in \mathscr{A}'_X$ and $P(X{=}x) > 0$ for all $x \in X(\Omega)$. Then

$$V \underset{P}{=} \sum_{x \in X(\Omega)} E^{X=x}(Y\,\vert\,Z) \cdot \mathbf{1}_{X=x}, \quad \forall\, V \in \mathscr{E}(Y\,\vert\,X,Z)\,. \tag{3.73}$$

◁

3.4 Summary and Conclusions

In this chapter we reviewed the most important concepts of probability theory. The emphasis has been on the *structure* or *architecture* of these concepts. Box 3.1 provides a glossary. We started with the three components of a *probability space* $(\Omega, \mathscr{A}, P)$, the *set* Ω *of possible outcomes*, the σ-*algebra* $\mathscr{A}$, a set of subsets of Ω with certain properties, and the

probability measure P, a function on $\mathscr{A}$ with values in the closed interval $[0,1]$, also satisfying certain properties, the Kolmogorov axioms of probability.

We continued with the concept of a *random variable Y*, a mapping on Ω with value space $(\Omega'_Y, \mathscr{A}'_Y)$, where Ω'_Y is a set and $\mathscr{A}'_Y$ a σ-algebra on Ω'_Y. Such a pair of a set and a σ-algebra on this set is called a *measurable space*. The crucial property of a random variable is that all inverse images $Y^{-1}(A')$, $A' \in \mathscr{A}'_Y$, are elements of $\mathscr{A}$. This guarantees that all these inverse images have probabilities assigned by P, which is used to define the *distribution* P_Y of the random variable Y.

Then we turned to the concept of an *X-conditional expectation* $E(Y|X)$ of a numerical random variable Y, that is, a random variable with value space $(\overline{\mathbb{R}}, \overline{\mathscr{B}})$. This concept contains the information on how the conditional expectation values $E(Y|X=x)$ depend on the values x of X. Because X can be multidimensional, such a conditional expectation $E(Y|X)$ and its values $E(Y|X=x)$ are what we try to estimate in many statistical procedures, for example, in regression analysis, analysis of variance, structural equation modeling, multilevel analysis, etc.

The examples presented in this chapter and in chapter 1 show that under some conditions (such as an experiment with randomized assignment of an observational unit to a treatment), conditional expectations can be used to describe causal dependencies [see $E(Y|X)$ in Table 3.1], whereas in other conditions (such as systematic self-selection) they can totally lead us astray if such a causal interpretation of a conditional expectation is intended [see $E(Y|X)$ in Table 1.2].

3.5 Proofs

Proof of Lemma 3.60

SN-Box 10.3 (ii) implies $Y \underset{P}{\geq} \alpha \Rightarrow E(Y|X) \underset{P}{\geq} \alpha$. Hence, if $Y \underset{P}{\geq} \alpha$, then $V \underset{P}{\geq} \alpha$ for all $V \in \mathscr{E}(Y|X)$. Now consider the event $C \in \sigma(X)$ defined by $C := \{\omega \in \Omega\colon V(\omega) = \alpha\}$. We prove $P(C) = 0$ by contradiction. Hence, assume $P(C) > 0$. Then

$$\begin{aligned}
\alpha \cdot P(C) &= \int 1_C\, \alpha \, dP && \text{[SN-Th. 3.36 (i), SN-(3.9)]}\\
&= \int 1_C\, V \, dP && [V(\omega) = \alpha \text{ if } \omega \in C]\\
&= \int 1_C\, Y \, dP && \text{[SN-(6.1), Def. 3.54 (b)]}\\
&> \alpha \cdot P(C). && \text{[SN-Th. 3.52 (ii)]}
\end{aligned}$$

This contradiction proves $P(C) = 0$ and Proposition (3.48). Note that SN-Theorem 3.52 can be applied because $\left(C, \mathscr{A}\big|_C, P\big|_C\right)$ is a measure space, where $\mathscr{A}\big|_C$ denotes the trace σ-algebra of $\mathscr{A}$ in C (see SN-Example 1.10), and $P\big|_C$ denotes the restriction of P on C (see SN-Example 1.61).

3.6 Exercises

▷ **Exercise 3-1** Let Ω be nonempty. Show that $\{\Omega, \emptyset\}$ is a σ-algebra on Ω.

▷ **Exercise 3-2** Let Ω be nonempty and $A \subset \Omega$. Show that $\{A, A^c, \Omega, \emptyset\}$ is a σ-algebra on Ω.

▷ **Exercise 3-3** Consider the random variable X in Example 3.28. Instead of $(\mathbb{R},\mathscr{B})$, choose $\big(\{0,1,2\}, \mathscr{P}(\{0,1,2\})\big)$ as the value space of X. Write down all inverse images $X^{-1}(B)$, $B \in \mathscr{P}(\{0,1,2\})$.

▷ **Exercise 3-4** Why is the conditional expectation value $E(Y|X=1, U=Joe)$ not uniquely defined in the example presented in Table 3.2?

▷ **Exercise 3-5** Consider the example presented in Table 3.1. Show that the random variable $E(Y|X) = P(Y=1|X)$ specified in this table satisfies conditions (a) and (b) of Definition 3.54.

▷ **Exercise 3-6** Specify the conditional probability measure $P^{X=1}$ in the example presented in Table 3.2.

▷ **Exercise 3-7** Consider example 3.85 and show that $P^{X=1}(\{\omega \in \Omega\colon V_1(\omega) = V_1^*(\omega)\}) = 1$, using the conditional probability measure specified in Exercise 3-6.

▷ **Exercise 3-8** What does it mean when we assume that the conditional expectation $E^{X=x}(Y|Z)$ is P-unique?

▷ **Exercise 3-9** Which are the values of $E^{X=0}(Y|U)$ and of the conditional expectation $E(Y|X,U)$ for $\omega_4 = (Joe, yes, +)$ in the example presented in Table 3.1?

▷ **Exercise 3-10** Which are the values of a Z-conditional expectation $E^{X=x}(Y|Z)$ of Y with respect to the conditional probability measure $P^{X=x}$?

▷ **Exercise 3-11** Compute the values of the conditional expectation $E(Y|X)$ in the example presented in Table 3.2.

▷ **Exercise 3-12** Compute the values of the conditional expectation $E(Y|X,U)$ in the example presented in Table 3.2.

Solutions

▷ **Solution 3-1** Obviously, conditions (a) and (b) of Definition 3.3 are satisfied, because $\Omega^c = \emptyset$ and $\emptyset^c = \Omega$, and $\Omega, \emptyset \in \{\Omega, \emptyset\}$. Condition (c) is satisfied as well. If Ω is at least one of the elements $A_1, A_2, \ldots \in \{\Omega, \emptyset\}$, then $\bigcup_{i=1}^{\infty} A_i = \Omega$, which is an element of $\{\Omega, \emptyset\}$. If Ω is not one of the elements $A_1, A_2, \ldots \in \{\Omega, \emptyset\}$, that is, if $A_1, A_2, \ldots = \emptyset, \emptyset, \ldots$, then $\bigcup_{i=1}^{\infty} A_i = \emptyset$, which is an element of $\{\Omega, \emptyset\}$, too.

▷ **Solution 3-2** Obviously, condition (a) of Definition 3.3 is satisfied, because $\Omega \in \{A, A^c, \Omega, \emptyset\}$. Condition (b) of Definition 3.3 is satisfied as well, because the complement of each element of $\{A, A^c, \Omega, \emptyset\}$ is also an element of this set. Finally, condition (c) is satisfied as well. If Ω is at least one of the elements $A_1, A_2, \ldots \in \{A, A^c, \Omega, \emptyset\}$ or if A and A^c are among the $A_1, A_2, \ldots$, then $\bigcup_{i=1}^{\infty} A_i = \Omega$, which is an element of $\{A, A^c, \Omega, \emptyset\}$. If neither Ω nor A or A^c are one of the elements $A_1, A_2, \ldots \in \{A, A^c, \Omega, \emptyset\}$, that is, if $A_1, A_2, \ldots = \emptyset, \emptyset, \ldots$, then $\bigcup_{i=1}^{\infty} A_i = \emptyset$, which is an element of $\{A, A^c, \Omega, \emptyset\}$, too. If A is among the elements $A_1, A_2, \ldots \in \{A, A^c, \Omega, \emptyset\}$ but neither A^c nor Ω, then $\bigcup_{i=1}^{\infty} A_i = A$, which is an element of $\{A, A^c, \Omega, \emptyset\}$. Finally, if A^c is among the elements $A_1, A_2, \ldots \in \{A, A^c, \Omega, \emptyset\}$ but neither A nor Ω, then $\bigcup_{i=1}^{\infty} A_i = A^c$, which is an element of $\{A, A^c, \Omega, \emptyset\}$.

▷ **Solution 3-3** These inverse image are identical to those listed in Equation (3.12).

▷ **Solution 3-4** In this example, $P(X=1, U=Joe) = 0$. This implies that the conditional probabilities $P(Y=y|X=1, U=Joe)$ that are used in the definition of $E(Y|X=1, U=Joe)$ [see Eqs. (3.27) and (3.29)] are not defined.

▷ **Solution 3-5** The σ-algebra generated by $E(Y|X)$ consists of the following four inverse images:

$$\forall B \in \mathscr{B}\colon \quad E(Y|X)^{-1}(B) \;=\; \begin{cases} \Omega_U \times \{no\} \times \Omega_Y = \{\omega_1, \omega_2, \omega_5, \omega_6\}, & \text{if } .6 \notin B \text{ and } .45 \in B \\ \Omega_U \times \{yes\} \times \Omega_Y = \{\omega_3, \omega_4, \omega_7, \omega_8\}, & \text{if } .6 \in B \text{ and } .45 \notin B \\ \Omega = \{\omega_1, \ldots, \omega_8\}, & \text{if } .6 \in B \text{ and } .45 \notin B \\ \emptyset, & \text{if } .6 \notin B \text{ and } .45 \notin B \end{cases}$$

(see Table 3.1). These four inverse images are identical to the elements of $\sigma(X)$:

$$\forall B \in \mathscr{B}\colon \quad X^{-1}(B) \;=\; \begin{cases} C_1 = \Omega_U \times \{no\} \times \Omega_Y = \{\omega_1, \omega_2, \omega_5, \omega_6\}, & \text{if } 1 \notin B \text{ and } 0 \in B \\ C_2 = \Omega_U \times \{yes\} \times \Omega_Y = \{\omega_3, \omega_4, \omega_7, \omega_8\}, & \text{if } 1 \in B \text{ and } 0 \notin B \\ C_3 = \Omega = \{\omega_1, \ldots, \omega_8\}, & \text{if } 1 \in B \text{ and } 0 \notin B \\ C_4 = \emptyset, & \text{if } 1 \notin B \text{ and } 0 \notin B \end{cases}$$

(see again Table 3.1). Hence, $\sigma\big(E(Y|X)\big) = \sigma(X)$ implying that condition (a) of Definition 3.54 holds.

Now we show that condition (b) of Definition 3.54 holds as well. Using Equation (3.24) for the random variable $1_{C_1} \cdot E(Y|X)$ and the probabilities listed in Table 3.1 yields

$$E\big(1_{C_1} \cdot E(Y|X)\big) \;=\; .45 \cdot (.09 + .21 + .24 + .06) \;+\; .6 \cdot 0 \;+\; 0 \cdot (.04 + .16 + .12 + .08) \;=\; .27.$$

Using the same formula and the same table for $1_{C_1} \cdot Y$ yields

$$E(1_{C_1} \cdot Y) \;=\; 0 \cdot (.09 + .24 + .04 + .12 + .16 + .08) \;+\; 1 \cdot (.21 + .06) \;=\; .27.$$

Hence $E\big(1_{C_1} \cdot E(Y|X)\big) = E(1_{C_1} \cdot Y)$ holds for $C_1 \in \sigma(X)$. The analog computations for C_2, C_3, and C_4 show that $E\big(1_C \cdot E(Y|X)\big) = E(1_C \cdot Y)$ also holds for the other three $C \in \sigma(X)$. Hence, condition (b) of Definition 3.54 is satisfied as well. This proves that the random variable $E(Y|X)$ specified in Table 3.1 is in fact a version of the X-conditional expectation of Y.

▷ **Solution 3-6** The measure $P^{X=1}$ is defined if the probabilities $P^{X=1}(\{\omega_i\})$ are specified for all eight elementary events $\{\omega_1\}, \ldots, \{\omega_8\}$. The probabilities of all other events can be computed from the probabilities of the elementary events [see SN-Box 4.1 (x)]. Hence, we just have to specify the following probabilities:

$$\text{For } i = 1, \ldots, 6\text{:} \quad P^{X=1}(\{\omega_i\}) \;=\; \frac{P(\{\omega_i\} \cap \{X=1\})}{P(X=1)} \;=\; \frac{0}{.12 + .08} \;=\; 0.$$

$$\text{For } i = 7\text{:} \quad P^{X=1}(\{\omega_i\}) \;=\; \frac{P(\{\omega_i\} \cap \{X=1\})}{P(X=1)} \;=\; \frac{.12}{.12 + .08} \;=\; .6.$$

$$\text{For } i = 8\text{:} \quad P^{X=1}(\{\omega_i\}) \;=\; \frac{P(\{\omega_i\} \cap \{X=1\})}{P(X=1)} \;=\; \frac{.08}{.12 + .08} \;=\; .4.$$

▷ **Solution 3-7** The values of V_1 and V_1^* only differ for the outcomes $\omega_1, \ldots, \omega_4$. According to Exercise 3-6, $P^{X=1}(\{\omega_1, \ldots, \omega_4\}) = 0$. Hence, $P^{X=1}(\{\omega \in \Omega\colon V_1(\omega) = V_1^*(\omega)\}) = 1$.

▷ **Solution 3-8** By definition, there may be different versions $E^{X=x}(Y|Z)$ of the Z-conditional expectation of Y with respect to the probability measure $P^{X=x}$. In general, all pairs of such versions are identical with probability $P^{X=x} = 1$ [see Eq. (3.65)]. If we additionally assume that $E^{X=x}(Y|Z)$ is P-unique, then we assume that all pairs of versions $V_x, V_x^* \in \mathscr{E}^{X=x}(Y|Z)$ are identical with probability $P = 1$ [see Eq. (3.66)].

▷ **Solution 3-9** $E^{X=0}(Y|U)(\omega_4) = E(Y|X=0, U=Joe) = .7$. In contrast, the value of $E(Y|X,U)$ is $E(Y|X,U)(\omega_4) = E(Y|X=1, U=Joe) = .8$ (see the fourth row in Table 3.1).

▷ **Solution 3-10** The values of $E^{X=x}(Y|Z)$ are identical with the conditional expectation values $E(Y|X{=}x, Z{=}z)$. In more formal terms, $E^{X=x}(Y|Z)(\omega) = E(Y|X{=}x, Z{=}z)$, if $\omega \in (X,Z)^{-1}(\{(x,z)\})$.

▷ **Solution 3-11** The values of the conditional expectation $E(Y|X)$ are the two conditional expectation values $E(Y|X=0)$ and $E(Y|X=1)$. Because $E(Y|X{=}x) = P(Y{=}1|X{=}x)$, they can be computed from Table 3.2 as follows:

$$P(Y{=}1|X=0) = \frac{P(Y{=}1, X=0)}{P(X=0)} = \frac{.35+.06}{.15+.35+.24+.06} = .5125,$$

and

$$P(Y{=}1|X=1) = \frac{P(Y{=}1, X=1)}{P(X=1)} = \frac{0+.08}{0+0+.12+.08} = .4.$$

▷ **Solution 3-12** The values of the conditional expectation $E(Y|X,U)$ in Table 3.2 are the four conditional expectation values $E(Y|X{=}x, U{=}u)$. Because $E(Y|X{=}x, U{=}u) = P(Y{=}1|X{=}x, U{=}u)$, they can be computed as follows:

$$P(Y{=}1|X=0, U{=}\mathit{Joe}) = \frac{P(Y{=}1, X=0, U{=}\mathit{Joe})}{P(X=0, U{=}\mathit{Joe})} = \frac{.35}{.15+.35} = .7,$$

$$P(Y{=}1|X=0, U{=}\mathit{Ann}) = \frac{P(Y{=}1, X=0, U{=}\mathit{Ann})}{P(X=0, U{=}\mathit{Ann})} = \frac{.06}{.24+.06} = .2,$$

$$P(Y{=}1|X=1, U{=}\mathit{Ann}) = \frac{P(Y{=}1, X=1, U{=}\mathit{Ann})}{P(X=1, U{=}\mathit{Ann})} = \frac{.08}{.12+.08} = .4.$$

The conditional expectation value $E(Y|X=1, U{=}\mathit{Joe}) = P(Y{=}1|X=1, U{=}\mathit{Joe})$ is undefined, because $P(X=1, U{=}\mathit{Joe}) = 0$. Choosing any number (such as 99) as a value of $E(Y|X,U)$ for $\omega_3 = (\mathit{Joe}, \mathit{yes}, -)$ and $\omega_4 = (\mathit{Joe}, \mathit{yes}, +)$ yields a version of $E(Y|X,U)$. Different versions of $E(Y|X,U)$ are identical almost surely with respect to the measure P [see Eq. (3.44)].

Part II
Basic Concepts

Chapter 4
Potential Confounder and True Outcome Variable

In chapter 1 it has been shown that the conditional expectation values $E(Y|X{=}x)$ of an outcome variable Y and their differences $E(Y|X{=}x) - E(Y|X{=}x')$, the *prima facie effects*, can be misleading in evaluating the causal total effect of a (treatment) variable X on an (outcome) variable Y. In chapter 2, we described random experiments of various research designs in which a causal total effect is of interest. In chapter 3, we reviewed the most important concepts of probability theory, which are treated more extensively in Steyer and Nagel (2017). The examples in chapters 1 and 3 show that the traditional probabilistic concepts alone cannot be used offhandedly to define the causal effects in which we are interested when we evaluate a treatment, an intervention, or an exposition.

In the present chapter, we introduce the concept of a *potential confounder*, which is crucial for the theory of probabilistic causality. In fact, conditional expectations such as $E(Y|X)$ and $E(Y|X,Z)$ that describe a causal dependence can be distinguished from conditional expectations that have no causal interpretation by their relationship to the potential confounders. A *global potential confounder* is a special potential confounder that is used in the definition of a *true outcome variable*. True outcome variables will be used in chapter 5 to define various kinds of causal total effects and in chapter 6 to define *unbiasedness* of the conditional expectations $E(Y|X)$ and $E(Y|X,Z)$, where Z denotes a covariate and potential confounder.

4.1 Potential Confounder and Global Potential Confounder

In chapter 2 we already discussed the role of potential confounders in single-unit trials. There, we already mentioned that we consider all random variables as potential confounders of a focussed cause X that are prior or at least simultaneous to X such that they can induce biased prima facie total effects of X.

We start with the formal framework in which we can define the concept of a potential confounder and a global potential confounder. This necessitates that we consider a random experiment that involves a time set T with elements 1, 2, and 3, allowing to distinguish between past, present, and future events, respectively, from the perspective of the treatment (intervention or exposition). Note that this does not preclude that the underlying process is more fine-grained in terms of the time points considered. For the present purpose, however, it suffices to distinguish between three time 'points' or 'phases' of the process considered. Presenting this framework we use the concept of a product σ-algebra, which is specified, for example, in SN-Definition 1.31.

Definition 4.1 (Causality Space)
Let $(\Omega, \mathscr{A}, P)$ be a probability space and let $(\Omega_t, \mathscr{A}_t)$, $t \in T = \{1,2,3\}$, be measurable spaces such that

$$\Omega = \Omega_1 \times \Omega_2 \times \Omega_3 \quad \text{and} \quad \mathscr{A} = \mathscr{A}_1 \otimes \mathscr{A}_2 \otimes \mathscr{A}_3, \tag{4.1}$$

where $\mathscr{A}_1 \otimes \mathscr{A}_2 \otimes \mathscr{A}_3$ is the product σ-algebra. For each $t \in T$, assume that $\mathscr{A}_t$ is such that $\{\omega_t\} \in \mathscr{A}_t$ for all $\omega_t \in \Omega_t$. Furthermore, for each $t \in T$, let $h_t : \Omega \to \Omega_t$ denote the projection defined by

$$h_t(\omega) = h_t(\omega_1, \omega_2, \omega_3) = \omega_t, \quad \forall\, \omega \in \Omega, \tag{4.2}$$

and assume that each h_t, $t \in T$, is $\mathscr{A}$-measurable. Let $X : \Omega \to \Omega'_X$ and $Y : \Omega \to \Omega'_Y$ be random variables on $(\Omega, \mathscr{A}, P)$ with value spaces $(\Omega'_X, \mathscr{A}'_X)$ and $(\Omega'_Y, \mathscr{A}'_Y)$, respectively, such that $\{x\} \in \mathscr{A}'_X$ for all $x \in \Omega'_X$. Finally, define $\mathscr{F}_1 = \sigma(h_1)$, $\mathscr{F}_2 = \sigma(h_1, h_2)$, and $\mathscr{F}_3 = \sigma(h_1, h_2, h_3)$, and assume that $\sigma(Y) \neq \{\Omega, \emptyset\} \neq \sigma(X)$, $\sigma(X) \subset \sigma(h_2)$, and $\sigma(Y) \subset \sigma(h_3)$. Then $\big((\Omega, \mathscr{A}, P), (\mathscr{F}_t, t \in T), X, Y\big)$ is called a (total effects) causality space.

Remark 4.2 (Intuitive Background of a Causality Space) In a causality space we refer to a probability space $(\Omega, \mathscr{A}, P)$ with a set of possible outcomes Ω that can be written as a Cartesian product of three sets Ω_1, Ω_2, and Ω_3. Of course, each of these sets can itself be a Cartesian product of other sets. Requiring $\mathscr{A} = \mathscr{A}_1 \otimes \mathscr{A}_2 \otimes \mathscr{A}_3$ and $\{\omega_t\} \in \mathscr{A}_t$ for all $\omega_t \in \Omega_t$ and each $t \in T$, we secure that $\mathscr{A}$ is fine-grained enough to contain all relevant events in the random experiment, which, in empirical applications, is represented by $(\Omega, \mathscr{A}, P)$. The projections h_1, h_2, and h_3 and the requirements for the σ-algebras $\sigma(X)$ and $\sigma(Y)$ are constructed such that the σ-algebras generated by these projections can, from the perspective of the treatment variable X, be interpreted as representing the sets of past, present, and future events, respectively, in the causal process considered. ◁

Remark 4.3 (Filtration) Using the projections h_1, h_2, and h_3 we specified a filtration $(\mathscr{F}_t, t \in T)$ by $\mathscr{F}_1 = \sigma(h_1)$, $\mathscr{F}_2 = \sigma(h_1, h_2)$, and $\mathscr{F}_3 = \sigma(h_1, h_2, h_3)$. Such a filtration satisfies $\mathscr{F}_s \subset \mathscr{F}_t$ for all $s, t \in T$ with $s \leq t$. With respect to such a filtration we can also define time order between events, sets of events (including σ-algebras), and random variables (see, e. g., SN-Rem. 4.18), so that phrases such as *X is prior to Y* have a well-defined meaning. In the framework of a causality space this phrase means $\sigma(X) \subset \mathscr{F}_2$ and $\mathscr{F}_2 \not\supset \sigma(Y) \subset \mathscr{F}_3$. ◁

Now we define a *potential confounder of X* and a *global potential confounder of X* as follows:

Definition 4.4 (Potential Confounder and Global Potential Confounder of X)
Let $\big((\Omega, \mathscr{A}, P), (\mathscr{F}_t, t \in T), X, Y\big)$ be causality space.

(i) A random variable $W : \Omega \to \Omega'_W$ with value space $(\Omega'_W, \mathscr{A}'_W)$ on $(\Omega, \mathscr{A}, P)$ is called a potential confounder of X if there is a mapping $g : \Omega_1 \to \Omega'_W$ such that $W = g \circ h_1$ and $g^{-1}(\mathscr{A}'_W) \subset \mathscr{A}_1$.

(ii) A potential confounder W of X is called global *if* $\sigma(W) = \sigma(h_1)$. *In this case, we often use the notation* C_X *instead of* W *and denote the value space of* C_X *by* $(\Omega'_{C_X}, \mathscr{A}'_{C_X})$.
(iii) We call $\mathscr{F}_1 = \sigma(h_1)$ *the* σ*-algebra of potential confounders of* X.

If there is ambiguity that we are considering total effects, and not direct or indirect effects, then we can add the term 'with respect to total effects' to the terms introduced in points (i) and (ii) of Definition 4.4.

Remark 4.5 (A Potential Confounder is h_1-Measurable) According to SN-Theorem 2.49, a potential confounder is measurable with respect to the projection h_1 introduced in Definition 4.1. From a substantive point of view, this secures that a potential confounder of X represents a random variable that is prior or simultaneous to X in the causal process considered, provided, of course, that, in an empirical application, Ω_1 and $\mathscr{A}_1$ are constructed such that the σ-algebra generated by h_1 represents all events that are prior or simultaneous to X. ◁

Remark 4.6 (An Example of a Global Potential Confounder) The projection h_1 itself is a global potential confounder of X. In this case, the identity mapping $id\colon \Omega_1 \to \Omega_1$ takes the role of the mapping g occuring in Definition 4.4 (i). ◁

Remark 4.7 (Uniqueness) Once the sets Ω_t, $t = 1,2,3$, are specified, then the projections $h_t\colon (\Omega, \mathscr{A}) \to (\Omega_t, \mathscr{A}_t)$ are uniquely defined by Equation (4.2). This implies that the σ-algebras generated by each h_t are uniquely defined as well. There can be several global potential confounders of X (see U and $1_{U=Joe}$ in Example 4.10). However, the σ-algebras they generate are identical to $\sigma(h_1)$, the σ-algebra of potential confounders of X. ◁

Remark 4.8 (Coarsening X) In Definition 4.4, requiring $\sigma(X) \subset \sigma(h_2)$ instead of $\sigma(X) = \sigma(h_2)$ allows us, for instance, to merge two or more (original) treatment conditions into a single one and compare the merged treatment to another one. For example, if originally there are two treatment conditions and a control condition, then we can define a treatment variable X with values 0 and 1 that compares treatment (no matter which of the two) ($X = 1$) to control ($X = 0$). ◁

Remark 4.9 (Covariate of X) If W is a potential confounder of X, then we also call it a *covariate of* X, in particular if we condition on W and X in a conditional expectation $E(Y|X,W)$ or in a conditional distribution $P_{Y|X,W}$. ◁

Example 4.10 (Joe and Ann With Randomized Assignment) We illustrate the concepts introduced in Definition 4.4 by the example presented in Table 3.1. This table refers to the following random experiment: *First*, we sample a unit u from the set $\Omega_U := \{Joe, Ann\}$. *Second*, each unit receives (*yes*) or does not receive a treatment (*no*), and *third*, it is observed whether (+) or not (−) a success criterion is reached at some appropriate time after treatment. Defining $\Omega_X := \{yes, no\}$ and $\Omega_Y := \{+,-\}$, the set of possible outcomes ω of this random experiment is

$$\begin{aligned} \Omega &= \Omega_1 \times \Omega_2 \times \Omega_3 := \Omega_U \times \Omega_X \times \Omega_Y \\ &= \big\{ (Joe, no, -),\ (Joe, no, +), \ldots,\ (Ann, yes, +) \big\}. \end{aligned} \tag{4.3}$$

In this example, the set Ω has eight elements, the triples $\omega_1 = (Joe, no, -)$, $\omega_2 = (Joe, no, +)$, …, $\omega_8 = (Ann, yes, +)$ (see the first column of Table 3.1 for a complete list of these elements). Furthermore, we define $\mathscr{A} := \mathscr{P}(\Omega)$ to be the power set of Ω. Finally, because each nonempty element $A \in \mathscr{A}$ is a union of singletons $\{\omega_i\}$, $\omega_i \in \Omega$, and because a measure is additive, the probability measure $P: \mathscr{A} \to [0,1]$ is uniquely defined by the second column of Table 3.1 [see SN-Box 4.1 (x)]. Hence, the probability space $(\Omega, \mathscr{A}, P)$ is completely specified.

Aside from the projections (see below), we consider three random variables: the *observational-unit variable* or *person variable* $U: \Omega \to \Omega_U$ with value space $\big(\Omega_U, \mathscr{P}(\Omega_U)\big)$, the *treatment variable* $X: \Omega \to \mathbb{R}$ with value space $(\mathbb{R}, \mathscr{B})$, and the *outcome* or *response variable* $Y: \Omega \to \mathbb{R}$, also with value space $(\mathbb{R}, \mathscr{B})$. Table 3.1 shows how each of these random variables assigns one of its values to each of the eight elements $\omega_i \in \Omega$.

In this example, the projection mapping h_1 is defined by

$$h_1(\omega_i) \;=\; h_1\big((u, \omega_X, \omega_Y)\big) \;=\; u, \quad \forall\, \omega_i \in \Omega.$$

Applying this assignment rule to all eight elements ω_i displayed in the first column of Table 3.1 shows that $h_1 = U$. Hence, according to Definition 4.4 (ii), U is a global potential confounder of X.

The σ-algebra $\mathscr{F}_1$ has only four elements, the event that *Joe is drawn*,

$$\{U = Joe\} \;:=\; U^{-1}(\{Joe\}) \;=\; \big\{(Joe, no, -),\, (Joe, no, +),\, (Joe, yes, -),\, (Joe, yes, +)\big\},$$

the event that *Ann is drawn*,

$$\{U = Ann\} \;:=\; U^{-1}(\{Ann\}) \;=\; \big\{(Ann, no, -),\, (Ann, no, +),\, (Ann, yes, -),\, (Ann, yes, +)\big\},$$

the *sure event* Ω, and the *impossible event* $\emptyset$. Note that U is a projection that maps each $\omega \in \Omega$ on the first factor set Ω_U of Ω. [It is identical to the projection mapping h_1 introduced in Def. 4.4, that is, $h_1 = U$ and $\sigma(h_1) = \sigma(U)$]. Furthermore, the projection mappings $h_2: \Omega \to \Omega_X$ and $h_3: \Omega \to \Omega_Y$ are defined by

$$h_2(\omega_i) \;=\; h_2\big((u, \omega_X, \omega_Y)\big) \;=\; \omega_X, \quad \forall\, \omega_i \in \Omega,$$

and

$$h_3(\omega_i) \;=\; h_3\big((u, \omega_X, \omega_Y)\big) \;=\; \omega_Y, \quad \forall\, \omega_i \in \Omega,$$

respectively.

Note that X and h_2 are not identical because they take on different values. Whereas the values of h_2 are *yes* or *no*, the values of X are 1 or 0, respectively. Nevertheless, the σ-algebras generated by X and generated by h_2 *are* identical (see Exercise 4-1). That is, $\sigma(X) = \sigma(h_2)$, and this implies $\sigma(X) \subset \sigma(h_2)$ (see Def. 4.1).

According to Definition 4.4 (i), each random variable $W: \Omega \to \Omega'_W$ on $(\Omega, \mathscr{A}, P)$ for which there is a mapping $g: \Omega_1 \to \Omega'_W$ such that $W = g \circ h_1$ and $g^{-1}(\mathscr{A}'_W) \subset \mathscr{A}_1$ is a potential confounder. The indicator $1_{U=Joe}$ defined by

$$1_{U=Joe}(\omega_i) \;=\; \begin{cases} 1, & \text{if } \omega_i \in \{U = Joe\} \\ 0, & \text{otherwise} \end{cases}$$

is an example in case. In this specific example, in which there are only two observational units, $1_{U=Joe}$ is also a global potential confounder of X (see Exercise 4-2).

The structure of this concrete random experiment is essentially the same for every simple experiment of the type described in section 2.1. Only the number of values of U, X, and Y might change if we consider more than two observational units, more than two treatment conditions, or more than two values of the outcome variable. ◁

Example 4.11 (Nonorthogonal Two-Factorial Experiment) In the example presented in Table 1.5, the person variable U is a global potential confounder, because U is only a more convenient notation for the projection $h_1 : \Omega \to \Omega_U$ (see Example 3.2).

The random variable Z occurring in Table 1.5 is a potential confounder, but not a global one. This random variable Z is also called a (qualitative) covariate if we consider the conditional expectation values $E(Y \mid X{=}x, Z{=}z)$ (see Remark 4.9 and Table 1.7). The definition of Z in Equation (3.13) shows that there is a mapping $g\colon \Omega_U \to \Omega'_Z = \{low, med, hi\}$ such that $Z = g \circ U$. This mapping is defined by

$$\begin{aligned} g(Tim) &= g(Tom) = low \\ g(Joe) = g(Jim) &= g(Ann) = g(Eva) = med \\ g(Sue) &= g(Mia) = hi. \end{aligned} \tag{4.4}$$

Obviously, $Z(\omega) = g\big(U(\omega)\big)$ [see again Eq. (3.13)].

The inverse images $g^{-1}(B)$, $B \in \mathscr{P}(\Omega'_Z)$, are

$$\begin{aligned} g^{-1}(\{low\}) &= \{Tim, Tom\} \\ g^{-1}(\{med\}) &= \{Joe, Jim, Ann, Eva\} \\ g^{-1}(\{hi\}) &= \{Sue, Mia\} \\ g^{-1}(\{low, med\}) &= \{Tim, Tom, Joe, Jim, Ann, Eva\} \\ g^{-1}(\{low, hi\}) &= \{Tim, Tom, Sue, Mia\} \\ g^{-1}(\{med, hi\}) &= \{Joe, Jim, Ann, Eva, Sue, Mia\} \\ g^{-1}(\Omega'_Z) &= \Omega_U \\ g^{-1}(\emptyset) &= \emptyset. \end{aligned} \tag{4.5}$$

These eight inverse images are the elements of $g^{-1}\big(\mathscr{P}(\Omega'_Z)\big)$ and this σ-algebra is a subset of $\mathscr{P}(\Omega_U)$. Hence, Z satisfies all requirements of a potential confounder [see Def. 4.4 (i)]. ◁

Example 4.12 (Experiments With Fallible Covariates) Which is the structure of the single-unit trial of experiments and quasi-experiments if we *do observe* at least one fallible covariate (see section 2.2)? This kind of random experiments consists of

(a) drawing a unit from Ω_U,
(b) observing an element ω_Z of Ω_Z (based on which the fallible covariate $Z\colon \Omega \to \Omega'_Z$ assigns a value $z \in \Omega'_Z$ to $\omega_i \in \Omega$),
(c) assigning the unit or observing its selection to one of the experimental conditions (represented by the value x of the treatment variable X), and
(d) recording the numerical value y of the outcome variable Y.

Hence,

$$\Omega \;=\; \Omega_1 \times \Omega_2 \times \Omega_3 \;=\; (\Omega_U \times \Omega_Z) \times \Omega_X \times \Omega_Y, \qquad \text{where } \Omega_1 = \Omega_U \times \Omega_Z. \tag{4.6}$$

Compared to Equation (4.3), now the set Ω involves the additional set Ω_Z, which together with the set Ω_U of observational units, defines $\Omega_1 = \Omega_U \times \Omega_Z$. Note that Z may also be multivariate, consisting of several univariate covariates. It consists of *fallible measures* of attributes of the units, for example, self-rated *motivation for therapy*, *personality*, or *ability test score variables*. Given a particular unit u, there is at least one value of such a fallible variable such that $0 < P(Z{=}z) < 1$. Because Z is fallible, there is no mapping $f\colon \Omega_U \to \Omega'_Z$ such that $Z = f(U)$. This includes the fallible measures of a latent potential confounder, say ξ, which itself, by definition, is a mapping of U (see, e. g., Steyer et al., 2015). Hence, the latent variable ξ itself is *not* fallible, it is measurable with respect to U. ◁

Example 4.13 (Joe With two Independent Treatments) Table 4.1 presents another random experiment in which a given unit, say Joe, may simultaneously receive either:

(a) no treatment at all,
(b) not a first treatment (e. g., no group therapy) but a second one (e. g., individual therapy),
(c) a first treatment (group therapy) but no second one (no individual therapy),
(d) both treatments (group therapy and individual therapy).

Furthermore, it is registered whether (+) or not (−) a success criterion is reached. The set Ω of all possible outcomes of the random experiment consists of the eight possible outcomes listed in the first column of the table.

In this example, we consider only one single person, which implies that the observational-unit variable is a constant and can be ignored. Therefore, it does not appear in Table 4.1. Furthermore, we can focus X as the treatment variable and Z as a potential confounder, but also vice versa. This choice has implications on how we structure the set Ω of possible outcomes.

Choosing X as a cause and Z as a potential confounder, Ω_Z, Ω_X, and Ω_Y take the roles of Ω_1, Ω_2, and Ω_3, respectively [see Eq. (4.1)]. That is,

$$\Omega \;=\; \Omega_1 \times \Omega_2 \times \Omega_3 \;=\; \Omega_Z \times \Omega_X \times \Omega_Y,$$

with $\Omega_Z = \Omega_X = \{yes, no\}$, and $\Omega_Y = \{+, -\}$.

We choose $\mathscr{A} = \mathscr{P}(\Omega)$ to be the power set of Ω. The probabilities of all $2^8 = 256$ events (i. e., of the elements of $\mathscr{A}$) are determined by the probabilities of the elementary events displayed in the first column of the table. The table also shows how the random variables X, Y, and Z are specified, that is, how their values are assigned to each $\omega_i \in \Omega$.

In this example, the random variable $Z\colon \Omega \to \Omega'_Z$ is identical to the projection h_1, because $\Omega'_Z = \Omega_1 = \Omega_Z = \{yes, no\}$ and

$$\mathscr{A}'_Z := \mathscr{P}(\Omega'_Z) \;=\; \big\{\{no\}, \{yes\}, \Omega'_Z, \emptyset\big\}$$

contains the singletons $\{no\}$ and $\{yes\}$ as elements. Furthermore, the treatment variable $X\colon \Omega \to \Omega'_X$ is identical to the projection h_2, because $\Omega'_X = \Omega_2 = \Omega_X = \{yes, no\}$ and

$$\mathscr{A}'_X := \mathscr{P}(\Omega'_X) \;=\; \big\{\{no\}, \{yes\}, \Omega'_X, \emptyset\big\}$$

Table 4.1. Joe with two independent treatments

Outcomes ω_i					Random variables			Conditional expectations				
Group therapy	Individual therapy	Success	$P(\{\omega_i\})$	$P^{X=0}(\{\omega_i\})$	$P^{X=1}(\{\omega_i\})$	Treatment variable Z	Treatment variable X	Outcome variable Y	$E(Y\|X,Z)$	$E(Y\|X)$	$E(X\|Z)$	$\tau_0 = E^{X=0}(Y\|Z)$	$\tau_1 = E^{X=1}(Y\|Z)$
$\omega_1 = (no,$	$no,$	$-)$	.09	.15	0	*no*	*no*	0	.7	.45	.4	.7	.8
$\omega_2 = (no,$	$no,$	$+)$	.21	.35	0	*no*	*no*	1	.7	.45	.4	.7	.8
$\omega_3 = (no,$	$yes,$	$-)$	.04	0	.1	*no*	*yes*	0	.8	.6	.4	.7	.8
$\omega_4 = (no,$	$yes,$	$+)$	.16	0	.4	*no*	*yes*	1	.8	.6	.4	.7	.8
$\omega_5 = (yes,$	$no,$	$-)$	.24	.4	0	*yes*	*no*	0	.2	.45	.4	.2	.4
$\omega_6 = (yes,$	$no,$	$+)$	.06	.1	0	*yes*	*no*	1	.2	.45	.4	.2	.4
$\omega_7 = (yes,$	$yes,$	$-)$	.12	0	.3	*yes*	*yes*	0	.4	6	.4	.2	.4
$\omega_8 = (yes,$	$yes,$	$+)$	.08	0	.2	*yes*	*yes*	1	.4	6	.4	.2	.4

contains the singletons $\{no\}$ and $\{yes\}$ as elements. The two possible values of this random variable, *yes* and *no* represent receiving and not receiving *individual therapy*, respectively.

Note that we could also have chosen co-domains for the true treatment variables that differ from $\{yes, no\}$, such as $\{0,1\}$ or $\mathbb{R}$. Although these co-domains would lead to other random variables that would not be identical any more to Z or X, respectively, the σ-algebras generated by these new treatment variables would still be identical to $\sigma(h_1)$ and $\sigma(h_2)$, respectively. $\triangleleft$

The σ-algebra $\mathscr{F}_1 = \sigma(h_1)$ has four elements, the event

$$\{(no, no, -), (no, no, +), (no, yes, -), (no, yes, +)\}$$

that no group therapy is obtained, the event

$$\{(yes, no, -), (yes, no, +), (yes, yes, -), (yes, yes, +)\}$$

that group therapy is obtained, the sure event Ω and the impossible event $\emptyset$. The σ-algebra $\mathscr{F}_1$ is also the σ-algebra of potential confounders of X.

The σ-algebra $\mathscr{F}_2$ has $2^4 = 16$ elements, the four events

$$\begin{array}{ll}
\{(no, no, -), (no, no, +)\} & \textit{'neither individual nor group therapy'},\\
\{(no, yes, -), (no, yes, +)\} & \textit{'no group but individual therapy'},\\
\{(yes, no, -), (yes, no, +)\} & \textit{'group but no individual therapy'},\\
\{(yes, yes, -), (yes, yes, +)\} & \textit{'individual and group therapy'},
\end{array}$$

all (pairwise, triple-wise, quadruple-wise) unions of these four events, and the impossible event $\emptyset$. The first set $\{(no, no, -), (no, no, +)\}$, for example, is the event that Joe neither receives treatment 1 nor treatment 2, whereas the second set $\{(no, yes, -), (no, yes, +)\}$ is

the event that Joe does not receive treatment 1 (group therapy) but receives treatment 2 (individual therapy).

The σ-algebra $\mathscr{F}_3$ has $2^8 = 256$ elements. It consists of the eight elementary events

$$\{(no,\ no,-)\},\ \{(no,\ no,+)\},\ldots,\ \{(yes,\ yes,+)\}$$

(see the first column of Table 4.1), all unions of these events, which includes Ω, and the impossible event Ø.

4.2 True Outcome Variable

In this section, we introduce the concept of a *(total effects) true outcome variable of Y given the value x of X*. This is a fundamental concept of the theory of causal effects. In its definition, we consider a C_X-conditional expectation of Y with respect to the conditional probability measure $P^{X=x}$ (see Def. 3.79), where C_X denotes a global potential confounder of X (see Def. 4.4). This is tantamount to conditioning on the event $\{X=x\}$ and on *all* potential confounders of X. To emphasize, defining a total effects true outcome variable, we do *not* condition on intermediate variables, and this is why we refer to total effects if the context is ambiguous.

Remark 4.14 (Intuitive Background) The intuitive background is as follows. Conditioning on a global potential confounder, we share John Stuart Mill's idea already described in the preface. However, we make a slight modification. Instead of comparing values of Y, we compare certain *conditional expectation values* of Y between treatment conditions.

As we have seen in Example 4.10, the person variable U can take the role of a global potential confounder of X. Now, suppose an observational unit u may receive treatment (intervention, exposition) $(X=1)$ or control (no treatment, an alternative treatment or exposition) $(X=0)$. If there is a difference between the $(X=1, U=u)$-conditional expectation value and the $(X=0, U=u)$-conditional expectation value, then this difference is due to (i. e., caused by) the treatment variable X. In this example, conditioning also on a person u means that 'everything else is invariant', for example, the severity of symptoms, the motivation for treatment, educational status, etc. Hence, this is a probabilistic version of the *ceteris paribus clause*.[1] Note that the treatment effect can be different for different values of the global potential confounder. ◁

Remark 4.15 (Two Concepts) The intuitive idea of a true outcome variable outlined in Remark 4.14 can mathematically be specified by *conditional expectations $E^{X=x}(Y|C_X)$ with respect to a conditional probability measure $P^{X=x}$* (see section 3.3.6). If $P(X=x)>0$, then this concept can be used to describe how a numerical random variable Y depends on a random variable C_X given the event $\{X=x\}$. The difference $E^{X=x}(Y|C_X) - E^{X=x'}(Y|C_X)$ then defines the true effect function (see Def. 5.2). Alternatively, we might use the *partial conditional expectations* $E(Y|X=x,C_X)$ and $E(Y|X=x',C_X)$ (see Def. 3.73) comparing treatment x to treatment x'. According to Theorem 3.91, if $P(X=x)>0$, then $E(Y|X=x,C_X)$ is also a version of the conditional expectation $E^{X=x}(Y|C_X)$ with respect to the conditional probability measure $P^{X=x}$. ◁

[1] This idea is already found in Splawa-Neyman (1923/1990). Later, it has been oversimplified by Rubin introducing the potential outcome variables (see, e. g., Rubin, 1974, 2005). The true outcome variables play a similar role as the potential outcome variables in Rubin's approach.

Definition 4.16 (True Outcome Variable)
Let the assumptions 4.1 hold, where Y is real-valued with $E(Y^2) < \infty$, let $x \in \Omega'_X$ be a value of X with $P(X{=}x) > 0$, and let C_X be a global potential confounder of X. Then

$$\tau_x := E^{X=x}(Y|C_X) \tag{4.7}$$

is called a true outcome variable of Y given the value x of X or given the event $\{X{=}x\}$.

Remark 4.17 (A Caveat on Notation) The shortcut τ_x is meaningful only if the references to a specified outcome variable Y, a specified cause X, and a global potential confounder C_X of X are unambiguous. ◁

Remark 4.18 (Value of a True Outcome Variable) Assume that x is a value of X and c a value of C_X such that $P(X{=}x, C_X{=}c) > 0$. Then a value of τ_x is called the *true (or expected) outcome of Y given the value x of X* or *given the event* $\{X{=}x\}$. According to Equations (3.69) and (4.7),

$$\forall\, \omega \in \Omega: \quad \tau_x(\omega) \;=\; E^{X=x}(Y|C_X)(\omega) \;=\; E^{X=x}(Y|C_X{=}c), \quad \text{if } \omega \in \{C_X{=}c\}. \tag{4.8}$$

Furthermore, if $P(X{=}x, C_X{=}c) > 0$, then

$$E^{X=x}(Y|C_X{=}c) \;=\; E(Y|X{=}x, C_X{=}c), \tag{4.9}$$

[see Eq. (3.70)]. That is, a value of a true outcome variable τ_x is identical to a conditional expectation value of Y given the value x of X and the value c of a global potential confounder C_X. As already mentioned in Remark 3.68, such a conditional expectation value is uniquely defined only if $P(X{=}x, C_X{=}c) > 0$. ◁

Remark 4.19 (τ_x is C_X-Measurable) Note that τ_x is a random variable on $(\Omega, \mathscr{A}, P)$ that is measurable with respect to C_X, that is, $\sigma(\tau_x) \subset \sigma(C_X)$. This follows from the fact that there is a mapping $g_x\colon \Omega'_{C_X} \to \overline{\mathbb{R}}$ such that $\tau_x = g_x \circ C_X$ and $g_x^{-1}(\overline{\mathscr{B}}) \subset \mathscr{A}'_{C_X}$ [see SN-Th. 2.49 and Def. 4.4 (ii)]. The existence of such a mapping g_x follows from the fact that $\tau_x = E^{X=x}(Y|C_X)$ is a conditional expectation (with respect to the measure $P^{X=x}$) and Corollary 3.63. Hence, we may rewrite Equation (4.8) as follows:

$$\begin{aligned} \forall\, \omega_i \in \Omega: \quad \tau_x(\omega_i) \;&=\; E^{X=x}(Y|C_X)(\omega_i) \;=\; g_x \circ C_X(\omega_i) \\ &=\; g_x\big(C_X(\omega_i)\big) \;=\; g_x(c) \;=\; E^{X=x}(Y|C_X{=}c), \quad \text{if } \omega_i \in \{C_X{=}c\}. \end{aligned} \tag{4.10}$$

According to this equation, in order to assign a value of τ_x to an outcome $\omega_i \in \Omega$ of the random experiment, first we may assign to ω_i a value c of the global potential confounder, and then assign to c via g_x the corresponding conditional expectation value $E^{X=x}(Y|C_X{=}c)$ (see Def. 3.87, Rem. 3.88, and Example 4.21). ◁

Remark 4.20 (True Outcome Variable of Y Given Treatment x) In Definition 4.16 we do *not* presume that X is a treatment variable in an experiment or quasi-experiment. If, however, X is a treatment variable, then we call τ_x a true outcome variable of *Y given treatment x*. If there is ambiguity, we use the term *total effects* true outcome variable in order to distinguish it from a *direct effects* true outcome variable (which is not treated in this volume). ◁

Example 4.21 (Joe and Ann With Randomized Assignment) The most important parameters of this random experiment are presented in Table 3.1. In Example 4.10 we already specified the set Ω of possible outcomes and the projections h_1, h_2, and h_3 for this random experiment. There, we also asserted that U is a global potential confounder of the treatment variable X, playing the role of C_X in the general theory. Therefore, $\tau_x = E^{X=x}(Y|C_X) = E^{X=x}(Y|U)$ for both treatments x=0 and x=1.

Now we specify the true outcome variables $\tau_x = E^{X=x}(Y|C_X) = E^{X=x}(Y|U)$ for the two treatment conditions $x = 0, 1$. The *true outcome variable* $\tau_0 = E^{X=0}(Y|U)$ of Y given control is specified by

$$\tau_0 \;=\; g_0(U), \tag{4.11}$$

where $U\colon \Omega \to \Omega_U$ with

$$U(\omega_i) \;=\; U\big((u, \omega_X, \omega_Y)\big) \;=\; u, \quad \text{for all } \omega_i \in \Omega, \tag{4.12}$$

and $g_0\colon \Omega_U \to \mathbb{R}$ with

$$g_0(u) \;=\; E(Y|X=0, U=u), \quad \text{for all } u \in \Omega_U \tag{4.13}$$

[see Eqs. (4.9) and (4.10)]. Hence, in order to assign a value of τ_0 to an outcome $\omega_i \in \Omega$ of the random experiment, first we have to assign to ω_i a value u (*Joe* or *Ann*) of the observational-unit variable U, and then assign to u via g_0 the corresponding conditional expectation value $E(Y|X=0, U=u)$ (see Def. 3.87 and Rem. 3.88).

If, for instance, $\omega_3 = (\mathit{Joe}, \mathit{yes}, -)$ (see the third row in Table 3.1), then $U(\omega_3) = \mathit{Joe}$, and the value of τ_0 is

$$\tau_0(\omega_3) \;=\; g_0\big(U(\omega_3)\big) \;=\; g_0(\mathit{Joe}) \;=\; E(Y|X=0, U=\mathit{Joe}) \;=\; P(Y=1\,|\,X=0, U=\mathit{Joe}) \;=\; .7.$$

This is true *even though* $X(\omega_3) = 1$ and the value of the conditional expectation $E(Y|X,U)$ is

$$E(Y|X,U)(\omega_3) \;=\; E(Y|X=1, U=\mathit{Joe}) \;=\; P(Y=1\,|\,X=1, U=\mathit{Joe}) \;=\; .8.$$

Hence, the true outcome variable τ_0 of treatment 0 takes on a well-defined value for ω_3 *even though* the unit drawn receives treatment 1. This illustrates the distinction between the random variables $\tau_0 = E^{X=0}(Y|U)$ and $E(Y|X,U)$. While τ_0 is solely a function of U [see Eq. (4.11)], the conditional expectation $E(Y|X,U)$ is a function of X and U. That is, there is a function $g\colon \Omega'_X \times \Omega_U \to \mathbb{R}$ such that $E(Y|X,U) = g(X,U)$ can be written as the composition of (X,U) and g.

Because, in this example, the outcome variable Y is binary with values 0 and 1, the conditional expectation value $E(Y|X=0, U=u)$ is also the conditional probability $P(Y=1|\,X=0, U=u)$ of success, and because, in this example, U has only two values, *Joe* and *Ann*, the true outcome variable τ_0 also has only two different values, the two conditional probabilities $P(Y=1\,|\,X=0, U=\mathit{Joe}) = .7$ and $P(Y=1\,|\,X=0, U=\mathit{Ann}) = .2$ (see Table 3.1).

Similarly, the *true outcome variable* $\tau_1 = E^{X=1}(Y|U)$ of treatment condition 1 is specified by

$$\tau_1 \;=\; g_1(U), \tag{4.14}$$

where $g_1\colon \Omega_U \to \mathbb{R}$ is defined by

$$g_1(u) \;=\; E(Y|X=1, U=u), \quad \text{for all } u \in \Omega_U. \tag{4.15}$$

Hence, if $\omega_i \in \{U=Joe\}$, then the value of τ_1 is

$$\tau_1(\omega_i) \;=\; g_1\big(U(\omega_i)\big) \;=\; g_1(Joe) \;=\; E(Y|X=1,U=Joe) \;=\; P(Y=1|X=1,U=Joe) \;=\; .8,$$

and if $\omega_i \in \{U=Ann\}$, then the value of τ_1 is

$$\tau_1(\omega_i) \;=\; g_1\big(U(\omega_i)\big) \;=\; g_1(Ann) \;=\; E(Y|X=1,U=Ann) \;=\; P(Y=1|X=1,U=Ann) \;=\; .4.$$

Table 3.1 shows which values τ_0 and τ_1 assign to each of the eight possible outcomes ω_i of the random experiment. ◁

Remark 4.22 (True Outcome vs. Potential Outcomes) Rubin (1974, 2005) assumes that given an observational unit u and a treatment condition x, the values of his potential outcome variables Y_0 and Y_1 are fixed numbers. In the example presented in Table 3.1, this would mean to replace the two true outcome variables τ_0 and τ_1 by the two potential outcome variables Y_0 and Y_1 that can take on only the values 0 and 1. Substantively speaking, this would mean that, given a concrete treatment and a concrete observational unit, the outcome is fixed to 0 or 1. For example, if the outcome is being alive ($Y=1$) or not ($Y=0$) at the age of 80 and the treatment is receiving ($X=1$) or no receiving ($X=0$) an anti-smoking therapy before the age of 40, then this deterministic idea is not in line with our knowledge of causes for being dead at the age of 80.

In contrast, the two true outcome variables τ_0 and τ_1 can take on any real number as values. In the example of Table 3.1, they can take on any value between 0 and 1, inclusively. In the smoking example, their values would be the person-specific probabilities of being alive at the age of 80, given treatment or no treatment. Therefore, the true outcomes variables can be considered to be a generalization of the potential outcome variables. Most important, in contrast to potential outcome variables, true outcomes variables are in line with the idea that events that occur in between treatment and outcomes may also affect the outcome variable Y. ◁

Example 4.23 (No Treatment for Joe) In Table 3.2 we presented an example, in which U is a global potential confounder and the true outcome variable $\tau_1 = E^{X=1}(Y|U)$ is neither uniquely defined nor P-unique. In Example 3.62 we showed that there are several versions of the (X,U)-conditional expectation of Y that are P-equivalent. In Example 3.85 we showed that $\tau_1 = E^{X=1}(Y|U)$ is $P^{X=1}$-unique but not P-unique. In contrast, in the same example, $\tau_0 = E^{X=0}(Y|U)$ is not only $P^{X=0}$-unique but also P-unique and even uniquely defined. ◁

Example 4.24 (Nonorthogonal Two-Factorial Experiment) In Example 4.11 we already showed that U is a global potential confounder of X and that Z is a potential confounder of X. In the last three columns of Table 1.5 we presented the true outcome variables $\tau_0 = E^{X=0}(Y|U)$, $\tau_1 = E^{X=1}(Y|U)$, and $\tau_2 = E^{X=2}(Y|U)$. In this example, all three true outcome variables are uniquely defined, because $P(X=1|U)>0$ (see Th. 3.84). ◁

4.3 Summary and Conclusions

In this chapter we set the stage for defining causal effects and causal probabilistic dependencies. We specified the mathematical structure, a *causality space*, and formulated the

assumptions under which we can define causal effects and meaningfully raise the question if conditional expectation values such as $E(Y|X=x)$ or $E(Y|X=x,Z=z)$ can be used to describe causal effects or if conditional distributions such as $P_{Y|X=x}$ or $P_{Y|X=x,Z=z}$ can be used to define causal dependencies.

We started with the fundamental concepts of a *potential confounder* and a *global potential confounder* of a focused cause (variable) X. Both concepts are essential for the theory of probabilistic causality because conditional expectations and distributions that have a causal interpretation differ from those having no such interpretation by their relationship to all potential confounders (see the chapters to come). The definition of these concepts presumes that the probability space considered has a certain mathematical structure that has been specified in Definition 4.1. From a substantive point of view, the most important idea is that we consider a random experiment in which a focussed cause (variable), say X, has a past, a present, and a future. Because we only deal with causal *total* effects of X, potential confounders pertain to the past of X, whereas the outcome variable Y pertains to the future of X. These intuitive ideas have been formalized in Def. 4.4) utilizing the concepts of measurability and measurable mappings that are well-known in mathematical measure and probability theory. (For an introduction into these concepts see SN-ch. 2.)

In the definition of a *true outcome variable*, we condition on a value x of X that has a positive probability and on a global potential confounder C_X. In this way we control *for all* potential confounders. Note that in this and the next chapter, we are still in the process of defining the parameters to be estimated in empirical studies on causal effects. Defining these parameters is necessary if we want to study the conditions under which these parameters can in fact be estimated empirically.

Assuming $P(X=x)>0$, a true outcome variable τ_x has been defined such that its values are the conditional expectation values $E(Y|X=x,C_X=c)$ of the outcome variable Y holding constant X at the value x and C_X at a value c. We do not condition on intermediate variables, and this is why we only consider *total effects*. Ignoring intermediate variables implies that there can be a positive $(X=x,C_X=c)$-conditional variance of the outcome variable Y, for example, due to mediator variables and events that may occur in between X and Y. [2] A limitation of true outcome variables is that they are defined only for values x of X that have a positive probability.

In chapter 5 we use true outcome variables to define various kinds of causal conditional and average total effects. In chapter 6, true outcome variables are used to introduce the concept of unbiasedness, a first causality condition, that is, a condition allowing to identify causal total effects from empirically estimable parameters. In chapters 7 to **??**, we study various other causality conditions that imply unbiasedness.

4.4 Exercises

▷ **Exercise 4-1** Consider the example presented in Table 3.1 and show that the σ-algebras generated by X and generated by h_2 *are* identical.

▷ **Exercise 4-2** Show that $1_{U=Joe}$ defined in Example 4.10 is a global potential confounder of X.

[2] This is also the reason why Rubin's potential outcome variables are inadequate. He assumes that the value of a potential outcome variable is fixed if we condition on a treatment x and a person u. This is a contradiction to the idea that intermediate variables might also affect the outcome variable Y. True outcome variables remedy this deficiency.

Box 4.1 Glossary of new concepts

(Total effects) causality space $\big((\Omega,\mathscr{A},P),(\mathscr{F}_t,t\in T),X,Y\big)$	The formal framework under which we can define causal total effects comparing a value x of X to another value x' of X with respect to the outcome variable Y.
Potential confounder	Under the assumptions of Definition 4.1, this is a random variable W on $(\Omega,\mathscr{A},P)$ with value space $(\Omega'_W,\mathscr{A}'_W)$ for which there is a mapping $g\colon \Omega_1\to\Omega'_W$ such that $W=g\circ h_1$ and $g^{-1}(\mathscr{A}'_W)\subset\mathscr{A}_1$.
Global potential confounder	A potential confounder of X (often denoted by C_X) generating the same σ-algebra as h_1.
True outcome variable τ_x	A version of the C_X-conditional expectation of Y with respect to the conditional probability measure $P^{X=x}$. That is, $\tau_x=E^{X=x}(Y\mid C_X)$. With a global potential confounder C_X we condition on *all* potential confounders of X.

▷ **Exercise 4-3** What are the assumptions based on which we can define a true outcome variable?

▷ **Exercise 4-4** What does it mean when we assume that a true outcome variable τ_x is uniquely defined up to P-equivalence?

▷ **Exercise 4-5** Which are the values of the true outcome variable τ_0 and of the conditional expectation $E(Y\mid X,U)$ for $\omega_4=(Joe,yes,+)$ in the example presented in Table 3.1?

▷ **Exercise 4-6** Which are the elements of the σ-algebra $\sigma(h_1)$ in Example 4.10?

▷ **Exercise 4-7** Compute the values of the true outcome variable $\tau_0=E^{X=0}(Y\mid U)$ in the example presented in Table 3.2.

Solutions

▷ **Solution 4-1** There are four different inverse images of sets $B\in\mathscr{B}$ under X:

$$\forall B\in\mathscr{B}\colon\quad X^{-1}(B)\;=\;\begin{cases}\{\omega_3,\omega_4,\omega_7,\omega_8\}, & \text{if } 0\notin B \text{ and } 1\in B\\ \{\omega_1,\omega_2,\omega_5,\omega_6\}, & \text{if } 0\in B \text{ and } 1\notin B\\ \Omega, & \text{if } 0\in B \text{ and } 1\in B\\ \emptyset, & \text{if } 0\notin B \text{ and } 1\notin B.\end{cases}$$

These four inverse images are the elements of $\sigma(X)=X^{-1}(\mathscr{B})$. Furthermore, we choose the σ-algebra $\mathscr{A}_2=\big\{\{no\},\{yes\},\Omega_2,\emptyset\big\}$, which satisfies the requirements made in the assumptions of Definition 4.1. There are four different inverse images of sets $A\in\mathscr{A}_2$ under h_2, namely

$$\begin{aligned}h_2^{-1}(\{yes\}) &= \{\omega_3,\omega_4,\omega_7,\omega_8\},\\ h_2^{-1}(\{no\}) &= \{\omega_1,\omega_2,\omega_5,\omega_6\},\\ h_2^{-1}(\Omega_2) &= \Omega,\end{aligned}$$

$$h_2^{-1}(\emptyset) = \emptyset.$$

These four sets are the elements of the σ-algebra $h_2^{-1}(\mathscr{A}_2)$ generated by h_2. Hence, $X^{-1}(\mathscr{B}) = h_2^{-1}(\mathscr{A}_2)$.

▷ **Solution 4-2** According to Definition 4.4 (i), each random variable $W\colon \Omega \to \Omega'_W$ on $(\Omega, \mathscr{A}, P)$ for which there is a mapping $g\colon \Omega_1 \to \Omega'_W$ such that $W = g \circ h_1$ and $g^{-1}(\mathscr{A}'_W) \subset \mathscr{A}_1$ is a potential confounder. In this example Ω_U takes the role of Ω_1 and $\mathscr{P}(\Omega_U)$ the role of $\mathscr{A}_1$. If we choose $\{0,1\}$ to be the co-domain of the indicator $1_{U=Joe}$, then the mapping $g\colon \Omega_U \to \{0,1\}$ is defined by

$$g(Ann) = 0 \quad \text{and} \quad g(Joe) = 1.$$

Obviously, g is such that $1_{U=Joe} = g \circ U$ and $g^{-1}\left(\mathscr{P}(\{0,1\})\right) = \mathscr{P}(\Omega_U)$ [see Eq. (3.15)].

▷ **Solution 4-3** *First*, we assume that the random experiment considered in an empirical application is represented by the probability space $(\Omega, \mathscr{A}, P)$. The set Ω of possible outcomes of the random experiment is structured such that $\Omega = \Omega_1 \times \Omega_2 \times \Omega_3$. The set Ω_1 is chosen such that the σ-algebra generated by the projection $h_1\colon \Omega \to \Omega_1$ contains all events that cannot be caused by X. The set Ω_2 is chosen such that X is measurable with respect to $h_2\colon \Omega \to \Omega_2$, and Ω_3 is chosen such that Y is measurable with respect to $h_3\colon \Omega \to \Omega_3$. *Second*, there are two random variables on the probability space $(\Omega, \mathscr{A}, P)$, say X and Y, where X represents the cause and Y the outcome variable.

▷ **Solution 4-4** By definition, there may be different versions of a true outcome variable. In general, two such versions τ_x and τ_x^* are identical almost surely with respect to the probability measure $P^{X=x}$. A more precise formulation is Equation (3.65) for $V_x = \tau_x$ and $V_x^* = \tau_x^*$. If we additionally assume that τ_x is P-unique, then two versions τ_x and τ_x^* are identical almost surely with respect to the probability measure P. Again, see the more precise formulation in Equation (3.66).

▷ **Solution 4-5** The value $\tau_0(\omega_4)$ of the true outcome variable τ_0 is $E(Y \mid X=0, U=Joe) = .7$. In contrast, the value $E(Y \mid X, U)(\omega_4)$ of the conditional expectation $E(Y \mid X, U)$ is $E(Y \mid X=1, U=Joe) = .8$ (see the fourth row in Table 3.1).

▷ **Solution 4-6** The σ-algebra $\sigma(h_1)$ has four elements. Aside from Ω and $\emptyset$, these are the event

$$A = \{Joe\} \times \Omega_X \times \Omega_Y = \big\{(Joe, no, -), (Joe, yes, -), (Joe, no, +), (Joe, yes, +)\big\}$$

that Joe is drawn and the event

$$A^c = \{Ann\} \times \Omega_X \times \Omega_Y = \big\{(Ann, no, -), (Ann, yes, -), (Ann, no, +), (Ann, yes, +)\big\}$$

that Ann is drawn.

▷ **Solution 4-7** The values of the true outcome variable $\tau_0 = E^{X=0}(Y \mid U)$ are the two conditional expectation values $E(Y \mid X=0, U=Joe)$ and $E(Y \mid X=0, U=Ann)$. Because Y is binary, $E(Y \mid X=0, U=u) = P(Y=1 \mid X=0, U=u)$, and these conditional probabilities can be computed from the probabilities of the elementary events presented in Table 3.2 as follows:

$$P(Y=1 \mid X=0, U=Joe) = \frac{P(Y=1, X=0, U=Joe)}{P(X=0, U=Joe)} = \frac{.35}{.15+.35} = .7$$

and

$$P(Y=1 \mid X=0, U=Ann) = \frac{P(Y=1, X=0, U=Ann)}{P(X=0, U=Ann)} = \frac{.06}{.24+.06} = .2.$$

Chapter 5
Causal Total Effects

In chapter 4, we defined a *true outcome variable* $\tau_x = E^{X=x}(Y|C_X)$ as a version of the C_X-conditional expectation of Y with respect to the $(X=x)$-conditional probability measure $P^{X=x}$. With C_X we condition on *all* potential confounders of X. Although, in empirical applications, the *values* of such a true outcome variable are rarely estimable under realistic assumptions, the expectation of τ_x as well as the conditional expectation of τ_x given another random variable V, such as a covariate of X, *can* be estimated under appropriate and realistic assumptions.

We start this chapter defining a *true total effect variable* and a $(C_X=c)$*-conditional true causal total effect*. The expectation of the true total effect variable is then defined to be the *causal average total effect* of Y comparing x to x' (another value of X). Then we turn to the definition of a *causal conditional total effect* of treatment x compared to treatment x' given the value v of a random variable V, as well as a *causal conditional total effect function* comparing treatment x to treatment x' conditioning on a random variable V. Each of these kinds of conditional total effects or effect functions provides specific information that might be of interest in empirical causal research and evaluation studies. In the first place, these parameters and effect functions are of a purely theoretical nature. However, in the chapters to come we study how these various kinds of causal effects can be identified by empirically estimable parameters and how the causal effect functions can be identified by empirically estimable functions.

Talking about an effect or an effect function, we compare in some way or other two values x and x' of X to each other. This makes sense only if the quantities that we compare are uniquely defined, at least up to some minimal degree of uniqueness. For example, comparing $E(Y|X=x)$ to $E(Y|X=x')$ makes sense only if both conditional expectation values are uniquely defined, an assumption that does not necessarily hold (see, e. g., Def. 3.66 and Rem. 3.68). In contrast, if we assume $P(X=x)>0$ and $P(X=x')>0$, then $E(Y|X=x)$ and $E(Y|X=x')$ are uniquely defined [see Eq. (3.27)] and it is meaningful to compare these two numbers to each other. This kind of uniqueness issue is the reason why we include $P(X=x), P(X=x')>0$ in the assumptions, which will often be referred to in this chapter. Note, however, that this kind of assumption is not necessary for all causality conditions for conditional expectations such as $E(Y|X)$ or $E(Y|X,Z)$ (see, e. g., chs. 8 and 9).

Assumptions 5.1
Let $\big((\Omega,\mathscr{A},P),(\mathscr{F}_t,t\in T),X,Y\big)$ *be a causality space, let* $x,x'\in\Omega'_X$ *be two values of* X *such that* $P(X=x),P(X=x')>0$*, and let* Y *be real-valued such that* $E(Y^2)<\infty$*. Furthermore, let* C_X *denote a global potential confounder of* X*. Finally, let* τ_x *and* $\tau_{x'}$ *denote true outcome variables of* Y *given the values* x *and* x' *of* X*, respectively.*

5.1 True Total Effect Variable

Now we introduce the concepts of a *true* $(C_X{=}c)$*-conditional total effect* and a *true total effect variable of* Y *comparing* x *to* x', the latter being two different values of the focused cause X. The basic idea of these concepts is to hold constant all other possible causes of Y at one combination of their values and then compare the conditional expectation values of Y between the two values x and x' of X. The intuitive version of this basic idea goes back at least to John Stuart Mill (1843/1865). It is often referred to as the *ceteris paribus clause* (all other things equal). Remember that a global potential confounder can be a multivariate random variable, consisting of several univariate random variables (potential confounders). Also note that a global potential confounder of X does not comprise intermediate variables, that is, variables that might be affected by X and might have an effect on Y. This would be necessary only if we would define *direct* effects instead of *total* effects.

A true $(C_X{=}c)$-conditional total effect is a uniquely defined number, the difference

$$E(Y\,|\,X{=}x, C_X{=}c) \;-\; E(Y\,|\,X{=}x', C_X{=}c).$$

This difference may also be called *the true total effect on* Y *comparing* x *to* x' *given (the event)* $\{C_X{=}c\}$. The two conditional expectation values in this difference are uniquely defined if we assume $P(X{=}x, C_X{=}c) > 0$ and $P(X{=}x', C_X{=}c) > 0$ [see Eq. (3.27)]. For different values c and c' of the global potential confounder C_X, the true conditional total effects can differ from each other. This necessitates the second concept, a true total effect variable of Y comparing x to x', that is,

$$E^{X=x}(Y|C_X) \;-\; E^{X=x'}(Y|C_X) \;=\; \tau_x \;-\; \tau_{x'}.$$

In general, this random variable is not uniquely defined so that there can be many versions of such a true total effect variable. However, assuming that $\tau_x = E^{X=x}(Y|C_X)$ and $\tau_{x'} = E^{X=x'}(Y|C_X)$ are P-unique implies that their difference is P-unique as well [see SN-Box 14.1 (viii)].

Note that the concepts of a true total effect and a true total effect variable are of a theoretical nature and can be estimated only under rather restrictive assumptions. However, other causal total effects such as the expectation of $\tau_x - \tau_{x'}$ can be estimated under less restrictive assumptions.

Definition 5.2 (True Total Effect and True Total Effect Variable)
Let the Assumptions 5.1 hold.

(i) *If c is a value of C_X such that $P(X{=}x, C_X{=}c)$, $P(X{=}x', C_X{=}c) > 0$, then we call*

$$CTE_{C_X;xx'}(c) \;:=\; E(Y\,|\,X{=}x, C_X{=}c) - E(Y\,|\,X{=}x', C_X{=}c) \tag{5.1}$$

the true total effect on Y given the value c of C_X comparing x to x'.

(ii) *Assume that τ_x and $\tau_{x'}$ are P-unique. Then we call $CTE_{C_X;xx'}\colon \Omega'_{C_X} \to \mathbb{R}$ a version of the true total effect function comparing x to x' (with respect to Y), if*

$$CTE_{C_X;xx'}(C_X) \;\underset{P}{=}\; \tau_x - \tau_{x'} \tag{5.2}$$

holds for the composition $CTE_{C_X;xx'}(C_X)$ of C_X and $CTE_{C_X;xx'}$.

(iii) We call the composition $CTE_{C_X;xx'}(C_X)$ a version of the true total effect variable comparing x to x' (with respect to Y).

Remark 5.3 ($CTE_{C_X;xx'}(C_X)$ Versus $CTE_{C_X;xx'}$) While $CTE_{C_X;xx'}(C_X)$ is a random variable on the probability space $(\Omega, \mathscr{A}, P)$, which assigns values to all $\omega \in \Omega$, the function $CTE_{C_X;xx'}$ is a random variable on $(\Omega'_{C_X}, \mathscr{A}'_{C_X}, P_{C_X})$, which assigns values to all $c \in \Omega'_{C_X}$. From a substantive point of view the two mappings contain the same information. ◁

Remark 5.4 (Uniqueness of a True Total Effect Variable) For simplicity, we denote a true total effect variable $CTE_{C_X;xx'}(C_X)$ also by

$$\delta_{xx'} \;=\; \tau_x - \tau_{x'}\,. \tag{5.3}$$

It is a function on Ω and a random variable on $(\Omega, \mathscr{A}, P)$. It is not necessarily unique. However, $P(X{=}x), P(X{=}x') > 0$ and P-uniqueness of τ_x and $\tau_{x'}$ imply that the difference variable $\delta_{xx'}$ is also P-unique [see SN-Box 14.1 (viii)]. That is, two versions of such a difference variable are identical P-almost surely. Hence, if $\delta_{xx'}$ and $\delta^*_{xx'}$ are two such versions, then $E(\delta_{xx'}) = E(\delta^*_{xx'})$ [see SN-Box 6.1 (ix) with $A = \Omega$]. Other implications are that $\delta_{xx'}$ and $\delta^*_{xx'}$ have identical distributions, variances, and covariances with other random variables (see Rem. 3.57). ◁

Remark 5.5 (Uniqueness of a True Total Effect Function) In contrast to $\delta_{xx'}$, which is a function on Ω, a true total effect *function* $CTE_{C_X;xx'}$ is a function on Ω'_{C_X}. It is also not necessarily unique. However, if there are two versions of $CTE_{C_X;xx'}$, then they are identical P_{C_X}-almost surely. This follows from $\delta_{xx'} \underset{P}{=} \delta^*_{xx'}$ (see Rem. 5.4) and SN-Theorem 2.86. ◁

Remark 5.6 (Values of a True Total Effect Variable) As mentioned above, $\delta_{xx'}$ is a random variable on $(\Omega, \mathscr{A}, P)$. According to Equations (4.7), (3.69), and (3.70),

$$\begin{aligned} \forall\, \omega \in \Omega\colon \quad \delta_{xx'}(\omega) \;&=\; E^{X=x}(Y|C_X)(\omega) - E^{X=x'}(Y|C_X)(\omega) \\ &=\; E^{X=x}(Y\,|\,C_X{=}c) - E^{X=x'}(Y\,|\,C_X{=}c) \\ &=\; E(Y\,|\,X{=}x, C_X{=}c) - E(Y\,|\,X{=}x', C_X{=}c), \quad \text{if } \omega \in \{C_X{=}c\}. \end{aligned} \tag{5.4}$$

Hence, a value of a true total effect variable $\delta_{xx'}$ is the difference between the conditional expectation values of Y given the values (x,c) and (x',c) of (X, C_X). If c is a value of C_X such that $P(X{=}x, C_X{=}c), P(X{=}x', C_X{=}c) > 0$, then

$$\forall\, \omega \in \Omega\colon \quad \delta_{xx'}(\omega) \;=\; CTE_{C_X;xx'}(c), \quad \text{if } \omega \in \{C_X{=}c\} \tag{5.5}$$

[see Def. 5.2 (i) and Def. 3.87]. In this case, this value is identical for all versions of the true total effect function $CTE_{C_X;xx'}$ and for all versions of the true total effect variable $CTE_{C_X;xx'}(C_X) = \delta_{xx'}$. ◁

Remark 5.7 (Probabilistic Ceteris Paribus Clause) Considering a value $CTE_{C_X;xx'}(c)$ is tantamount to comparing x to x' (with respect to the outcome variable Y) keeping constant the global potential confounder C_X, and with it, *keeping constant all potential confounders of X*. Keeping constant C_X is the translation of the *ceteris paribus clause* for total effects into probability theory. ◁

5.2 Causal Average Total Effect

Now we define a *causal average total effect* by the expectation

$$E(\delta_{xx'}) \;=\; E(\tau_x - \tau_{x'}) \;=\; E(\tau_x) \;-\; E(\tau_{x'}) \tag{5.6}$$

of a true total effect variable [see Eq. (5.3) and SN-Box 6.1 (vi)].

Note that the causal average total effect is an *expectation* of a true total effect variable δ_{10}; it is not an unweighted average as the name might suggest.

As mentioned before, this expectation can be estimated under assumptions that are less restrictive than those that allow estimating the total effect variable $\delta_{xx'}$ itself. This will be detailed in the chapters on unconfoundedness and its sufficient conditions.

Definition 5.8 (Causal Average Total Effect)
Let the Assumptions 5.1 hold and let $\delta_{xx'} = \tau_x - \tau_{x'}$. Furthermore, assume that τ_x and $\tau_{x'}$ are P-unique. Then

$$ATE_{xx'} \;:=\; E(\delta_{xx'}), \tag{5.7}$$

is called the causal average total effect on Y comparing x to x' (with respect to P).

Taking the expectation of $\delta_{xx'}$ (with respect to P) means that we consider the average total effect with respect to the measure P, that is,

$$ATE_{xx'} \;=\; E(\delta_{xx'}) \;=\; E\big(CTE_{C_X;xx'}(C_X)\big). \tag{5.8}$$

In principle, we can also consider the average total effect with respect to another measure than P (see Rem. 5.24).

Remark 5.9 (Expectation of $CTE_{C_X;xx'}$ With Respect to the Distribution of C_X) According to SN-Theorem 6.13, $ATE_{xx'}$ is identical to the expectation of a true total effect function $CTE_{C_X;xx'}$ with respect to the distribution of the global potential confounder C_X, that is,

$$ATE_{xx'} \;=\; E(\delta_{xx'}) \;=\; E_{C_X}(CTE_{C_X;xx'}). \tag{5.9}$$

◁

Remark 5.10 (P-Uniqueness of τ_x and $\tau_{x'}$) In Definition 5.8, we assume $P(X{=}x) > 0$ and $P(X{=}x') > 0$ as well as P-uniqueness of τ_x and $\tau_{x'}$. According to Theorem 3.84, P-uniqueness of τ_x and $\tau_{x'}$ is equivalent to absolute continuity of P_{C_X} with respect to $P_{C_X|X=x}$ and $P_{C_X|X=x'}$, and to

$$P(X{=}x\,|\,C_X) \gtrsim_P 0 \quad \text{and} \quad P(X{=}x'\,|\,C_X) \gtrsim_P 0.$$

According to the same theorem, P-uniqueness of τ_x is also equivalent to

$$\forall\, \tau_x, \tau_x^* \in \mathscr{E}^{X=x}(Y\,|\,C_X): \quad E(\tau_x) \;=\; E(\tau_x^*), \tag{5.10}$$

because we presume that Y is real-valued with a finite second moment (see Ass. 5.1 and Exercise 5-6.) Hence, under the assumptions of Definition 5.8, the expectation $E(\tau_x)$ is a

uniquely defined number, and the same applies to $\tau_{x'}$ and its expectation $E(\tau_{x'})$.[1] This implies that, under the assumptions of Definition 5.8, $E(\delta_{xx'})$ is a uniquely defined number as well. If the true outcome variables τ_x and $\tau_{x'}$ are not P-unique, then the causal average total effect is not defined. In this case we also say that it does not exist. ◁

Remark 5.11 (Substantive Meaning) If X represents a treatment variable, then $ATE_{xx'}$ is also called the 'average causal effect' or the '(causal) average treatment effect' comparing treatment x to treatment x', which is unambiguous as long as no direct and/or indirect treatment effects are discussed in the same context. An average total effect is among the parameters that one might want to estimate from sample data. ◁

Remark 5.12 (Hypothesis of a t-Test in a Randomized Experiment) The causal average total effect is what is tested by a t-test of the hypothesis $\mu_0 = \mu_1$ [in our notation $E(Y|X=0) = E(Y|X=1)$] in an experiment with randomized assignment of an observational unit to a treatment condition (see Cor. 8.17). The expectations $E(\tau_x)$ and $E(\tau_{x'})$ occurring in Equation (5.6) are estimated, for example, by the sample group means in such a randomized experiment involving two treatment conditions x and x'. If certain assumptions hold, then these expectations are also estimated by the *adjusted means* in analysis of covariance and its generalizations, as well as in procedures based on propensity scores. (For more details see the chapters to come.) ◁

Example 5.13 (Joe and Ann With Randomized Assignment) In Table 3.1 we presented an example in which the person sampled is assigned to one of the two treatment conditions by randomization. This implies that $P(X{=}x|U{=}u) = P(X{=}x)$ for all persons $u \in \Omega_U$, that is, all persons have the same probability to be assigned to treatment x. In example 4.21, we specified the true outcome variables $\tau_0 = g_0(U)$ and $\tau_1 = g_1(U)$ and their values $g_0(u)$ and $g_1(u)$ for this example.

Now we illustrate the *causal average total effect* comparing treatment to control, which can be computed as follows:

$$
\begin{aligned}
ATE_{10} \;&=\; E\big(E^{X=1}(Y|U) - E^{X=0}(Y|U)\big) \;=\; E(\tau_1 - \tau_0) \;=\; E(\tau_1) - E(\tau_0) \\
&= \sum_u g_1(u)\cdot P(U{=}u) - \sum_u g_0(u)\cdot P(U{=}u) \\
&= P(Y{=}1\,|\,X{=}1, U{=}\mathit{Joe})\cdot P(U{=}\mathit{Joe}) \;+\; P(Y{=}1\,|\,X{=}1, U{=}\mathit{Ann})\cdot P(U{=}\mathit{Ann}) \\
&\quad - \big(P(Y{=}1\,|\,X{=}0, U{=}\mathit{Joe})\cdot P(U{=}\mathit{Joe}) \;+\; P(Y{=}1\,|\,X{=}0, U{=}\mathit{Ann})\cdot P(U{=}\mathit{Ann})\big) \\
&= .80\cdot.50 \;+\; .40\cdot.50 \;-\; (.70\cdot.50 \;+\; .20\cdot.50) \;=\; .15,
\end{aligned}
$$

using the transformation theorem (see SN-Th. 6.13 and SN-Rem. 6.15).

Figure 5.1 visualizes the causal average total effect, which, in this example, is identical to the difference

$$E(Y\,|\,X{=}1) - E(Y\,|\,X{=}0) \;=\; P(Y{=}1\,|\,X{=}1) - P(Y{=}0\,|\,X{=}0) \;=\; .6 - .45 \;=\; .15$$

(see ch. 8 for the reason why $E(Y\,|\,X{=}1) - E(Y\,|\,X{=}0) = ATE_{10}$). Figure 5.1 also illustrates various conditional probabilities of success. The points marked by the dashed line represent the probabilities $P(Y{=}1\,|\,U{=}\mathit{Joe}, X{=}1)$ and $P(Y{=}1\,|\,U{=}\mathit{Joe}, X{=}0)$ of success for

[1] The expectation $E(\tau_x)$ corresponds to the term $E[Y|do(x)]$ in Pearl's and to $E(Y_x)$ in Rubin's terminologies (see, e. g., Pearl, 2009, p. 108 and Rubin, 2005, p. 323).

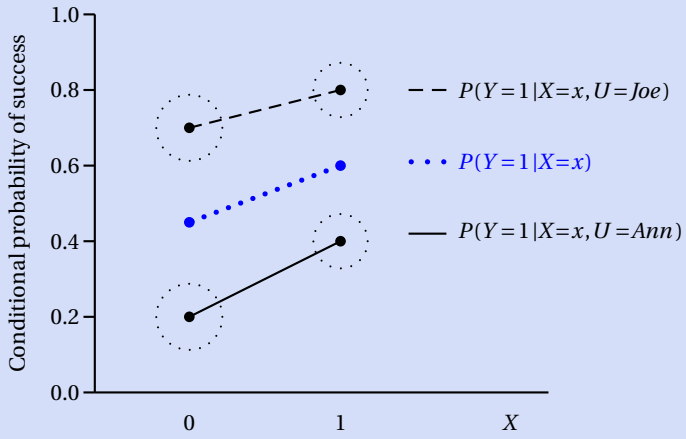

Figure 5.1. Conditional probabilities of success given treatment and person

Joe given that he is treated and given that he is not treated, respectively. Similarly, the two points marked by the solid line indicate the probabilities $P(Y=1\,|\,U=Ann\,,X=1)$ and $P(Y=1\,|\,U=Ann\,,X=0)$ of success for Ann given that she is treated and given that she is not treated, respectively. The points marked by the dotted line represent the conditional probabilities $P(Y=1\,|\,X=1)$ and $P(Y=1\,|\,X=0)$ of success given treatment and given control, respectively. The size of the area of the dotted circles is proportional to the conditional probabilities $P(U=u\,|\,X=x)$ that are used in the computation of the conditional expectation values

$$E(Y\,|\,X=x) \;=\; \sum_{u} E(Y\,|\,X=x,U=u)\cdot P(U=u\,|\,X=x) \tag{5.11}$$

[see SN-Box 9.2 (ii)]. ◁

5.3 Conditional Total Effect and Conditional Total Effect Function

So far we considered the random variables X, Y, and C_X, a global potential confounder of X. Now we bring into play an additional random variable V on $(\Omega,\mathscr{A},P)$ and introduce the concepts of the *causal conditional total effect* given the value v of a V and of a *causal V-conditional total effect function*. Often V is a covariate of X, such as a pretest that measures the 'same' attribute as the outcome variable Y, only prior to treatment. In other examples, V could be X or any other random variable on $(\Omega,\mathscr{A},P)$. First, we explain the assumptions based on which we can introduce these concepts, present their definitions, and then turn to re-aggregating conditional effects.

5.3.1 Notation, Assumptions and Definitions

Assume that V is a random variable on $(\Omega, \mathscr{A}, P)$ with value space $(\Omega'_V, \mathscr{A}'_V)$, let $v \in \Omega'_V$ such that $P(V=v) > 0$, and let $P^{V=v}$ denote the $(V=v)$-conditional probability measure on $(\Omega, \mathscr{A})$ [see Def. 3.17 for $B = \{V=v\} = \{\omega \in \Omega\colon V(\omega) = v\}$]. In Definition 5.17, we will define a $(V=v)$-*conditional total effect* to be the conditional expectation value $E(\tau_x - \tau_{x'} | V=v)$.

Remark 5.14 ($P^{V=v}$-Uniqueness of a True Total Effect Variable) If $P(V=v) > 0$, then $\tau_x = E^{X=x}(Y|C_X)$ is called $P^{V=v}$-*unique*, if

$$P^{V=v}\big(\{\omega \in \Omega\colon \tau_x(\omega) \neq \tau_x^*(\omega)\}\big) \;=\; 0, \quad \forall\, \tau_x, \tau_x^* \in \mathscr{E}^{X=x}(Y|C_X)\,. \tag{5.12}$$

◁

Remark 5.15 (An Implication of $P^{V=v}$-Uniqueness of τ_x and $\tau_{x'}$) The conditional expectation value $E(\tau_x - \tau_{x'}|V=v)$ is a uniquely defined number if $P(V=v) > 0$ and we assume $P^{V=v}$-uniqueness of τ_x and $\tau_{x'}$. In other words, if τ_x, $\tau_{x'}$ are $P^{V=v}$-unique and $\delta_{xx'}$, $\delta^*_{xx'}$ are different versions of the true total effect variable $CTE_{C_X;xx'}(C_X)$ [see Def. 5.2 (iii)], then

$$E^{V=v}(\delta_{xx'}) \;=\; E^{V=v}(\delta^*_{xx'}) \;=\; E(\delta_{xx'}|V=v) \;=\; E(\delta^*_{xx'}|V=v) \tag{5.13}$$

[see SN-Box 14.1 (viii)]. ◁

Remark 5.16 (P-Uniqueness Implies $P^{V=v}$-Uniqueness) If $P(V=v) > 0$, then P-uniqueness of $E^{X=x}(Y|C_X)$ implies that $E^{X=x}(Y|C_X)$ is also $P^{V=v}$-unique [see SN-Box 14.1 (v)]. In contrast, $P^{V=v}$-uniqueness of $E^{X=x}(Y|C_X)$ does not imply that it is also P-unique. Hence, $P^{V=v}$-uniqueness of $\tau_x = E^{X=x}(Y|C_X)$ is a weaker assumption than P-uniqueness of τ_x. ◁

Definition 5.17 (Causal Conditional Total Effect Function)
Let the Assumptions 5.1 hold and let V be a random variable on $(\Omega, \mathscr{A}, P)$ with value space $(\Omega'_V, \mathscr{A}'_V)$.

(i) Assume that $P(V=v) > 0$ and that τ_x and $\tau_{x'}$ are $P^{V=v}$-unique. Then we call

$$CTE_{V;xx'}(v) \;:=\; E(\tau_x - \tau_{x'}|V=v) \tag{5.14}$$

the causal $(V=v)$-conditional total effect on Y comparing x to x'.

(ii) Assume that τ_x and $\tau_{x'}$ are P-unique. If the function $CTE_{V;xx'}\colon \Omega'_V \to \mathbb{R}$ is such that $CTE_{V;xx'}{}^{-1}(\mathscr{B}) \subset \mathscr{A}'_V$ and

$$CTE_{V;xx'}(V) \;\underset{P}{=}\; E(\tau_x - \tau_{x'}|V), \tag{5.15}$$

then $CTE_{V;xx'}$ is called a version of the causal V-conditional total effect function comparing x to x' (with respect to Y).

(iii) The composition $CTE_{V;xx'}(V)$ is called a version of the causal V-conditional total effect variable comparing x to x' (with respect to Y).

Remark 5.18 (A Characterization of $CTE_{V;xx'}(V)$) SN-Theorem 2.49 and SN-Remark 10.14 imply that $CTE_{V;xx'}(V)$ is V-measurable and is a version of the V-conditional expectation of $\tau_x - \tau_{x'}$, that is,

$$CTE_{V;xx'}(V) \in \mathscr{E}(\tau_x - \tau_{x'}|V) \tag{5.16}$$

[see SN-Prop. (10.12)]. ◁

Remark 5.19 ($CTE_{V;xx'}(V)$ Versus $CTE_{V;xx'}$) While $CTE_{V;xx'}(V)$ is a random variable on the probability space $(\Omega, \mathscr{A}, P)$, which assigns values to all $\omega \in \Omega$, the function $CTE_{V;xx'}$ is a random variable on $(\Omega'_V, \mathscr{A}'_V, P_V)$, which assigns values to all $v \in \Omega'_V$. It is the *factorization of* the conditional expectation $CTE_{V;xx'}(V)$. ◁

Remark 5.20 (Conditioning on C_X Versus Conditioning on V) So far, we considered two kinds of random variables on which we condition. The first and most fine-grained one is C_X, a global potential confounder of X. Such a variable is essential for the theory and in particular for translating the ceteris paribus clause into the language of probability theory. It has been used in chapter 4 to define a true outcome variable $\tau_x = E^{X=x}(Y|C_X)$. Estimating the true outcome variables τ_x and $\tau_{x'}$ or their difference $\delta_{xx'}$ requires strong assumptions. Considering a causal V-conditional total effect function, we reaggregate the true total effect variable. This yields a less fine-grained or coarsened total effect variable, but it is still a *causal* conditional total effect variable. In contrast to a true total effect variable, a V-conditional total effect function often can be identified under realistic assumptions by empirically estimable conditional expectations (see, e. g., chs. 6 to **??**). ◁

Remark 5.21 ($E(\tau_x|V)$ Versus $E^{X=x}(Y|V)$) Note the distinction between the V-conditional expectations $E(\tau_x|V)$ and $E(\tau_{x'}|V)$ on one side and the two V-conditional expectations $E^{X=x}(Y|V)$ and $E^{X=x'}(Y|V)$ on the other side. The difference between the first two conditional expectations is a causal V-conditional total effect variable, that is,

$$CTE_{V;xx'}(V) \underset{P}{=} E(\tau_x - \tau_{x'}|V) \underset{P}{=} E(\tau_x|V) - E(\tau_{x'}|V). \tag{5.17}$$

In contrast, in general, the conditional expectations $E^{X=x}(Y|V)$ and $E^{X=x'}(Y|V)$, and their difference have no causal meaning. The difference $E^{X=x}(Y|V) - E^{X=x'}(Y|V)$ is just a prima facie effect function, which can be seriously misleading if erroneously interpreted as a causal total effect variable. ◁

Remark 5.22 (Coarsening the True Total Effect Function $\tau_x - \tau_{x'}$) With $E(\tau_x|V)$ we coarsen (or reaggregate) the true outcome variable $\tau_x = E^{X=x}(Y|C_X)$. Conditioning on the global potential confounder C_X we control for all potential confounders of X. Therefore the conditional expectations $E^{X=x}(Y|C_X)$ and $E^{X=x'}(Y|C_X)$ inform us how Y depends on the values x and x' controlling for all potential confounders of X. Hence, considering the V-conditional expectation of the difference variable $\tau_x - \tau_{x'}$ does not introduce bias. It just coarsens the most fine-grained total effects to causal total effects that are less fine-grained. In contrast, considering the conditional expectations $E^{X=x}(Y|V)$ and $E^{X=x'}(Y|V)$ and their difference, we only control for V, possibly neglecting important potential confounders. [In chapter 6 we define $E^{X=x}(Y|V)$ to be *unbiased* or *biased* depending on whether or not $E^{X=x}(Y|V) \underset{P}{=} E(\tau_x|V)$.] ◁

Remark 5.23 (Values of a Causal Conditional Total Effect Variable) Note that $CTE_{V;xx'}(V)$, the composition of the conditional total effect function $CTE_{V;xx'}$ and V, is a random variable on $(\Omega, \mathscr{A}, P)$, and according to Equation (3.51),

$$\begin{aligned}\forall\ \omega\in\Omega\colon\quad CTE_{V;xx'}(V)(\omega) &= E(\tau_x-\tau_{x'}\,|\,V)(\omega)\\ &= E(\tau_x\,|\,V)(\omega)-E(\tau_{x'}\,|\,V)(\omega)\\ &= E(\tau_x\,|\,V=v)-E(\tau_{x'}\,|\,V=v),\quad \text{if } v\in\{V=v\}\\ &= E(\tau_x-\tau_{x'}\,|\,V=v)\\ &= CTE_{V;xx'}(v).\end{aligned}\tag{5.18}$$

Hence, a value of the causal conditional total effect variable $CTE_{V;xx'}(V)$ is the difference between a $(V=v)$-conditional expectation value of τ_x and of $\tau_{x'}$. If v is a value of V such that the assumptions of Definition 5.17 (i) are satisfied, then $CTE_{V;xx'}(v)$ is uniquely defined and it is identical to the $(V=v)$-conditional expectation value of $\tau_x-\tau_{x'}$ given the event $\{V=v\}=\{\omega\in\Omega\colon V(\omega)=v\}$. ◁

Remark 5.24 (Average Total Effect With Respect to $P^{V=v}$) In Definition 5.17 (i) it is assumed that $P(V=v)>0$ and that τ_x and $\tau_{x'}$ are $P^{V=v}$-unique. Therefore, the causal $(V=v)$-conditional total effect $CTE_{V;xx'}(v)$ on Y comparing x to x' is identical to the causal average total effect on Y comparing x to x' with respect to the measure $P^{V=v}$. ◁

Remark 5.25 (Causal Conditional Versus Causal Average Total Effects) A causal conditional total effect variable is more informative than the causal average total effect. If V is a mapping of the observational-unit variable U such as $V :=$ *sex* or $V :=$ *educational status*, then the causal $(V=v)$-conditional total effect is the causal average total effect given that we sample a person *from the subpopulation* represented by the value v of V. If the variable V is a pretest that measures the 'same' (e. g., *live satisfaction*) as the outcome variable Y (the post-test), but prior to the onset of the treatment, then comparing the conditional total effects $CTE_{V;xx'}(v)$ and $CTE_{V;xx'}(v')$ shows if these conditional total effects are different for different values v and v' of this pretest. If they are, then the numbers $CTE_{V;xx'}(v)$ and $CTE_{V;xx'}(v')$ may inform about the differential indication of the treatment. That is they answer questions such as "Which treatment is good for which kind of persons?" ◁

5.3.2 Causal $(X=x^*)$-Conditional Total Effect

A special case of a $(V=v)$-conditional effect is the $(X=x^*)$-conditional effect comparing x to x'. In this case, the X does not only play the role of the focused cause, but also of the variable on whose values we condition. Note that x^* can be identical to x, x', or to a third value of X. For $V=X$ and $v=x^*$, Definition 5.17 (i) yields

$$CTE_{X;xx'}(x^*) \;=\; E(\tau_x-\tau_{x'}\,|\,X=x^*),\tag{5.19}$$

the *causal* $(X=x^*)$*-conditional total effect* on Y comparing x to x'.

Remark 5.26 (Substantive Meaning) Suppose X represents a treatment variable in an experiment or in a quasi-experiment. If there are two treatment conditions, treatment $(X=1)$ and control $(X=0)$, then we may consider $CTE_{X;10}(1)$, the $(X=1)$-conditional total effect comparing of treatment $(X=1)$ to control $(X=0)$, and $CTE_{X;10}(0)$, the $(X=0)$-conditional total effect comparing treatment $(X=1)$ to control $(X=0)$. These effects are also known as the 'average effect on the treated' and the 'average effect on the untreated', respectively. ◁

Remark 5.27 (Pre-Facto Perspective) At first sight, the concept of an $(X=x^*)$*-conditional total effect* comparing x to x' seems strange, in particular if $x^*=x'$. If, for example, x'

represents 'no treatment', how can we talk about the causal average (or conditional) total treatment effect on the *untreated*? Remember that we are not talking about data that resulted from an experiment — an interpretation that is suggested by the term 'treatment effect on the untreated'. Instead we are considering a random experiment *that is still to be conducted*, that is, we look at the random experiment from the *pre-facto perspective*. This is what probabilistic theories are about: a random experiment that is not yet conducted. Talking about the probability of an event does not make sense for an event that already occurred, unless we *do as if* it did not yet occur, that is, unless we take the pre-facto perspective. Hence, we can talk about a causal individual total effect although the individual is not yet treated and even if it will never be treated, just in the same way as we can talk about the probability of flipping 'heads', even if the coin is never flipped. ◁

Remark 5.28 (Causal $(X=x^*)$-Conditional Total Treatment Effects) The causal conditional total effects given a specific value x^* of the treatment variable X are often more informative than the causal average total effects, especially, if the X-conditional expectations of the true outcome variables τ_x and $\tau_{x'}$ actually depend on the values of X. If, however,

$$E(\tau_x|X) \underset{P}{=} E(\tau_x) \quad \text{and} \quad E(\tau_{x'}|X) \underset{P}{=} E(\tau_{x'}), \tag{5.20}$$

then

$$CTE_{X;xx'}(X) \underset{P}{=} E(\tau_x - \tau_{x'}|X) \underset{P}{=} E(\tau_x|X) - E(\tau_{x'}|X) \underset{P}{=} E(\tau_x) - E(\tau_{x'}) = ATE_{xx'}. \tag{5.21}$$

In this case the causal $(X=x^*)$-conditional total treatment effects $CTE_{X;xx'}(x^*)$ are identical for all values x^* of X for which $P(X=x^*)>0$. A sufficient condition of Proposition (5.20) is stochastic independence of X and the global potential confounder C_X (see Exercise 5-10), a condition that is created in the randomized experiment. (For more details see ch. 8.) ◁

Remark 5.29 ($CTE_{X;10}(1)$ Versus $CTE_{X;10}(0)$) Suppose we are interested in the effects of a treatment (represented by $X=1$) compared to a control (represented by $X=0$) with respect to the outcome variable Y, say *well-being*. Because there is no random assignment of persons to treatments, the persons that tend to take the treatment may differ in their well-being before treatment and in other pre-treatment variables from those who tend to be in the control condition. In this case, there might be large differences in the causal $(X=1)$-conditional total effect $CTE_{X;10}(1)$ compared to the causal $(X=0)$-conditional total effect $CTE_{X;10}(0)$. In this scenario the causal average total effect ATE_{10} would not be of much interest. The causal $(X=1)$-conditional effect $CTE_{X;10}(1)$ helps us evaluating how good the treatment is on average for those that tend to go to this treatment. In contrast, $CTE_{X;10}(0)$ informs us about the average effect of the treatment on those who tend to be in control — under the present side conditions determining the participation in the treatment. Hence, if the causal conditional total effect $CTE_{X;10}(1)$ is smaller than the causal conditional total effect $CTE_{X;10}(0)$, one may raise the question whether or not it would be worthwhile to change the regime of assigning units to treatment, if this regime is under our control. ◁

Remark 5.30 (Multivariate Random Variable V) Also note that the concept of a causal $(V=v)$-conditional total effect is not restricted to a univariate random variable V. Instead, $V=(V_1,\ldots,V_m)$ may also be an m-variate random variable such that a value $v=(v_1,\ldots,v_m)$ of V is an m-tupel of values of the random variables $V_1,\ldots,V_m$. ◁

5.3.3 Reaggregation

In the following theorem, we consider the expectation of a causal total effect function $CTE_{V;xx'}$. According to this theorem this expectation is identical to the causal average total effect on Y comparing x to x'.

Theorem 5.31 (Reaggregating a Conditional Total Effect Variable)
Let the assumptions of Definition 5.17 hold and let $CTE_{V;xx'}(V)$ denote a causal V-conditional total effect variable comparing x to x'. Then

$$E\big(CTE_{V;xx'}(V)\big) = ATE_{xx'}. \tag{5.22}$$

(Proof p. 118)

Remark 5.32 (Expectation With Respect to the Distribution of V) According to Theorem 5.31, the expectation of a causal V-conditional total effect function is identical to the average total effect. To emphasize, the expectation is with respect to the distribution of V, that is,

$$E\big(CTE_{V;xx'}(V)\big) = E_V(CTE_{V;xx'}) = ATE_{xx'}, \tag{5.23}$$

which immediately follows from SN-Equation (6.13). Taking this expectation means to reaggregate the $(V=v)$-conditional total effects to a single number, the causal average total effect. ◁

Remark 5.33 (The Proper Way of Reaggregation) Inserting the definition of $CTE_{V;xx'}(V)$ [see Eq. (5.15)] into Equation (5.22) yields

$$\begin{aligned}
& ATE_{xx'} \\
&= E\big(CTE_{V;xx'}(V)\big) && [(5.22)] \\
&= E\big(E(\tau_x - \tau_{x'}|V)\big) && [(5.15)] \\
&= E(\tau_x - \tau_{x'}) && [\text{SN-Box 10.2 (iv)}] \\
&= E\big(E^{X=x}(Y|C_X) - E^{X=x'}(Y|C_X)\big) && [(4.7)] \\
&= E\big(E^{X=x}(Y|C_X)\big) - E\big(E^{X=x'}(Y|C_X)\big) && [\text{SN-Box 6.1 (vii)}]
\end{aligned} \tag{5.24}$$

These equations reveal that reaggregation means to consider the expectation of V-conditional expectations. If Y is binary, then, according to the last of these equations, we take the expectation of C_X-conditional probabilities $P^{X=x}(Y=1|C_X)$ and $P^{X=x'}(Y=1|C_X)$, and not of their log odds ratios or other transformations of these probabilities. A reaggregation of such transformed probabilities does not yield the causal average total effect (see Exercise 5-12). ◁

Now we turn to a less rigorous reaggregation of a causal total effect variable $CTE_{V;xx'}(V)$, considering a W-conditional expectation of $CTE_{V;xx'}(V)$, which, according to this theorem, is a W-conditional total effect variable $CTE_{W;xx'}(W)$. Furthermore, the conditional expectation value $E\big(CTE_{V;xx'}(V)\,\big|\,W=w\big)$ is identical to the causal $(W=w)$-conditional total effect on Y comparing x to x', if we assume $P(W=w)>0$.

Theorem 5.34 (Reaggregating a Conditional Total Effect Variable)
Let the assumptions of Definition 5.17 hold and let $CTE_{V;xx'}(V)$ denote a causal V-conditional total effect variable on Y comparing x to x'. Furthermore, let W be a random variable on $(\Omega, \mathscr{A}, P)$ with value space $(\Omega'_W, \mathscr{A}'_W)$ and let $f\colon \Omega'_V \to \Omega'_W$ be a mapping such that $f^{-1}(\mathscr{A}'_W) \subset \mathscr{A}'_V$ and $W = f \circ V$.

(i) Then

$$E\left(CTE_{V;xx'}(V) \,\middle|\, W\right) \underset{P}{=} CTE_{W;xx'}(W). \tag{5.25}$$

(ii) If $w \in \Omega'_W$ is a value of W such that $P(W=w) > 0$, then

$$E\left(CTE_{V;xx'}(V) \,\middle|\, W=w\right) = CTE_{W;xx'}(w). \tag{5.26}$$

(Proof p. 118)

Remark 5.35 (Partial Reaggregation) According to Theorem 5.34 (i), the W-conditional expectation of a causal V-conditional total effect function is almost surely (with respect to P) identical to a causal W-conditional total effect function, provided that $W = f(V)$ is a measurable mapping of V. If $W = f(V)$ is a measurable mapping of V, then $\sigma(W) \subset \sigma(V)$ (see SN-Th. 2.49), and if $\sigma(W) \neq \sigma(V)$, then $E\left(CTE_{V;xx'}(V) \,\middle|\, W\right)$ may be called a *partial reaggregation* of the original causal total effect function $CTE_{V;xx'}(V)$. It is tantamount to *coarsening* the original causal total effect function $CTE_{V;xx'}(V)$ to a less fine-grained causal total effect function $CTE_{W;xx'}(W)$. ◁

Remark 5.36 (The Proper Way of Partial Reaggregation) Inserting the definition of the causal conditional total effect variable $CTE_{V;xx'}(V)$ [see Eq. (5.15)] into Equation (5.25) yields

$$\begin{aligned}
& CTE_{W;xx'}(W) \\
\underset{P}{=}\ & E\left(CTE_{V;xx'}(V) \,\middle|\, W\right) && \text{[(5.25)]} \\
\underset{P}{=}\ & E\left(E(\tau_x - \tau_{x'}|V) \,\middle|\, W\right) && \text{[(5.15)]} \\
\underset{P}{=}\ & E(\tau_x - \tau_{x'}|W) && \text{[SN-Box 10.2 (v)]} \\
\underset{P}{=}\ & E\left(E^{X=x}(Y|C_X) - E^{X=x'}(Y|C_X) \,\middle|\, W\right) && \text{[(4.7)]} \\
\underset{P}{=}\ & E\left(E^{X=x}(Y|C_X) \,\middle|\, W\right) - E\left(E^{X=x'}(Y|C_X) \,\middle|\, W\right) && \text{[SN-Box 10.2 (xv)]}
\end{aligned} \tag{5.27}$$

These equations reveal that partial reaggregation means to consider a W-conditional expectation of a V-conditional expectation, presuming that $W = f(V)$ is a measurable function of V. If Y is binary, then, according to the last of these equations, we take a W-conditional expectation of V-conditional probabilities $P^{X=x}(Y=1|C_X)$ and $P^{X=x'}(Y=1|C_X)$, and *not of their log odds ratios* or other transformations of these probabilities. A reaggregation of such transformed probabilities does not yield a causal conditional total effect function. ◁

5.4 Example: Joe and Ann With Bias at the Individual Level

Now we illustrate the various causal total effects by an example in which there are two treatment variables. Such a two-factorial experiment has already been discussed at an informal level in section 2.3 and example that is very similar has been presented in Table 4.1. However, now the outcome variable is not binary any more. Furthermore, we exemplify that bias can occur at the individual level at which the person variable is kept constant.

5.4.1 The Random Experiment

The parameters displayed in Table 5.1 refer to an experiment in which a person is drawn from a set of two persons, each of which may or may not receive a first treatment, say *group therapy* (represented by Z), and, at the same time, receive or not receive a second treatment, say *individual therapy* (represented by X). Furthermore, both treatments affect a quantitative outcome variable Y (e. g., *well-being, life satisfaction,* or a score on a *symptom checklist*).

The set of possible outcomes of the random experiment is

$$\Omega \;=\; \Omega_U \times \Omega_Z \times \Omega_X \times \Omega_Y,$$

where

$$\Omega_U \;:=\; \{\mathit{Joe}, \mathit{Ann}\},$$

$$\Omega_Z \;:=\; \{\mathit{group\ therapy}, \mathit{no\ group\ therapy}\},$$

$$\Omega_X \;:=\; \{\mathit{individual\ therapy}, \mathit{no\ individual\ therapy}\},$$

and Ω_Y is the set of possible observations based on which the score of the outcome variable Y is computed. If Y is discrete, then we may choose $\mathscr{A} = \mathscr{P}(\Omega)$, where $\mathscr{P}(\Omega)$ denotes the power set of Ω. However, if Y is continuous, then $\mathscr{A}$ is the product σ-algebra $\mathscr{A} = \mathscr{P}(\Omega_U) \otimes \mathscr{P}(\Omega_Z) \otimes \mathscr{P}(\Omega_X) \otimes \mathscr{B}$, where $\mathscr{B}$ denotes the Borel σ-algebra on $\mathbb{R}$ (see SN-section 1.2.3).

In this example, the individual *probabilities* for the two kinds of treatments are as follows: Joe receives group therapy ($Z=1$) with probability $P(Z=1\,|\,U=\mathit{Joe}) = 1/2$ or no group therapy ($Z=0$) with probability $P(Z=0\,|\,U=\mathit{Joe}) = 1/2$. Furthermore, Joe receives individual therapy ($X=1$) with probability $P(X=1\,|\,U=\mathit{Joe}, Z=0) = 3/4$ if he does not receive group therapy ($Z=0$), and with probability $P(X=1\,|\,U=\mathit{Joe}, Z=1) = 1/4$ if he receives group therapy ($Z=1$) as well. Similarly, Ann receives individual therapy ($X=1$) with probability $P(X=1\,|\,U=\mathit{Ann}, Z=0) = 3/4$ if she does not receive group therapy ($Z=0$), and with probability $P(X=1\,|\,U=\mathit{Ann}, Z=1) = 1/4$ if she also receives group therapy ($Z=1$).

5.4.2 Choosing the Focused Treatment Variable

There are *several* true total causal effects we might look at. In principle, we might be interested in

- (a_1) the individual total effect on Y of group therapy ($Z=1$) compared to no group therapy ($Z=0$) given that Joe (Ann) also receives individual therapy ($X=1$),
- (b_1) the individual total effect on Y of group therapy ($Z=1$) compared to no group therapy ($Z=0$) given that Joe (Ann) does *not* receive individual therapy ($X=0$), and

Table 5.1. Joe and Ann with bias at the individual level

		Fundamental parameters					Derived parameters		
Person variable U	*Group therapy* Z	$P(U=u)$	$P(Z=z\|U)$	$P(X=1\|U,Z)$	$\tau_0 = E^{X=0}(Y\|U,Z)$	$\tau_1 = E^{X=1}(Y\|U,Z)$	$\delta_{10} = \tau_1 - \tau_0$	$P^{X=0}(Z=z\|U)$	$P^{X=1}(Z=z\|U)$
Joe	0	1/2	1/2	3/4	68	82	14	1/4	3/4
	1		1/2	1/4	96	100	4	3/4	1/4
Ann	0	1/2	1/2	3/4	80	98	18	1/4	3/4
	1		1/2	1/4	104	106	2	3/4	1/4

(c_1) the average of these individual total effects, averaging over the two values of X (individual therapy).

Of course, the latter effect is certainly less informative than the two conditional effects.

Similarly, we may also be interested in

(a_2) the individual total effect on Y of individual therapy ($X=1$) compared to no individual therapy ($X=0$) given that Joe (Ann) also receives group therapy ($Z=1$),
(b_2) the individual total effect of individual therapy ($X=1$) compared to no individual therapy ($X=0$) on Y given that Joe (Ann) does *not* receive group therapy ($Z=0$), and
(c_2) the average of these individual total effects of individual therapy, averaging over the two values of Z (group therapy).

Looking at the effects (a_1) to (c_1), we consider *individual therapy* to be a (qualitative) covariate and *group therapy* to be the treatment variable asking for the conditional effects of group therapy given individual therapy and their average, the 'main effect' of group therapy. In contrast, looking at the effects (a_2) to (c_2), we consider *group therapy* to be a (qualitative) covariate and *individual therapy* to be the treatment variable.

In principle, both treatment variables, X and Z, may take the role of a covariate (and potential confounder), *depending on which treatment effects we are studying*, the effects of (values of) X on Y or the effects of (values of) Z on Y. In this example, we focus on X as a cause. In this case,

$$\Omega_1 = \Omega_U \times \Omega_Z, \quad \Omega_2 = \Omega_X, \quad \text{and} \quad \Omega_3 = \Omega_Y$$

[see Eq. (4.1)], and the bivariate random variable (U,Z) is a global potential confounder of X.

5.4.3 True Outcome Variables and True Total Effects

In the example presented in Table 5.1, $C_X = (U,Z)$ is a global potential confounder of X and $CTE_{U,Z;10}(U,Z)$ is C_X-measurable. Hence, the true outcome variables are

$$\tau_0 \;=\; E^{X=0}(Y\,|\,C_X) \;=\; E^{X=0}(Y\,|\,U,Z),$$
$$\tau_1 \;=\; E^{X=1}(Y\,|\,C_X) \;=\; E^{X=1}(Y\,|\,U,Z).$$

The values of these two true outcome variables are displayed in Table 5.1.

Obviously, Joe's total effect of the individual therapy is

$$E(Y\,|\,X=1,Z=0,U=\mathit{Joe}) \;-\; E(Y\,|\,X=0,Z=0,U=\mathit{Joe}) \;=\; 82 \;-\; 68 \;=\; 14, \tag{5.28}$$

if he does not receive group therapy ($Z=0$), whereas it is

$$E(Y\,|\,X=1,Z=1,U=\mathit{Joe}) \;-\; E(Y\,|\,X=0,Z=1,U=\mathit{Joe}) \;=\; 100 \;-\; 96 \;=\; 4, \tag{5.29}$$

if he does ($Z=1$). Similarly, Ann's total effect of the individual therapy is

$$E(Y\,|\,X=1,Z=0,U=\mathit{Ann}) \;-\; E(Y\,|\,X=0,Z=0,U=\mathit{Ann}) \;=\; 98 \;-\; 80 \;=\; 18, \tag{5.30}$$

if she does not receive group therapy ($Z=0$), whereas it is

$$E(Y\,|\,X=1,Z=1,U=\mathit{Ann}) \;-\; E(Y\,|\,X=0,Z=1,U=\mathit{Ann}) \;=\; 106 \;-\; 104 \;=\; 2, \tag{5.31}$$

if she does ($Z=1$) (see Table 5.1). Hence, all these ($U=u,Z=z$)-conditional total effects are positive, and they are the true total effects on Y [see Def. 5.2 (i)]. However, in this example, these effects *are not* the individual effects (see section 5.4.6).

5.4.4 Causal Average Total Effect

Using the true total effects obtained in Equations (5.28) to (5.31), the causal average total effect of *individual therapy* ($X=1$) compared to *no individual therapy* ($X=0$) can be computed by

$$\begin{aligned}
&E\big(\,CTE_{U,Z;10}(U,Z)\big)\\
&= \sum_u \sum_z \big(E(Y\,|\,X=1,U=u,Z=z) - E(Y\,|\,X=0,U=u,Z=z)\big)\cdot P(U=u,Z=z)\\
&= 14\cdot\frac{1}{4} + 4\cdot\frac{1}{4} + 18\cdot\frac{1}{4} + 2\cdot\frac{1}{4} \;=\; 9.5.
\end{aligned}$$

In these computations we used SN-Equation (6.15), and

$$P(U=u,Z=z) = P(Z=z\,|\,U=u)\cdot P(U=u) = 1/2\cdot 1/2 = 1/4$$

for all values (u,z) of the global potential confounder (U,Z).

5.4.5 ($U=u$)-Conditional Prima Facie Effects

In this example, the differences

$$E(Y\,|\,X=1,U=\mathit{Joe}) \;-\; E(Y\,|\,X=0,U=\mathit{Joe})$$

and

$$E(Y\,|\,X=1,U=\mathit{Ann}) \;-\; E(Y\,|\,X=0,U=\mathit{Ann})$$

are the individual prima facie effects. They are not identical to the causal individual total effects, that is, they are not the values of the causal U-conditional total effect function. In this example, these individual prima facie effects are biased, in the sense that will be explicated in chapter 6.

In order to compute the conditional expectation values of Y given treatment and unit, we use the equation

$$E(Y \mid X{=}x, U{=}u) \;=\; \sum_z E(Y \mid X{=}x, U{=}u, Z{=}z) \cdot P(Z{=}z \mid X{=}x, U{=}u), \tag{5.32}$$

which is always true if $P(X{=}x, U{=}u) > 0$ and Z is discrete with $P(X{=}x, U{=}u, Z{=}z) > 0$ for all values of Z [see SN-Box 9.2 (ii)]. Both kinds of parameters occurring on the right-hand side of this equation are displayed in Table 5.1. The conditional expectation values $E(Y \mid X{=}x, U{=}u, Z{=}z)$ are among the fundamental parameters.[2] The probabilities $P(Z{=}z \mid X{=}x, U{=}u)$ have been computed from the fundamental parameters via:

$$P(Z{=}z \mid X{=}x, U{=}u) \;=\; \frac{P(X{=}x \mid U{=}u, Z{=}z) \cdot P(Z{=}z \mid U{=}u)}{\sum_z P(X{=}x \mid U{=}u, Z{=}z) \cdot P(Z{=}z \mid U{=}u)} \tag{5.33}$$

(see Exercises 5-11 and 5-13).

Hence, using Equation (5.32), the $(X{=}x, U{=}\mathit{Joe})$-conditional expectation values for Joe are

$$\begin{aligned} E(Y \mid X{=}1, U{=}\mathit{Joe}) \;&=\; 82 \cdot \frac{3}{4} + 100 \cdot \frac{1}{4} \;=\; 86.5, \\ E(Y \mid X{=}0, U{=}\mathit{Joe}) \;&=\; 68 \cdot \frac{1}{4} + 96 \cdot \frac{3}{4} \;=\; 89, \end{aligned}$$

and his individual prima facie effect is

$$E(Y \mid X{=}1, U{=}\mathit{Joe}) \;-\; E(Y \mid X{=}0, U{=}\mathit{Joe}) \;=\; 86.5 \;-\; 89 \;=\; -2.5. \tag{5.34}$$

In this example, the individual prima facie effect of *individual therapy* compared to *no individual therapy* are negative, namely -2.5, although all $(U{=}\mathit{Joe}, Z{=}z)$-conditional effects are positive, namely 14 for $U{=}\mathit{Joe}$ and $Z{=}0$ (e. g., given *Joe and no group therapy*) and 4 for $U{=}\mathit{Joe}$ and $Z{=}1$ (e. g., given *Joe and group therapy*).

Similarly, using Equation (5.32), the $(X{=}x, U{=}\mathit{Ann})$-conditional expectation values for Ann are

$$\begin{aligned} E(Y \mid X{=}1, U{=}\mathit{Ann}) \;&=\; 98 \cdot \frac{3}{4} + 106 \cdot \frac{1}{4} \;=\; 100, \\ E(Y \mid X{=}0, U{=}\mathit{Ann}) \;&=\; 80 \cdot \frac{1}{4} + 104 \cdot \frac{3}{4} \;=\; 98, \end{aligned}$$

and her individual prima facie effect is

$$E(Y \mid X{=}1, U{=}\mathit{Ann}) \;-\; E(Y \mid X{=}0, U{=}\mathit{Ann}) \;=\; 100 \;-\; 98 \;=\; 2. \tag{5.35}$$

This prima facie effect does not have a causal interpretation. It is not identical to the causal $(U{=}\mathit{Ann})$-conditional total effect of X on Y, which will be computed in section 5.4.6.

[2] The term 'fundamental parameter' has no deeper meaning. It simply refers to the fact that these parameters can be used to compute the joint and marginal distributions of U, Z, and X, as well as the values of the conditional expectation $E(Y \mid U, Z, X)$.

5.4.6 Causal $(U=u)$-Conditional Total Effects

Because $C_X = (U, Z)$ is a global potential confounder in the example presented in Table 5.1, the causal $(U=u)$-conditional (or individual) total effects of *individual therapy* $(X=1)$ compared to *no individual therapy* $(X=0)$ can be computed for Joe by

$$\begin{aligned}
CTE_{U;10}(Joe) &= E\big(CTE_{U,Z;10}(U,Z)\,|\,U=Joe\big)\\
&= E\big(E^{X=1}(Y|U,Z) - E^{X=0}(Y|U,Z)\;\big|\;U=Joe\big)\\
&= \sum_u \sum_z \big(E(Y|X=1,U=u,Z=z) - E(Y|X=0,U=u,Z=z)\big)\cdot P(U=u,Z=z\,|\,U=Joe)\\
&= (82-68)\cdot\frac{1}{2} + (100-96)\cdot\frac{1}{2} + (98-80)\cdot 0 + (106-104)\cdot 0 \;=\; 9
\end{aligned}$$

[see Eq. (5.26) and SN-Eq. (9.19)], and for Ann by

$$\begin{aligned}
CTE_{U;10}(Ann) &= E\big(CTE_{U,Z;10}(U,Z)\,|\,U=Ann\big)\\
&= E\big(E^{X=1}(Y|U,Z) - E^{X=0}(Y|U,Z)\;\big|\;U=Ann\big)\\
&= \sum_u \sum_z \big(E(Y|X=1,U=u,Z=z) - E(Y|X=0,U=u,Z=z)\big)\cdot P(U=u,Z=z\,|\,U=Ann)\\
&= (82-68)\cdot 0 + (100-96)\cdot 0 + (98-80)\cdot\frac{1}{2} + (106-104)\cdot\frac{1}{2} \;=\; 10.
\end{aligned}$$

Hence, in this example, the two individual total effects for Joe and Ann are both positive.

Comparing the causal individual effect $CTE_{U;10}(Joe) = 9$ to the corresponding prima facie effect -2.5 [see Eq. (5.34)] shows that the individual prima facie effect $E(Y\,|\,X=1, U=Joe) - E(Y\,|\,X=0, U=Joe)$ strongly differs from its causal counterpart, and the same applies to the individual prima facie effect of Ann. This is evident if we compare her prima facie effect $E(Y\,|\,X=1, U=Ann) - E(Y\,|\,X=0, U=Ann) = 2$ to Ann's causal individual effect $CTE_{U;10}(Ann) = 10$.

According to Equation (5.22), the expectation of these causal individual effects,

$$\begin{aligned}
E\big(CTE_{U;10}(U)\big) &= E\big(E^{X=1}(Y|U,Z) - E^{X=0}(Y|U,Z)\;\big|\;U\big)\\
&= \sum_u CTE_{U;10}(u)\cdot P(U=u) \;=\; 9\cdot\frac{1}{2} \;+\; 10\cdot\frac{1}{2} \;=\; 9.5,
\end{aligned}$$

is the causal average total effect ATE_{10} of *individual therapy* compared to *no individual therapy*.

Causal individual total effects are more informative than the causal average total effect and usually more informative than causal conditional total effects given a value of a pretest or a second treatment variable as in this example. However, note again that causal individual (i. e., $(U=u)$-conditional) total effects are not necessarily the most fine-grained causal total effects. In this example, there is a second treatment variable, denoted Z, that contributes to the variation of the outcome variable Y beyond the individual level. This is exemplified comparing the causal $(U=u, Z=z)$-conditional total effects to the causal $(U=u)$-conditional total effects.

5.4.7 Causal $(Z=z)$-Conditional Total Effects

In the example presented in Table 5.1, the causal $(Z=0)$-conditional (i. e., given *no group therapy*) total effect of *individual therapy* $(X=1)$ compared to *no individual therapy* $(X=0)$ can be computed by

$$\begin{aligned}
&CTE_{Z;10}(0) \;=\; E\big(CTE_{U,Z;10}(U,Z)\,|\,Z=0\big)\\
&= E\big(E^{X=1}(Y|U,Z) - E^{X=0}(Y|U,Z)\;\big|\;Z=0\big)\\
&= \sum_u\sum_z \big(E(Y|X=1,U=u,Z=z) - E(Y|X=0,U=u,Z=z)\big)\cdot P(U=u,Z=z|Z=0)\\
&= (82-68)\cdot\frac{1}{2} + (100-96)\cdot 0 + (98-80)\cdot\frac{1}{2} + (106-104)\cdot 0 \;=\; 16
\end{aligned}$$

[see Eqs. (5.26) and SN-(9.19)]. The corresponding causal $(Z=1)$-conditional (i. e., given *group therapy*) total effect is

$$\begin{aligned}
&CTE_{Z;10}(1) \;=\; E\big(CTE_{U,Z;10}(U,Z)\,|\,Z=1\big)\\
&= E\big(E^{X=1}(Y|U,Z) - E^{X=0}(Y|U,Z)\;\big|\;Z=1\big)\\
&= \sum_u\sum_z \Big(E(Y|X=1,U=u,Z=z) - E(Y|X=0,U=u,Z=z)\Big)\cdot P(U=u,Z=z|Z=1)\\
&= (82-68)\cdot 0 + (100-96)\cdot\frac{1}{2} + (98-80)\cdot 0 + (106-104)\cdot\frac{1}{2} \;=\; 3.
\end{aligned}$$

According to Equation (5.22), taking the expectation

$$E\big(CTE_{Z;10}(Z)\big) \;=\; \sum_z E\big(CTE_{Z;10}(z)\,|\,Z=z\big)\cdot P(Z=z) \;=\; 16\cdot\frac{1}{2} \;+\; 3\cdot\frac{1}{2} \;=\; 9.5 \qquad (5.36)$$

again yields the average total effect. In this equation, we used the theorem of total probability in order to compute $P(Z=z) = \sum_u P(Z=z\,|\,U=u)\cdot P(U=u)$ (see SN-Th. 4.25), which yields $P(Z=0) = P(Z=1) = 1/2$.

5.4.8 Causal $(X=x)$-Conditional Total Effects

Consider again the example presented in Table 5.1. Because $C_X = (U,Z)$, the causal $(X=0)$-conditional total effect of *individual therapy* $(X=1)$ compared to *no individual therapy* $(X=0)$ can be computed by

$$\begin{aligned}
&CTE_{X;10}(0) \;=\; E\big(CTE_{U,Z;10}(U,Z)\,|\,X=0\big)\\
&= \sum_u\sum_z \big(E(Y|X=1,U=u,Z=z) - E(Y|X=0,U=u,Z=z)\big)\cdot P(U=u,Z=z|X=0)\\
&= (82-68)\cdot\frac{1}{8} + (100-96)\cdot\frac{3}{8} + (98-80)\cdot\frac{1}{8} + (106-104)\cdot\frac{3}{8} \;=\; 6.25
\end{aligned}$$

[see again Eqs. (5.26) and SN-(9.19)]. In contrast,

$$CTE_{X;10}(1) \;=\; E\big(CTE_{U,Z;10}(U,Z)\,|\,X=1\big)$$

$$= \sum_u \sum_z \big(E(Y \mid X=1, U=u, Z=z) - E(Y \mid X=0, U=u, Z=z)\big) \cdot P(U=u, Z=z \mid X=1)$$
$$= (82-68) \cdot \frac{3}{8} + (100-96) \cdot \frac{1}{8} + (98-80) \cdot \frac{3}{8} + (106-104) \cdot \frac{1}{8} \;=\; 12.75$$

yields the $(X=1)$-conditional total effect of *individual therapy* $(X=1)$ compared to *no individual therapy* $(X=0)$. In these equations, we used

$$P(U=u, Z=z \mid X=x) \;=\; \frac{P(X=x \mid U=u, Z=z) \cdot P(U=u, Z=z)}{P(X=x)}. \tag{5.37}$$

According to Equation (5.22), taking the expectation

$$E\big[\, CTE_{X;10}(X) \big] \;=\; \sum_x CTE_{X;10}(x) \cdot P(X=x) \;=\; 6.25 \cdot \frac{1}{2} \;+\; 12.75 \cdot \frac{1}{2} \;=\; 9.5$$

yields the causal average total effect. In this equation, we again used the theorem of total probability, that is, $P(X=x) = \sum_u \sum_z P(X=x \mid U=u, Z=z) \cdot P(U=u, Z=z)$ (see SN-Th. 4.25), which yields $P(X=0) = P(X=1) = 1/2$.

According to these results, the assignment regime as described by the individual treatment probabilities in Table 5.1 seems to be reasonable, because the causal $(X=1)$-conditional total effect $CTE_{X;10}(1) = 12.75$ [of treatment $(X=1)$ compared to control $(X=0)$] is *greater* than the corresponding causal $(X=0)$-conditional total effect $CTE_{X;10}(0) = 6.25$. If, in contrast, $CTE_{X;10}(1) < CTE_{X;10}(0)$ would be true, then it might be worthwhile to change the assignment regime for the treatment represented by $X=1$ assigning those persons to treatment 1 that are now assigned to treatment 0.

5.4.9 $(X=x^*, V=v)$-Conditional Total Effects

If X represents a treatment variable and V is a mapping of the person variable U, this allows us, for instance, to ask for the causal conditional effects of a treatment x compared to treatment x' given treatment x^* *in specific subpopulations* represented by the values v of V. The corresponding question can also be raised if V is a fallible covariate of X. In these cases we can consider

$$CTE_{X,V;xx'}(x^*, v), \tag{5.38}$$

the *causal* $(X=x^*, V=v)$*-conditional total effect* of x compared to x', provided that we can assume $P(X=x^*, V=v) > 0$.

Suppose that X represents a treatment variable. If there are two treatment conditions, say control $(X=0)$ and treatment $(X=1)$, then, according to Equation (5.38), we may consider $CTE_{10}(1, v)$, the *causal conditional total effect given treatment and* $V=v$, as well as $CTE_{X,V;10}(0, v)$, the *causal conditional total effect given control and* $V=v$. If, for example, V is the potential confounder *sex*, then $CTE_{X,V;10}(1, m)$ is the causal $(X=1, V=m)$-conditional total treatment effect (in the *male* subpopulation), whereas $CTE_{X,V;10}(1, f)$, is the causal $(X=1, V=f)$-conditional total treatment effect (in the *female* subpopulation). As mentioned before, in more precise terms conditioning on a subpopulation means that we condition on the event $\{V=v\} = \{\omega \in \Omega\colon V(\omega) = v\}$ that the person drawn is an element of the subpopulation represented by the value v of V.

Example 5.37 (Joe and Ann With Bias at the Individual Level) In the example presented in Table 5.1, $C_X = (U,Z)$ is a global potential confounder of X and $CTE_{U,Z;10}(U,Z)$ is C_X-measurable. Hence, there is a measurable function $CTE_{U,Z;10}\colon(\Omega_U,\Omega'_Z)\to(\mathbb{R},\mathscr{B})$ such that $CTE_{U,Z;10}(U,Z)$ is the composition of (U,Z) and $CTE_{U,Z;10}$ (see SN-Lemma 2.52). Because (U,Z) is a global potential confounder,

$$CTE_{U,Z;10}(U,Z) \;=\; \delta_{10} \;=\; \tau_1 \,-\, \tau_0,$$

that is, $CTE_{U,Z;10}(U,Z)$ is a version of the true effect variable. In this example, δ_{10} is uniquely defined, because the probabilities $P(X{=}x,U{=}u,Z{=}z)$ are positive for all triples of values of X, U, and Z. The values of $CTE_{U,Z;10}(U,Z)$ are:

$$\begin{aligned}
&CTE_{U,Z;10}(U,Z)(\omega)\\
&= CTE_{U,Z;10}(u,z)\\
&= E^{X=1}(Y|U{=}u,Z{=}z) - E^{X=0}(Y|U{=}u,Z{=}z), \quad \forall\omega \text{ with } (U,Z)(\omega) = (U{=}u,Z{=}z)
\end{aligned}$$

(see Rem. 4.18). According to Table 5.1, these values are 14, 4, 18, and 2. Hence, the causal average total effect of *individual therapy* $(X=1)$ compared to *no individual therapy* $(X=0)$ can be computed by

$$\begin{aligned}
E\big(CTE_{U,Z;10}(U,Z)\big) \;&=\; \sum_u\sum_z \big(E^{X=1}(Y|U{=}u,Z{=}z) - E^{X=0}(Y|U{=}u,Z{=}z)\big)\cdot P(U{=}u,Z{=}z)\\
&=\; 14\cdot\frac{1}{4} + 4\cdot\frac{1}{4} + 18\cdot\frac{1}{4} + 2\cdot\frac{1}{4} \;=\; 9.5,
\end{aligned}$$

using SN-Equation (6.15), and

$$P(U{=}u,Z{=}z) = P(Z{=}z|U{=}u)\cdot P(U{=}u) = 1/2\cdot 1/2 = 1/4,$$

for all values (u,z) of $C_X = (U,Z)$. ◁

5.4.10 Causal $(X{=}x,Z{=}z)$-Conditional Total Effects

Because, in the example presented in Table 5.1, (U,Z) is a global potential confounder, the causal conditional total effect $CTE_{X,Z;10}(x,z)$ of *individual therapy* $(X=1)$ compared to *no individual therapy* $(X=0)$ *given* $X{=}x$ and $Z{=}z$ can be computed by

$$\begin{aligned}
&CTE_{X,Z;10}(x,z) \;=\; E\big(CTE_{U,Z;10}(U,Z)\,\big|\,X{=}x,Z{=}z\big)\\
&= \sum_u\sum_{z'}\big[E^{X=1}(Y|U{=}u,Z{=}z') - E^{X=0}(Y|U{=}u,Z{=}z')\big]\cdot P(U{=}u,Z{=}z'|X{=}x,Z{=}z).
\end{aligned}$$

If $z'{\neq}z$, then the conditional probabilities $P(U{=}u,Z{=}z'|X{=}x,Z{=}z)$ are zero. Otherwise they can be computed via

$$P(U{=}u,Z{=}z|X{=}x,Z{=}z) \;=\; \frac{P(X{=}x|U{=}u,Z{=}z)\cdot P(U{=}u,Z{=}z)}{P(X{=}x|Z{=}z)\cdot P(Z{=}z)}, \tag{5.39}$$

where

$$P(X{=}x|Z{=}z) \;=\; \sum_u P(X{=}x|U{=}u,Z{=}z)\cdot P(U{=}u|Z{=}z), \tag{5.40}$$

with $P(U{=}u\,|\,Z{=}z) = P(Z{=}z\,|\,U{=}u)\cdot P(U{=}u)/P(Z{=}z)$. In this example, $P(U{=}u|Z{=}z) = 1/2$, for all values of U and Z. Hence, Equation (5.40) yields $P(X=0\,|\,Z=0) = 1/4\cdot 1/2 + 1/4\cdot 1/2 = 1/4$, and using Equation (5.39) we receive:

$$P(U{=}u, Z{=}0\,|\,X{=}0, Z{=}0) \;=\; \frac{1/4\cdot 1/4}{1/4\cdot 1/2} \;=\; \frac{1}{2},$$

for $u{=}$*Joe* and for $u{=}$*Ann*. Hence, the equation for $E\big(CTE_{U,Z;10}(U,Z)\,\big|\,X{=}x, Z{=}z\big)$ yields

$$\begin{aligned}&E\big(CTE_{U,Z;10}(U,Z)\,\big|\,X=0, Z=0\big)\\ &= (82-68)\cdot\frac{1}{2} + (100-96)\cdot 0 + (98-80)\cdot\frac{1}{2} + (106-104)\cdot 0 \;=\; 16.\end{aligned}$$

For $X=1$ and $Z=0$, we receive

$$\begin{aligned}&E\big(CTE_{U,Z;10}(U,Z)\,\big|\,X=1, Z=0\big)\\ &= (82-68)\cdot\frac{1}{2} + (100-96)\cdot 0 + (98-80)\cdot\frac{1}{2} + (106-104)\cdot 0 \;=\; 16,\end{aligned}$$

for $X=0$ and $Z=1$, we receive

$$\begin{aligned}&E\big(CTE_{U,Z;10}(U,Z)\,\big|\,X=0, Z=1\big)\\ &= (82-68)\cdot 0 + (100-96)\cdot\frac{1}{2} + (98-80)\cdot 0 + (106-104)\cdot\frac{1}{2} \;=\; 3,\end{aligned}$$

and for $X=1$ and $Z=1$:

$$\begin{aligned}&E\big(CTE_{U,Z;10}(U,Z)\,\big|\,X=1, Z=1\big)\\ &= (82-68)\cdot 0 + (100-96)\cdot\frac{1}{2} + (98-80)\cdot 0 + (106-104)\cdot\frac{1}{2} \;=\; 3.\end{aligned}$$

Hence, in this special example, the conditional total effects $E\big(CTE_{U,Z;10}(U,Z)\,\big|\,Z{=}z\big)$ and $E\big(CTE_{U,Z;10}(U,Z)\,\big|\,X{=}x, Z{=}z\big)$ are identical.

According to Equation (5.22), taking the expectation

$$\begin{aligned}&E\Big(E\big(CTE_{U,Z;10}(U,Z)\,\big|\,X,Z\big)\Big)\\ &= \sum_x\sum_z E\big(CTE_{U,Z;10}(U,Z)\,\big|\,X{=}x, Z{=}z\big)\cdot P(X{=}x, Z{=}z)\\ &= 16\cdot\frac{1}{4} \;+\; 3\cdot\frac{1}{4} \;+\; 16\cdot\frac{1}{4} \;+\; 3\cdot\frac{1}{4} \;=\; 9.5,\end{aligned}$$

yields the average total effect. In this equation, we again used the theorem of total probability, that is, $P(X{=}x, Z{=}z) = \sum_u P(X{=}x, Z{=}z\,|\,U{=}u)\cdot P(U{=}u)$ (see SN-Th. 4.25), which yields $P(X{=}x, Z{=}z) = 1/4$, for all values of X and Z.

5.5 Summary and Conclusions

In this chapter we defined several kinds of causal total effects of x compared to x' based on the true total effect variable

Box 5.1 Glossary of New Concepts

Let the Assumptions 5.1 hold and let V be a random variable on $(\Omega, \mathscr{A}, P)$ with value space $(\Omega'_V, \mathscr{A}'_V)$.

$ATE_{xx'}$ *Causal average total effect comparing x to x'.* If τ_x and $\tau_{x'}$ are P-unique, then

$$ATE_{xx'} := E(\tau_x - \tau_{x'}).$$

$CTE_{V;xx'}(V)$ *Causal V-conditional total effect variable comparing x to x'.* If τ_x and $\tau_{x'}$ are P-unique, then

$$CTE_{V;xx'}(V) := E(\tau_x - \tau_{x'} \mid V).$$

$CTE_{V;xx'}(v)$ *Causal $(V=v)$-conditional total effect comparing x to x'.* If τ_x and $\tau_{x'}$ are $P^{V=v}$-unique, then

$$CTE_{V;xx'}(v) := E(\tau_x - \tau_{x'} \mid V=v).$$

$CTE_{U;xx'}(u)$ *Causal individual total effect comparing x to x' for unit u.* It is a special case of $CTE_{V;xx'}(v)$ for $U=V$.

$CTE_{X;xx'}(x^*)$ *Causal $(X=x^*)$-conditional total effect comparing x to x'.* It is a special case of $CTE_{V;xx'}(v)$ for $X=V$.

$$\delta_{xx'} = \tau_x - \tau_{x'}.$$

The definitions of a *causal average total effect* $ATE_{xx'}$ and of a *causal V-conditional total effect variable* $CTE_{V;xx'}(V)$ (see Box 5.1) are based on the assumption that the two true outcome variables τ_x and $\tau_{x'}$ are P-unique. Defining the *causal $(V=v)$-conditional total effect* $CTE_{V;xx'}(v)$ we only assume that τ_x and $\tau_{x'}$ are $P^{V=v}$-unique. The term 'total' is used in order to distinguish these effects from direct and indirect effects that are not considered in this volume.

While C_X is a global potential confounder of X on which we condition in order to control for all potential confounders of X, the variable V may be used to reaggregate the $(C_X=c)$-conditional total effects in order to consider less fine-grained causal conditional total effects. Examples of V are the observational-unit variable U, a pre-treatment variable Z, and the treatment variable X.

Causal Average Total Effect

Often we have to content ourselves with the causal average total effect or causal conditional total effects. Note, however, that a causal average total effect may not apply to any unit at all. There might very well be cases in which half of the units have positive causal individual total effects and the other half negative ones. The causal average total effect can then be zero. This is not a paradox but the nature of an average. Also remember that a causal average total effect is already much more informative for causal inference than an ordinary true mean difference $E(Y \mid X=1) - E(Y \mid X=0)$, the prima facie effect considered

in chapter 1. These prima facie effects have no causal interpretation at all, unless they are identical to the causal average total effect (see Def. 6.2).

Main Effects Versus Conditional Effects

Conceptually, the *causal average total effect* is what is tested in a two-group t-test and as the *main effect* of the 'treatment factor' in orthogonal analysis of variance, provided that the data are sampled in a perfect randomized experiment. Note that the causal average total effect is uniquely defined even if there are inter-individual differences in the individual total effects, and even if there is interaction between X and a potential confounder V in the sense that the effect of X depends on the values of V. The causal average total effect is uniquely defined even if X and V are correlated or stochastically dependent. If V is a qualitative covariate, it is considered a second 'factor' in analysis of variance, and the causal average total effects are what we test as the *main effect* of the 'treatment factor'. However, only in the randomized experiment we can be sure that with the main effects in analysis of variance we test the causal average total effects.

Of course, the causal conditional effects given the values of a covariate are usually more informative than their average, that is, than the causal average total effect; but sometimes averaging is useful in order to avoid information overload, and sometimes we may be able to estimate precisely enough only the causal average effect, but not the causal conditional effects, for example, because of small sample sizes. If V is a nonnumerical random variable with a few number of values, it is often considered a second factor in analysis of variance. In this case, the $(V=v)$-conditional total effects are often called the 'simple main effects' (see, e. g., Woodward & Bonett, 1991).

Pre-Facto Versus Counterfactual Perspective

Note that our definitions of the various kinds of total effects solely use concepts of probability theory. No concepts had to be borrowed from philosophy or any other science — although the basic idea goes back at least to Mill (1843/1865). We did not take a counterfactual but a *pre-facto perspective*, which is the perspective taken in *every* application of probability theory. Causal total effects are parameters, just in the same way as the probability of flipping 'heads' is a parameter about which we can talk *before* the coin is flipped and even if the coin is *never* flipped. Therefore, it is also meaningful to talk about the causal individual effects of a treatment for a unit which is actually never treated, and about the *causal average effect of a treatment* including also those that are not treated. It is even meaningful to talk about the *causal conditional effect of a treatment given control* [see Eq. (5.19) and Rem. 5.26].

Outlook

Note that all concepts introduced in this chapter such as the *causal average total effects*, *causal conditional total effects*, *causal individual total effects*, and so on, are of a purely theoretical nature. They explicate what exactly we are looking for when we ask for the causal effects, for example, of a treatment variable or of another discrete cause. This also applies to the examples treated in chapters 4 and 5. For example, Table 4.1 and Table 5.1 do not show data that might be obtained in a data sample. Instead, they contain the theoreti-

cal parameters we would like to estimate from sample data, and this includes the various causal effects. In terms of the metaphor presented in the preface, these causal effects are the *size* of the invisible man. In contrast, in chapter 1, we only dealt with the prima facie effects: (a) ordinary conditional expectation values $E(Y|X{=}x)$ of an outcome variable Y given treatment x, (b) conditional expectation values $E(Y|X{=}x,Z{=}z)$ of the outcome variable given treatment x and value z of a covariate Z, (c) differences between these (conditional) expectation values, the (conditional) *prima facie effects*, and (d) averages over these conditional prima facie effects. The conditional expectation values $E(Y|X{=}x)$ and $E(Y|X{=}x,Z{=}z)$ are easily estimated under the usual assumptions made for a sample, such as the assumption of independent, identically distributed observations. However, they are only like the length of the invisible man's shadow; depending on the angle of the sun, they can be seriously biased if mistaken for the size of the invisible man itself. Consequently, the next chapter is devoted to unbiasedness of the conditional expectation values such as $E(Y|X{=}x)$ and $E(Y|X{=}x,Z{=}z)$ and their differences between different values x and x' of X.

5.6 Proofs

Proof of Theorem 5.31

$$\begin{aligned} E\big(CTE_{V;xx'}(V)\big) &= E\big(E(\tau_x-\tau_{x'}|V)\big) && \text{[(5.16)]}\\ &= E(\tau_x-\tau_{x'}). && \text{[SN-Box 10.2 (iv)]} \end{aligned}$$

Proof of Theorem 5.34

(i)

$$\begin{aligned} E\big(CTE_{V;xx'}(V)\,\big|\,W\big) &\underset{P}{=} E\big(E(\tau_x-\tau_{x'}|V)\,\big|\,W\big) && \text{[(5.16)]}\\ &\underset{P}{=} E(\tau_x-\tau_{x'}|W). && \text{[SN-Box 10.2 (v)]} \end{aligned}$$

(ii)

$$\begin{aligned} E\big(CTE_{V;xx'}(V)\,\big|\,W{=}w\big) &= E\big(E(\tau_x-\tau_{x'}|V)\,\big|\,W{=}w\big) && \text{[(5.16)]}\\ &= E(\tau_x-\tau_{x'}|W{=}w). && \text{[SN-(10.37)]} \end{aligned}$$

5.7 Exercises

▷ **Exercise 5-1** Suppose X is a treatment variable and Y an outcome variable. Why are the conditional expectation values $E(Y|X{=}x)$ and their differences $E(Y|X{=}x)-E(Y|X{=}x')$, the prima facie effects, often useless in the evaluation of treatment effects?

▷ **Exercise 5-2** Suppose X is a treatment variable and Y an outcome variable. If the conditional expectation values $E(Y|X{=}x)$ and their differences $E(Y|X{=}x)-E(Y|X{=}x')$ do not represent the treatment effects we are interested in, then what *are* the treatment effects we would like to study?

▷ **Exercise 5-3** What is the difference between the causal average total effect $ATE_{xx'}$ and the prima facie effect $PFE_{xx'}$?

▷ **Exercise 5-4** What is the *causal conditional total effect* $CTE_{V;xx'}(v)$ on Y comparing x to x' *given the value* v *of a random variable* V?

▷ **Exercise 5-5** What is the *causal conditional total effect* $CTE_{X;xx'}(x^*)$ on outcome variable Y comparing x to x' *given treatment* x^*?

▷ **Exercise 5-6** Show that $E(Y^2) < \infty$ implies $\exists\, V_x \in \mathscr{E}^{X=x}(Y \mid C_X)$ such that $E(V_x) < \infty$.

▷ **Exercise 5-7** What is the *causal conditional total effect* $CTE_{X,V;xx'}(x^*, v)$ on Y comparing x to x' *given treatment* x^* *and value* v *of the covariate* V?

▷ **Exercise 5-8** Use SN-Theorem 4.25 to compute the probability $P(X = 1)$ for the example displayed in Table 5.1.

▷ **Exercise 5-9** What are the elements of the σ-algebra generated by Z in Example 5.4?

▷ **Exercise 5-10** Show that proposition (5.20) follows from independence of X and C_X.

▷ **Exercise 5-11** Compute the probability $P(U{=}Ann, Z = 0 \mid X = 1)$ in the example of Table 5.1.

▷ **Exercise 5-12** Use the *Aggregation*[0,1]*Xplorer* (see 'Tools' at www.causal-effects.de), choose the example 'Treatment effect, X and Z are dependent [Inversion of SME and MEM]', and click on 'Compute aggregated effects and visualizations'.

▷ **Exercise 5-13** Compute the probabilities $P(Z{=}z \mid X{=}x, U{=}u)$ in Table 5.1.

▷ **Exercise 5-14** Compute the causal average total effect ATE_{10} for the random experiment presented in Table 5.1.

▷ **Exercise 5-15** Compute the causal conditional total effect $CTE_{Z;10}(0)$ given *no group therapy* for the random experiment presented in Table 5.1.

▷ **Exercise 5-16** Let Z represent *sex* with values m (males) and f (females). Furthermore, suppose $CTE_{Z;20}(m) = 11$, $CTE_{Z;20}(f) = 5$, $P(Z{=}m) = 1/3$, and $P(Z{=}f) = 2/3$. Which is the causal average total effect ATE_{10}?

Solutions

▷ **Solution 5-1** Certain other variables, the potential confounders, may determine both the probability of being treated *and* the $(X{=}x)$-conditional expectation values of the outcome variable. This implies that the conditional expectation values $E(Y \mid X{=}x)$ and their differences $E(Y \mid X{=}x) - E(Y \mid X{=}x')$ are not identical to the treatment effects to be studied. An example of such a potential confounder is *severity of the symptoms*. If there is self-selection or if there is systematic selection to treatment by experts that is also determined by the severity of the symptoms, then the variable *severity of the symptoms* will both affect the treatment probability and the conditional expectation values of the outcome variable (e. g., *severity of the symptoms after treatment*). Simpson's paradox presented in chapter 1 is another example.

▷ **Solution 5-2** The basic idea is to consider the true total effect variable, that is, the difference $\tau_x - \tau_{x'} = E^{X=x}(Y|C_X) - E^{X=x'}(Y|C_X)$, where we condition on a global potential confounder C_X. This means controlling for *all* potential confounders and then taking the expectation over the distribution of C_X, which yields $E(\tau_x - \tau_{x'})$, the causal average total effect comparing x to x'. Note that this concept presumes that τ_x and $\tau_{x'}$ are P-unique. This assumption implies that the expectation $E(\tau_x - \tau_{x'})$ is identical for all versions of τ_x and $\tau_{x'}$ (see Defs. 3.79 and 4.16).

▷ **Solution 5-3** The causal average total effect $ATE_{xx'}$ comparing treatment x to treatment x' has been defined by Equation (5.7) (see also the solution to Exercise 5-2). It is these causal average total effects that might be of interest in the empirical sciences if our goal is to evaluate the treatment conditions x and x' using the outcome variable Y. In contrast, the prima facie effects $PFE_{xx'}$ comparing x to x' are usually not of interest for the evaluation of such treatment effects because they can be biased. Both terms differ from each other because $PFE_{xx'} = E(Y|X=x) - E(Y|X=x')$ is not neccesarily identical to $ATE_{xx'}$. Note, however, that *there are* conditions under which $PFE_{xx'} = ATE_{xx'}$. Such conditions, which are called *causality conditions*, are studied in some detail in the chapters to come.

▷ **Solution 5-4** The causal conditional total effect (on the outcome variable Y) comparing x to x' given the value v of a random variable V is the $(V=v)$-conditional expectation value of the true total effect variable $\delta_{xx'} = \tau_x - \tau_{x'}$, that is,

$$CTE_{V;xx'}(v) \;=\; E(\delta_{xx'}|V=v).$$

If v represents a subpopulation such as males, then $CTE_{V;xx'}(v)$ is the average total effect in this subpopulation. It is presumed that $P(V=v) > 0$ and that τ_x and $\tau_{x'}$ are $P^{V=v}$-unique.

▷ **Solution 5-5** The causal conditional total effect (on outcome variable Y) comparing x to x' given treatment x^* is the $(X=x^*)$-conditional expectation value of $\delta_{xx'} = \tau_x - \tau_{x'}$, that is,

$$CTE_{X;xx'}(x^*) \;=\; E(\delta_{xx'}|X=x^*),$$

where we presume that $P(X=x^*) > 0$ and that τ_x and $\tau_{x'}$ are $P^{X=x^*}$-unique. If X represents a treatment variable, $x^*=x$ and $X=0$ represents a control group, then $CTE_{X;x0}(x)$ is the causal conditional total effect comparing treatment x to control *given treatment x*. If $x'=x^*$ and $X=0$ represents a control group, then $CTE_{X;x0}(0)$ is the causal conditional total effect comparing treatment x to control *given control*. Although this sounds paradoxical, the term $CTE_{X;x0}(0)$ is meaningful and well-defined, because it refers to a random experiment to be conducted *in the future*. This means that they are well-defined, even if the experiment is not yet conducted, or is never conducted (see section 5.3.2 for more details).

▷ **Solution 5-6**

$$\begin{aligned}
& E(Y^2) < \infty \\
\Rightarrow\quad & E(Y) < \infty && \text{[SN-Rem. 6.25 (iii)]} \\
\Rightarrow\quad & E^{X=x}(Y) < \infty && \text{[SN-Th. 9.4 (ii), SN-(9.5), SN-(9.6)]} \\
\Rightarrow\quad & E^{X=x}\big(E^{X=x}(Y|C_X)\big) = E^{X=x}(Y) < \infty. && \text{[SN-Box 10.2 (iv)]}
\end{aligned}$$

▷ **Solution 5-7** The causal conditional total effect $CTE_{X,V;xx'}(x^*,v)$ (on the outcome variable Y) comparing x to x' given treatment x^* and value v of the covariate V is the $(X=x^*, V=v)$-conditional expectation value of the true-effect variable, that is,

$$CTE_{X,V;xx'}(x^*,v) \;=\; E(\delta_{xx'}|X=x^*, V=v),$$

where we presume that $P\big((X,V)=(x^*,v)\big) > 0$ and that τ_x and $\tau_{x'}$ are $P^{X=x^*,V=v}$-unique. If X represents a treatment variable, $x^*=x$, the value 0 of X represents a control group, and m is the value for

males of the covariate $V = \textit{sex}$, then $CTE_{X,V;x0}(x,m)$ is the causal conditional total effect comparing treatment x to control given treatment x and a person is sampled from the male subpopulation. If $x^* = 0$, then $CTE_{X,V;x0}(0,m)$ is the causal conditional total effect comparing treatment x to control given control and a person is sampled from the male subpopulation.

▷ **Solution 5-8** Note that the four pairs (u,z) of values of U and Z are disjoint and all these pairs of values have positive probabilities. Hence we can apply the theorem of total probability:

$$\begin{aligned} P(X=1) &= \sum_u \sum_z P(X=1 \,|\, U=u, Z=z) \cdot P(U=u, Z=z) \\ &= \left(\frac{3}{4} + \frac{1}{4} + \frac{3}{4} + \frac{1}{4}\right) \cdot \frac{1}{4} = \frac{1}{2}. \end{aligned}$$

Note that $P(X=1) = E(1_{X=1}) = E[E(1_{X=1} \,|\, U,Z)]$ [see SN-Box 10.2 (iv)], and that the probabilities $P(X=1 \,|\, U=u, Z=z)$ are the values of the conditional expectation $E(1_{X=1} \,|\, U,Z)$. Then using SN-Equation (6.15) yields the same formula. This second way makes clear that the unconditional probability $P(X=1)$ is the expectation of the conditional probability $P(X=1 \,|\, U,Z)$.

▷ **Solution 5-9** The set of possible outcomes of the random experiment is

$$\Omega := \Omega_U \times \Omega_Z \times \Omega_X \times \Omega_Y,$$

where $\Omega_Z := \{\textit{group therapy}, \textit{no group therapy}\}$. Hence, the σ-algebra generated by Z and $\mathscr{A}'_Z = \big\{\{0\},\{1\},\{0,1\},\emptyset\big\}$ is

$$\sigma(Z) = \big\{\Omega,\ \emptyset,\ \Omega_U \times \{\textit{group therapy}\} \times \Omega_X \times \Omega_Y,\ \Omega_U \times \{\textit{no group therapy}\} \times \Omega_X \times \Omega_Y\big\}.$$

Hence, this set has four elements. Aside from Ω and the empty set Ø, this σ-algebra contains the event

$$\Omega_U \times \{\textit{group therapy}\} \times \Omega_X \times \Omega_Y$$

that the person drawn receives group therapy, and the event

$$\Omega_U \times \{\textit{no group therapy}\} \times \Omega_X \times \Omega_Y$$

that the person drawn does *not* receive group therapy.

▷ **Solution 5-10** Independence of X and C_X implies that also X and $E^{X=x}(Y|C_X)$ are independent, because $E^{X=x}(Y|C_X)$ is C_X-measurable [see SN-Box 16.3 (vi)]. Hence,

$$\begin{aligned} E(\tau_x \,|\, X) &\underset{P}{=} E\big(E^{X=x}(Y|C_X) \,|\, X\big) && \big[\tau_x \underset{P}{=} E^{X=x}(Y|C_X)\big] \\ &\underset{P}{=} E(\tau_x). && \text{[SN-Box 10.2 (vi)]} \end{aligned}$$

▷ **Solution 5-11** We have to use the equation

$$P(U=u, Z=z \,|\, X=x) = \frac{P(X=x, U=u, Z=z)}{P(X=x)} = \frac{P(X=x \,|\, U=u, Z=z) \cdot P(U=u, Z=z)}{P(X=x)}.$$

For $U=\textit{Ann}$, $Z=0$, and $X=1$ this equation yields

$$\begin{aligned} P(U=\textit{Ann}, Z=0 \,|\, X=1) &= \frac{P(X=1 \,|\, U=\textit{Ann}, Z=0) \cdot P(U=\textit{Ann}, Z=0)}{P(X=1)} \\ &= \frac{3/4 \cdot 1/4}{1/2} = \frac{3}{8}. \end{aligned}$$

▷ **Solution 5-12** No solution provided. See what happens with the various techniques of aggregating conditional effects, for example, aggregating the log odds ratios. Play with other parameter constellations.

▷ **Solution 5-13** The complementary probability to $P(X=1\,|\,U=u,Z=0)=3/4$ is $P(X=0\,|\,U=u,Z=0)=1/4$ for both units u. Similarly the complementary probability to $P(X=1\,|\,U=u,Z=1)=1/4$ is $P(X=0\,|\,U=u,Z=1)=3/4$. Now we can use the equation

$$P(Z=z\,|\,X=x,U=u) \;=\; \frac{P(X=x\,|\,U=u,Z=z)\cdot P(Z=z\,|\,U=u)}{\sum\limits_{z} P(X=x\,|\,U=u,Z=z)\cdot P(Z=z\,|\,U=u)}.$$

For example, the probability of not receiving group therapy ($Z=0$), if Ann is drawn ($U=Ann$) and does not receive individual therapy ($X=0$), is

$$P(Z=0\,|\,X=0,U=Ann) \;=\; \frac{P(X=0\,|\,U=Ann,Z=0)\cdot P(Z=0\,|\,U=Ann)}{\sum\limits_{z} P(X=0\,|\,U=Ann,Z=z)\cdot P(Z=z\,|\,U=Ann)},$$

where $P(X=0\,|\,U=Ann,Z=0)=1/4,\;\; P(Z=0\,|\,U=Ann)=1/2$, and

$$\begin{aligned}
&\sum_{z} P(X=0\,|\,U=Ann,Z=z)\cdot P(Z=0\,|\,U=Ann)\\
&= P(X=0\,|\,U=Ann,Z=0)\cdot P(Z=0\,|\,U=Ann) \;+\; P(X=0\,|\,U=Ann,Z=1)\cdot P(Z=1\,|\,U=Ann)\\
&= \frac{1}{4}\cdot\frac{1}{2} + \frac{3}{4}\cdot\frac{1}{2} \;=\; \frac{1}{2}.
\end{aligned}$$

Inserting this result yields the probability $P(Z=0\,|\,X=0,U=Ann)=1/4$. Using the same procedure for all values of U, X, and Z leads to the probabilities displayed in Table 5.1.

▷ **Solution 5-14**

$$\begin{aligned}
ATE_{10} \;&=\; \sum_{u}\sum_{z}\big(E(Y\,|\,X=1,U=u,Z=z) - E(Y\,|\,X=0,U=u,Z=z)\big)\cdot P(U=u,Z=z)\\
&=\; \big((82-68)+(100-96)+(98-80)+(106-104)\big)\cdot\frac{1}{4} \;=\; 9.5.
\end{aligned}$$

▷ **Solution 5-15**

$$\begin{aligned}
CTE_{Z;10}(0) \;&=\; \sum_{u}\sum_{z}\big(E(Y\,|\,X=1,U=u,Z=z) - E(Y\,|\,X=0,U=u,Z=z)\big)\cdot P(U=u,Z=z\,|\,Z=0)\\
&=\; (82-68)\cdot\frac{1}{2} + (100-96)\cdot 0 + (98-80)\cdot\frac{1}{2} + (106-104)\cdot 0 \;=\; 16.
\end{aligned}$$

▷ **Solution 5-16** Using Equation (5.22), we can compute the causal average total effect as follows:

$$ATE_{20} \;=\; CTE_{Z;20}(m)\cdot\frac{1}{3} + CTE_{Z;20}(f)\cdot\frac{2}{3} \;=\; 11\cdot\frac{1}{3} + 5\cdot\frac{2}{3} \;=\; 7.$$

Chapter 6
Unbiasedness and Identification of Causal Effects

In chapter 4, we introduced the concepts of a *causality space*, a *potential confounder*, a *covariate*, a *global potential confounder*, and a *true outcome variable*. In chapter 5, we turned to the concepts of a *true total effect variable*, a *causal average total effect*, a *causal conditional total effect variable*, and a *causal conditional total effect*. All these parameters and variables are of a theoretical nature. It is not evident how they can be computed (identified) from parameters of the joint distribution of observable random variables such as X, Y, and a (possibly multivariate) covariate Z, that is, from those parameters that can be estimated in a data sample.

In this chapter we introduce and study *unbiasedness* of various conditional expectation values, conditional expectations, prima facie effects, and prima facie effect functions. In particular, we study how these terms can be used to identify the corresponding causal effects and effect functions by estimable parameters and functions. Hence, this chapter provides the link between causal effects and causal effect functions on one side and parameters and functions that can be estimated on the other side. The *unbiasedness conditions* are the first and logically weakest kind of causality conditions, which, together with the structural components listed in a causality space, distinguish *causal* stochastic dependencies from *ordinary* stochastic dependencies.

We start defining *unbiasedness* of the conditional expectation values $E(Y|X{=}x)$, unbiasedness of the conditional expectation $E(Y|X)$, and unbiasedness of the corresponding prima facie effects, that is, of the differences $E(Y|X{=}x) - E(Y|X{=}x')$. We also treat a first way to identify the causal average total effect $ATE_{xx'}$. Then we turn to unbiasedness of a conditional expectation value $E(Y|X{=}x, Z{=}z)$, unbiasedness of the conditional expectations $E^{X=x}(Y|Z)$ and $E(Y|X,Z)$, as well as the prima facie effect variables $E^{X=x}(Y|Z) - E^{X=x'}(Y|Z)$. We also show how to identify the causal average total effect $ATE_{xx'}$ as well as a causal conditional total effect function $CTE_{V;xx'}$ and a causal conditional total effect $CTE_{V;xx'}(v)$ given a value v of a random variable V. Next, we illustrate these concepts by some numerical examples. Finally, we show that unbiasedness can be accidental, presenting an example in which the conditional expectation values $E(Y|X{=}x)$ are unbiased, whereas the conditional expectation values $E(Y|X{=}x, Z{=}z)$ are not.

In this chapter we often will refer to the following notation and assumptions.

Notation and Assumptions 6.1
Let $\big((\Omega, \mathscr{A}, P), (\mathscr{F}_t, t \in T), X, Y\big)$ be a causality space, let X be discrete, let $P(X{=}x) > 0$ for all values in the image $X(\Omega) = \{0, 1, \ldots, J\}$, and let Y be real-valued with $E(Y^2) < \infty$. Furthermore, let C_X be a global potential confounder of X and let $\tau_x = E^{X=x}(Y|C_X)$ and $\tau_{x'} = E^{X=x'}(Y|C_X)$ denote true outcome variables of Y given $x, x' \in X(\Omega)$, respectively.

6.1 Unbiasedness of $E(Y|X)$ and its Values $E(Y|X=x)$

In the section 5.2 we defined the average total effect by $ATE_{xx'} = E(\delta_{xx'})$. Inserting the definition of the true total effect variable $\delta_{xx'} = \tau_x - \tau_{x'}$ and the definition of true outcome variables yields

$$\begin{aligned} ATE_{xx'} &= E(\delta_{xx'}) = E(\tau_x - \tau_{x'}) \\ &= E(\tau_x) - E(\tau_{x'}) = E\big(E^{X=x}(Y|C_X)\big) - E\big(E^{X=x'}(Y|C_X)\big). \end{aligned} \tag{6.1}$$

According to Theorem 3.84 (e) the expectation $E\big(E^{X=x}(Y|C_X)\big)$ appearing on the right-hand side of this equation is a uniquely defined number if $\tau_x = E^{X=x}(Y|C_X)$ is P-unique. This motivates the following definition.

Definition 6.2 (Unbiasedness of $E(Y|X=x)$ and $E(Y|X)$)
Let the Assumptions 6.1 hold.

(i) We call the conditional expectation value $E(Y|X=x)$ unbiased if τ_x is P-unique and

$$E(Y|X=x) = E(\tau_x). \tag{6.2}$$

(ii) We call the conditional expectation $E(Y|X)$ unbiased if, for all $x \in X(\Omega)$, the conditional expectation values $E(Y|X=x)$ are unbiased.

Remark 6.3 (An Equivalent Condition of P-Uniqueness) According to Theorem 3.84 (b), P-uniqueness of τ_x is equivalent to

$$P(X=x|C_X) \gneq_P 0. \tag{6.3}$$

Hence, we may replace the assumption of P-uniqueness of τ_x by $P(X=x|C_X) \gneq_P 0$, which is a shortcut for

$$P\big(\{\omega \in \Omega\colon P(X=x|C_X)(\omega) > 0\}\big) = 1. \tag{6.4}$$

◁

Remark 6.4 (Unbiased With Respect to $\mathscr{F}_1$) Note that Definition 6.2 refers to *total effects* true outcome variables (see Rem. 4.20). Considering these true outcome variables $\tau_x = E^{X=x}(Y|C_X)$ we condition on a global potential confounder C_X, and with it on its generated σ-algebra $\sigma(C_X)$ (see SN-Def. 10.2 and SN-Rem. 10.3), which is identical to $\mathscr{F}_1$ [see Def. 4.4 (ii)]. Therefore, if it is not clear that we are talking about unbiasedness with respect to total effects, then we may say *unbiased with respect to $\mathscr{F}_1$*, because for direct and indirect effects we would consider other true outcome variables, which would be defined by conditioning on another σ-algebra of potential confounders. Such a σ-algebra would also include those events that are in between X and the intermediate variable considered. ◁

Example 6.5 (No Treatment for Joe) In the example displayed in Table 3.2, the person variable U plays the role of C_X, and $\mathscr{F}_1 = \sigma(U)$. Furthermore, the co-domain Ω'_X of X is any subset of $\mathbb{R}$ containing the elements 0 and 1. The values of the two true outcome variables τ_0 and τ_1 are displayed in the table. Whereas τ_0 is the only element in the set $\mathscr{E}^{X=0}(Y|U)$

(which implies that τ_0 is P-unique), τ_1 is not the only element in the set $\mathscr{E}^{X=1}(Y|U)$, because the random variable τ_1^* displayed in the last column of the table is also an element of $\mathscr{E}^{X=1}(Y|U)$. Furthermore, τ_1 and τ_1^* are not identical with probability 1. Instead, $\tau_1(\omega) \neq \tau_1^*(\omega)$ for $\omega \in \{\omega_1, \ldots, \omega_4\}$ and $P(\{\omega_1, \ldots \omega_4\}) = .5$. Hence, in this example, the conditional expectation value $E(Y|X=0)$ is unbiased, whereas $E(Y|X=1)$ as well as the conditional expectation $E(Y|X)$ are not unbiased, because τ_1 is not P-unique. ◁

Remark 6.6 (Unbiased Estimators and Unbiased Parameters) Unbiasedness in statistics usually refers to *estimators* of a parameter. However, if the expectation of the estimator is not identical to the parameter to be estimated, then the expectation of the estimator itself is also biased if used for drawing inferences on the original parameter. In particular, although the sample mean is an unbiased estimator of $E(Y|X{=}x)$, it can be a biased estimator of $E(\tau_x)$, the parameter of interest. Hence, the conditional expectation value $E(Y|X{=}x)$ can also be biased in the sense that it is not identical to $E(\tau_x)$, and this is the basic idea of Definition 6.2. ◁

Remark 6.7 (Expectations with Respect to the Measure $P^{X=x}$) If $P(X{=}x) > 0$, then

$$E(Y|X{=}x) \;=\; E^{X=x}(Y) \tag{6.5}$$

[see Eqs. (3.23), (3.27), or SN-Cor. 9.5]. Hence, if $P(X{=}x) > 0$, then the conditional expectation value $E(Y|X{=}x)$ is unbiased if and only if the expectation $E^{X=x}(Y)$ of Y with respect to the conditional probability measure $P^{X=x}$ is unbiased, that is, if and only if τ_x is P-unique and

$$E^{X=x}(Y) \;=\; E(\tau_x). \tag{6.6}$$

◁

Remark 6.8 (Identification of $E(\tau_x)$) Although trivial, let us emphasize that unbiasedness of $E(Y|X{=}x)$ is important because it gives us access to the expectation of the true outcome variable τ_x. If $E(Y|X{=}x)$ is unbiased, then, according to Equation (6.2), an estimate of $E(Y|X{=}x)$ is also an estimate of $E(\tau_x)$. ◁

In the following theorem, we present three conditions that are equivalent to unbiasedness of $E(Y|X{=}x)$.

Theorem 6.9 (Equivalent Conditions of Unbiasedness of $E(Y|X{=}x)$)
Let the assumptions 6.1 hold and assume that τ_x is P-unique. Then each of the following three equations is equivalent to unbiasedness of $E(Y|X{=}x)$.

$$E^{X=x}(\tau_x) \;=\; E(\tau_x). \tag{6.7}$$

$$E(\tau_x|1_{X=x}) \;=\; E(\tau_x). \tag{6.8}$$

$$E^{X=x}(\varepsilon_x) \;=\; E(\varepsilon_x) \;=\; 0, \quad \textit{where} \quad \varepsilon_x \;:=\; \tau_x - E^{X=x}(Y). \tag{6.9}$$

(Proof p. 146)

In the subsequent sections and chapters we will often refer to Equation (6.8) stating *mean-independence* of τ_x from the indicator $1_{X=x}$. Remember, the shortcut for this equation is $\tau_x \vdash 1_{X=x}$. That is,

$$\tau_x \vdash 1_{X=x} \quad \Leftrightarrow \quad E(\tau_x | 1_{X=x}) \;=\; E(\tau_x) \tag{6.10}$$

[see Proposition (3.55)]. Hence, we may read $\tau_x \vdash 1_{X=x}$ as τ_x *is mean-independent of* $1_{X=x}$.

Remark 6.10 (Sufficient Conditions of Unbiasedness) In Corollary 7.27, it will be shown that $E^{X=x}(\tau_x) = E(\tau_x)$ as well as $\tau_x \vdash 1_{X=x}$ follow from independence of τ_x and $1_{X=x}$, which itself follows from $C_X \perp\!\!\!\perp 1_{X=x}$, that is from independence of a global potential confounder C_X of X and the indicator $1_{X=x}$ (see Theorem 8.11). Note that $C_X \perp\!\!\!\perp 1_{X=x}$ can be created by randomized assignment of the observational unit to treatment x (see the example presented in Table 6.2). In the subsequent chapters we will also treat other sufficient conditions of unbiasedness of $E(Y|X=x)$. ◁

In empirical applications in which there is no randomized assignment of the observational unit to one of the treatment conditions, unbiasedness of $E(Y|X=x)$ is not very likely. However, if we additionally consider a (uni- or multivariate) covariate Z of X and the conditional expectation values $E(Y|X=x,Z=z)$, then unbiasedness of these parameters and of the conditional expectation $E(Y|X,Z)$ (see section 6.2) is much more realistic, even beyond experiments with randomized assignment of the unit to a treatment condition, that is, even in quasi-experiments.

6.2 Unbiasedness of $E(Y|X,Z)$ and its Values $E(Y|X=x,Z=z)$

Now we extend the concept of unbiasedness to conditioning on a on a (possibly multivariate) random variable Z. Note that $Z := (Z_1,\ldots,Z_m)$ is a random variable on $(\Omega,\mathscr{A},P)$ if and only if $Z_1,\ldots,Z_m$ are random variables on $(\Omega,\mathscr{A},P)$ (see SN-Th. 2.38 and SN-Def. 5.1).

Reading the following definition, note that

$$P^{X=x}(Z=z)>0 \quad \Leftrightarrow \quad P^{Z=z}(X=x)>0 \quad \Leftrightarrow \quad P(X=x,Z=z)>0 \tag{6.11}$$

and that $P(X=x,Z=z)>0$ implies

$$E^{X=x}(Y|Z=z) \;=\; E^{Z=z}(Y|X=x) \;=\; E(Y|X=x,Z=z)\,. \tag{6.12}$$

Also remember that $P^{Z=z}$*-uniqueness of* τ_x means

$$P^{Z=z}\big(\{\,\omega\in\Omega\colon \tau_x(\omega)=\tau_x^*(\omega)\}\big) \;=\; 1, \quad \forall\,\tau_x,\tau_x^*\in\mathscr{E}^{X=x}(Y|C_X), \tag{6.13}$$

and that this property follows from P-uniqueness of τ_x [see SN-Box 14.1 (v)]. Also note that $P^{Z=z}$-uniqueness of τ_x is equivalent to $P(X=x|C_X) \underset{P^{Z=z}}{>} 0$, which is a shortcut for

$$P^{Z=z}\big(\{\,\omega\in\Omega\colon P(X=x|C_X)>0\}\big) \;=\; 1. \tag{6.14}$$

Definition 6.11 (Unbiasedness of $E(Y|X=x,Z=z)$, $E^{X=x}(Y|Z)$, and $E(Y|X,Z)$)
Let the Assumptions 6.1 hold and let Z be a random variable on $(\Omega,\mathscr{A},P)$ with value space $(\Omega'_Z,\mathscr{A}'_Z)$. Then

(i) $E(Y|X=x,Z=z)$ *is called unbiased if* $P(X=x,Z=z)>0$, τ_x *is* $P^{Z=z}$*-unique, and*

$$E(Y|X=x,Z=z) \;=\; E(\tau_x|Z=z). \tag{6.15}$$

(ii) $E^{X=x}(Y|Z)$ *is called unbiased if* τ_x *is* P*-unique and*

$$E^{X=x}(Y|Z) \underset{P}{=} E(\tau_x|Z). \tag{6.16}$$

(iii) $E(Y|X,Z)$ *is called unbiased if, for all values* $x \in X(\Omega)$*, the conditional expectations* $E^{X=x}(Y|Z)$ *are unbiased.*

Remark 6.12 (Reference to $\mathscr{F}_1$) Again, if there is ambivalence, then we may add the reference to $\mathscr{F}_1$ in order to clarify that we are talking about total effects. The background of this note have been detailed already in Remark 6.4. ◁

Remark 6.13 (Identification of $E(\tau_x|Z=z)$ and $E(\tau_x|Z)$) According to Equation (6.15), if $E(Y|X=x,Z=z)$ is unbiased, then we can identify the $(Z=z)$-conditional expectation value $E(\tau_x|Z=z)$ by $E(Y|X=x,Z=z)$. Hence, an estimate of $E(Y|X=x,Z=z)$ is also an estimate of $E(\tau_x|Z=z)$ if $E(Y|X=x,Z=z)$ is unbiased. Analogously, If $E^{X=x}(Y|Z)$ is unbiased, then, according to Equation (6.16), we can identify the Z-conditional expectation $E(\tau_x|Z)$ by the Z-conditional expectation $E^{X=x}(Y|Z)$ of Y with respect to the conditional probability measure $P^{X=x}$. Hence, an estimate of $E^{X=x}(Y|Z)$ is also an estimate of $E(\tau_x|Z)$ provided that $E^{X=x}(Y|Z)$ is unbiased. ◁

Now we turn to some conditions that are equivalent to unbiasedness of a conditional expectation $E^{X=x}(Y|Z)$. The following lemma will be useful in the proof of the Theorem 6.15. Note that a crucial assumption in the following lemma is that Z is a covariate of X, which implies C_X-measurability of Z.

Lemma 6.14 (A Decomposition of the True Outcome Variable)
Let the Assumptions 6.1 hold and let Z *be a covariate of* X *with value space* $(\Omega'_Z, \mathscr{A}'_Z)$*. Then*

$$\tau_x := E^{X=x}(Y|C_X) \underset{P^{X=x}}{=} E^{X=x}(Y|Z) + \varepsilon_x \tag{6.17}$$

with

$$\varepsilon_x \underset{P^{X=x}}{=} \tau_x - E^{X=x}(\tau_x|Z) \tag{6.18}$$

and

$$E^{X=x}(\varepsilon_x|Z) \underset{P^{X=x}}{=} 0. \tag{6.19}$$

(Proof p. 148)

In the following theorem we present three conditions, each of which is equivalent to unbiasedness of $E^{X=x}(Y|Z)$. These conditions are also used in the proofs of sufficient conditions of unbiasedness (see chs. 8 to **??**).

Theorem 6.15 (Equivalent Conditions of Unbiasedness of $E^{X=x}(Y|Z)$)
Let the Assumptions 6.1 hold, let Z be a covariate of X with value space $(\Omega'_Z, \mathscr{A}'_Z)$, and assume that τ_x is P-unique. Then each of the following three equations is equivalent to $E^{X=x}(Y|Z)$ being unbiased.

$$E^{X=x}(\tau_x|Z) \underset{P}{=} E(\tau_x|Z). \tag{6.20}$$

$$E(\tau_x|1_{X=x}, Z) \underset{P}{=} E(\tau_x|Z). \tag{6.21}$$

$$E(\varepsilon_x|Z) \underset{P}{=} 0, \quad \textit{where} \quad \varepsilon_x := \tau_x - E^{X=x}(Y|Z). \tag{6.22}$$

(Proof p. 148)

In the subsequent sections and chapters we will often refer to Equation (6.21) stating Z-conditional mean-independence of τ_x from the indicator $1_{X=x}$. The shortcut for this equation is $\tau_x \vdash 1_{X=x}|Z$. That is,

$$\tau_x \vdash 1_{X=x}|Z \quad \Leftrightarrow \quad E(\tau_x|1_{X=x}, Z) \underset{P}{=} E(\tau_x|Z) \tag{6.23}$$

[see Prop. (3.56)]. Hence, we may read $\tau_x \vdash 1_{X=x}|Z$ as *τ_x is Z-conditionally mean-independent of $1_{X=x}$*.

Remark 6.16 (Unbiasedness of a Conditional Expectation Value) If $P(X{=}x, Z{=}z) > 0$, then

$$E(Y|X{=}x, Z{=}z) = E^{X=x}(Y|Z{=}z) \tag{6.24}$$

[see Rem. 3.88 and SN-Eq. (14.35)]. If we additionally presume that Z is a covariate of X and that τ_x is $P^{Z=z}$-unique, then Equation (6.24) and Definition 6.11 (i) imply that the $(X{=}x, Z{=}z)$-conditional expectation value $E(Y|X{=}x, Z{=}z)$ of Y is unbiased if and only if

$$E^{X=x}(Y|Z{=}z) = E(\tau_x|Z{=}z). \tag{6.25}$$

◁

Theorem 6.17 (Unbiasedness of a Conditional Expectation Value)
Let the Assumptions 6.1 hold, let Z be a covariate of X with value space $(\Omega'_Z, \mathscr{A}'_Z)$, assume $P(X{=}x, Z{=}z) > 0$, and that τ_x is $P^{Z=z}$-unique. Then $E(Y|X{=}x, Z{=}z)$ is unbiased if and only if

$$E^{X=x}(\tau_x|Z{=}z) = E(\tau_x|Z{=}z). \tag{6.26}$$

(Proof p. 149)

Note that, under the assumptions of Theorem 6.17, Equation (6.26) is equivalent to

$$E(\tau_x|X{=}x, Z{=}z) = E(\tau_x|Z{=}z) \tag{6.27}$$

(see Rem. 3.88).

6.3 Unbiasedness of Prima Facie Effects

In the Definition 6.18 we introduce the concepts of unbiasedness of the *prima facie effect*

$$PFE_{xx'} := E(Y|X{=}x) - E(Y|X{=}x'). \tag{6.28}$$

For a covariate Z of X with value space $(\Omega'_Z, \mathscr{A}'_Z)$ we also extend unbiasedness to a *Z-conditional prima facie effect function* $PFE_{Z;xx'}\colon \Omega'_Z \to \mathbb{R}$ satisfying

$$PFE_{Z;xx'}(Z) \underset{P}{=} E^{X=x}(Y|Z) - E^{X=x'}(Y|Z), \tag{6.29}$$

where $PFE_{Z;xx'}(Z)$ denotes the composite function of Z and $PFE_{Z;xx'}$. The function $PFE_{Z;xx'}$ assigns to each value $z \in \Omega'_Z$ a $(Z{=}z)$-conditional prima facie effect of x compared to x'. For simplicity, we also call the composition $PFE_{Z;xx'}(Z)$ a Z-conditional prima facie effect variable. Finally, presuming $P(X{=}x, Z{=}z),\ P(X{=}x', Z{=}z) > 0$, we define the *$(Z{=}z)$-conditional prima facie effect*

$$PFE_{Z;xx'}(z) := E(Y|X{=}x, Z{=}z) - E(Y|X{=}x', Z{=}z). \tag{6.30}$$

Definition 6.18 (Unbiasedness of a Prima Facie Effect)
Let the Assumptions 6.1 hold.

(i) The prima facie effect $PFE_{xx'}$ is called unbiased, if τ_x and $\tau_{x'}$ are P-unique, and

$$PFE_{xx'} = E(\tau_x - \tau_{x'}). \tag{6.31}$$

(ii) Additionally, let Z be a random variable on $(\Omega, \mathscr{A}, P)$ with value space $(\Omega'_Z, \mathscr{A}'_Z)$. A Z-conditional prima facie effect function $PFE_{Z;xx'}$ is called unbiased if τ_x and $\tau_{x'}$ are P-unique, and

$$PFE_{Z;xx'}(Z) \underset{P}{=} E(\tau_x - \tau_{x'}|Z). \tag{6.32}$$

(iii) Finally, let z be a value of Z such that $P(X{=}x, Z{=}z), P(X{=}x', Z{=}z) > 0$. Then $PFE_{Z;xx'}(z)$ is called unbiased, if τ_x and $\tau_{x'}$ are $P^{Z=z}$-unique, and

$$PFE_{Z;xx'}(z) = E(\tau_x - \tau_{x'}|Z{=}z). \tag{6.33}$$

Note again that all these concepts of unbiasedness refer to total effects. If this is ambiguous we may say 'unbiased with respect to $\mathscr{F}_1$', because for direct and indirect effects we would consider other true outcome variables, conditioning on another σ-algebra of potential confounders, which would also include those events that are in between X and the intermediate variable considered.

Example 6.19 (Conditioning on $\{Z{=}z\}$) If we consider a random experiment in which we sample a person u from a set Ω_U of persons (the population), and if z represents a subpopulation, that is, if z represents a subset of Ω_U such as 'males', then $PFE_{Z;xx'}(z)$ is identical to the difference

$$E^{X=x}(Y|Z{=}z) - E^{X=x'}(Y|Z{=}z) = E(Y|X{=}x, Z{=}z) - E(Y|X{=}x', Z{=}z),$$

Box 6.1 Unbiasedness

Let the Assumptions 6.1 hold and let Z be a random variable on $(\Omega, \mathscr{A}, P)$ with value space $(\Omega'_Z, \mathscr{A}'_Z)$. Then unbiasedness of various conditional expectations, their values, and their differences, is defined as follows:

Unbiasedness of …	**is defined by** …
$E(Y \mid X{=}x)$	τ_x is P-unique and satisfies $E(Y \mid X{=}x) = E(\tau_x)$.
$E(Y \mid X)$	For all $x \in X(\Omega)$: $E(Y \mid X{=}x)$ is unbiased.
$E(Y \mid X{=}x, Z{=}z)$	z is a value of Z such that $P(X{=}x, Z{=}z) > 0$, τ_x is $P^{Z=z}$-unique and satisfies $E(Y \mid X{=}x, Z{=}z) = E(\tau_x \mid Z{=}z)$.
$E^{X=x}(Y \mid Z)$	τ_x is P-unique and satisfies $E^{X=x}(Y \mid Z) \underset{P}{=} E(\tau_x \mid Z)$.
$E(Y \mid X, Z)$	For all $x \in X(\Omega)$: $E^{X=x}(Y \mid Z)$ is unbiased.
$PFE_{xx'}$	τ_x and $\tau_{x'}$ are P-unique and $PFE_{xx'} = E(\tau_x - \tau_{x'})$.
$PFE_{Z;xx'}(z)$	z is a value of Z satisfying $P(X{=}x, Z{=}z)$, $P(X{=}x', Z{=}z) > 0$, τ_x and $\tau_{x'}$ are $P^{Z=z}$-unique and $PFE_{Z;xx'}(z) = E(\tau_x - \tau_{x'} \mid Z{=}z)$.
$PFE_{Z;xx'}$, $PFE_{Z;xx'}(Z)$	τ_x and $\tau_{x'}$ are P-unique and $PFE_{Z;xx'}(Z) \underset{P}{=} E(\tau_x - \tau_{x'} \mid Z)$.

where we condition on the two values x and x' of the treatment variable and the event $\{Z{=}z\} = \{\omega \in \Omega\colon Z(\omega) = z\}$ that the person drawn is an element of the subpopulation represented by z. ◁

Box 6.1 summarizes the definitions of unbiasedness of various conditional expectations, their values, and prima facie effects.

Remark 6.20 (Estimability of Conditional Expectations) While true outcome variables τ_x and their values are not directly estimable in a data sample unless rather restrictive assumptions are introduced, the conditional expectation values $E(Y \mid X{=}x)$, $E(Y \mid X{=}x, Z{=}z)$, and the conditional expectations $E^{X=x}(Y \mid Z)$ and $E(Y \mid X, Z)$ *can* be estimated under realistic assumptions, and the same is true for the conditional and unconditional prima facie effects. As we shall see, this implies that causal average total effects and causal $(Z{=}z)$-conditional total effects can be estimated as well, provided that we assume that the conditional expectations $E(Y \mid X)$, $E^{X=x}(Y \mid Z)$, and $E(Y \mid X, Z)$ mentioned above are unbiased (see section 6.4 for details). ◁

Remark 6.21 (Unbiasedness and Randomization) In chapter 8 we show that unbiasedness of the conditional expectations, their values, and the prima facie effects can be created by randomized assignment of the observational unit to one of the treatment conditions. Unbiasedness can also be strived for by covariate selection, that is, we may try to select covariates $Z_1, \ldots, Z_m$ such that unbiasedness of the conditional expectations $E^{X=x}(Y \mid Z)$ holds for the m-variate covariate $Z := (Z_1, \ldots, Z_m)$ and all values x of X. ◁

Remark 6.22 (Unbiasedness and Covariate Selection) Unfortunately, unbiasedness cannot be used as a criterion for covariate selection. The reason is that it cannot be tested

empirically, because the definitions involve the true outcome variables τ_x. These variables even cannot be estimated unless very restrictive assumptions are introduced. This has been discussed in some detail by Holland (1986) and has been called the "fundamental problem of causal inference" (see also the preface). However, in chapters 8 to **??** we introduce other causality conditions that *can* be tested empirically and that imply unbiasedness. ◁

6.4 Identification of Causal Total Effects

In chapter 5 we introduced causal average total effects and causal conditional total effect functions, which, in the first place, are of a purely theoretical nature. They just define what we are interested in, for example, in studies evaluating the effects of a treatment, an intervention, or an exposition. Now we study how causal total effects can be identified by estimable parameters, and how the causal conditional total effect functions can be identified by estimable variables.

6.4.1 Identification of the Causal Average Total Effect

In Definition 5.8 we defined the causal average total effect

$$ATE_{xx'} = E(\tau_x - \tau_{x'}), \tag{6.34}$$

presuming that τ_x and $\tau_{x'}$ are P-unique. In Definition 6.18 we defined unbiasedness of the prima facie effect $PFE_{xx'}$ by P-uniqueness of τ_x and $\tau_{x'}$ and

$$PFE_{xx'} = E(\tau_x - \tau_{x'}). \tag{6.35}$$

Hence, these two definitions immediately yield the following corollary.

Corollary 6.23 (Identifying the Causal Average Effect by $PFE_{xx'}$)
Let the Assumptions 6.1 hold and assume that $PFE_{xx'}$ is unbiased. Then

$$ATE_{xx'} = PFE_{xx'}. \tag{6.36}$$

Remark 6.24 (Complete Reaggregation) Looking at Equation (6.34) shows that we can reaggregate the true total effect variable $\delta_{xx'} = \tau_x - \tau_{x'}$. Note that this does not mean to ignore the potential confounders of X, which, by definition, are measurable with respect to the global potential confounder C_X (see Def. 4.4). In general,

$$E(Y|X{=}x) - E(Y|X{=}x') \neq E(\tau_x - \tau_{x'}) = E(\tau_x) - E(\tau_{x'}).$$

However, if the prima facie effect $PFE_{xx'} = E(Y|X{=}x) - E(Y|X{=}x')$ is unbiased, then both sides of this inequality are identical. Considering the expectation $E(\tau_x - \tau_{x'})$ we (completely) *reaggregate* the true total effect variable $\delta_{xx'}$. That is, the true total effect variable is coarsened to one single number, the expectation of the true total effect variable. With such a reaggregation, we lose information. However, the resulting causal average total effect is still unbiased. ◁

Remark 6.25 (Ignoring Potential Confounders of X) In contrast to reaggregation, just considering the prima facie effect $E(Y|X{=}x) - E(Y|X{=}x')$, and in this sense, *ignoring potential confounders of X*, may lead to a completely wrong conclusion about the causal average total effect of x compared to x' on Y, unless $E(Y|X{=}x) - E(Y|X{=}x')$ is unbiased. In Table 6.1 we present an example in which the effect of treatment 1 compared to treatment 0 is reversed. While the causal average total effect, that is, the expectation of the true total effect variable is $E(\delta_{10}) = 10$, the corresponding prima facie effect is $PFE_{10} = E(Y|X=1) - E(Y|X=0) = -5.857$. Considering Box 6.2 and comparing Equations (i) and (ii) to each other reveals why such a reversal of effects can occur. ◁

According to the following theorem we can also identify the causal average total effect $ATE_{xx'}$ if the prima facie effect $PFE_{xx'}$ is not unbiased. It suffices to assume that the Z-conditional prima facie effect variable

$$PFE_{Z;xx'}(Z) \underset{P}{=} E^{X=x}(Y|Z) - E^{X=x'}(Y|Z) \tag{6.37}$$

is unbiased. This theorem is the theoretical foundation for the analysis of causal average total effects beyond the randomized experiment.

Theorem 6.26 (Identifying the Causal Average Total Effect by $PFE_{Z;xx'}$)
Let the Assumptions 6.1 hold, let Z be a random variable on $(\Omega, \mathscr{A}, P)$ with value space $(\Omega'_Z, \mathscr{A}'_Z)$, and assume that $PFE_{Z;xx'}$ is unbiased. Then

$$ATE_{xx'} = E\big(PFE_{Z;xx'}(Z)\big). \tag{6.38}$$

(Proof p. 150)

Equation (6.38) reveals that we can reaggregate the prima facie effect function $PFE_{Z;xx'}$ to obtain a single number. If $PFE_{Z;xx'}$ is unbiased, then this does *not* mean to ignore the potential confounders of X. Instead, we just reaggregate (coarsen) the causal conditional total effects to obtain a single number, the causal average total effect $ATE_{xx'}$.

Remark 6.27 (Z-Adjusted Conditional Expectation Value of Y) Inserting Equation (6.37) into the right-hand side of Equation (6.37) yields

$$ATE_{xx'} = E\big(E^{X=x}(Y|Z)\big) - E\big(E^{X=x'}(Y|Z)\big). \tag{6.39}$$

The first term on the right-hand side of this equation is called the *Z-adjusted $(X{=}x)$-conditional expectation value of Y*, and the second term the *Z-adjusted $(X{=}x')$-conditional expectation value of Y*. ◁

6.4.2 Identification of a Causal Conditional Total Effect Function

In Definition 5.17 we introduced the causal conditional total effect function $CTE_{V;xx'}$ and its composite $CTE_{V;xx'}(V)$ by

$$CTE_{V;xx'}(V) \underset{P}{=} E(\tau_x - \tau_{x'}|V), \tag{6.40}$$

presuming that τ_x and $\tau_{x'}$ are P-unique. In Definition 6.18 (ii) we defined *unbiasedness* of the prima facie effect function $PFE_{Z;xx'}$ by

$$PFE_{Z;xx'}(Z) \underset{P}{=} E(\tau_x - \tau_{x'}|Z), \tag{6.41}$$

again presuming P-uniqueness of τ_x and $\tau_{x'}$. Hence, for $V = Z$, these two definitions immediately imply the following corollary, according to which a conditional prima facie effect function $PFE_{Z;xx'}$ is a causal conditional total effect function $CTE_{Z;xx'}$, if $PFE_{Z;xx'}$ is unbiased.

Corollary 6.28 (Identifying the Causal Z-Conditional Effect Function by $PFE_{Z;xx'}$) *Let the Assumptions 6.1 hold, let Z be a random variable on $(\Omega, \mathscr{A}, P)$ with value space $(\Omega'_Z, \mathscr{A}'_Z)$, and assume that $PFE_{Z;xx'}$ is unbiased. Then*

$$CTE_{Z;xx'}(Z) \underset{P}{=} PFE_{Z;xx'}(Z) \underset{P}{=} E^{X=x}(Y|Z) - E^{X=x'}(Y|Z). \tag{6.42}$$

Remark 6.29 (A Measurability Assumption for V) According to Theorem 6.26, the expectation of an unbiased prima facie effect variable $PFE_{Z;xx'}(Z)$ is identical to the causal average total effect $ATE_{xx'}$. In Theorem 6.30 we extend this result to a causal V-conditional total effect variable $CTE_{V;xx'}(V)$. In this theorem we do not only assume $E^{X=x}(Y|Z)$ and $E^{X=x'}(Y|Z)$, the two subtrahends of $PFE_{Z;xx'}(Z)$, to be unbiased, but also that $\sigma(V) \subset \sigma(1_{X=x}, Z)$ and $\sigma(V) \subset \sigma(1_{X=x'}, Z)$. This second assumption is satisfied, for example, if V is Z-measurable, or for $V = X$ and X is dichotomous such that $1_{X=x'} = 1 - 1_{X=x}$. ◁

Theorem 6.30 (Identifying a Causal V-Conditional Total Effect Function via $PFE_{Z;xx'}$) *Let the Assumptions 6.1 hold, let Z be a random variable on $(\Omega, \mathscr{A}, P)$ with value space $(\Omega'_Z, \mathscr{A}'_Z)$, let V be a random variable on $(\Omega, \mathscr{A}, P)$ with value space $(\Omega'_V, \mathscr{A}'_V)$ such that $\sigma(V) \subset \sigma(1_{X=x}, Z)$, $\sigma(V) \subset \sigma(1_{X=x'}, Z)$, and assume that $E^{X=x}(Y|Z)$ and $E^{X=x'}(Y|Z)$ are unbiased. Then*

$$CTE_{V;xx'}(V) \underset{P}{=} E\left(PFE_{Z;xx'}(Z) \,\middle|\, V\right) \tag{6.43}$$

$$\underset{P}{=} E\left(E^{X=x}(Y|Z) \,\middle|\, V\right) - E\left(E^{X=x'}(Y|Z) \,\middle|\, V\right). \tag{6.44}$$

(Proof p. 150)

Remark 6.31 (Generalizing Identification of a Causal V-Conditional Total Effect Function) Under slightly more restrictive assumptions, we can extend reaggregation of $PFE_{Z;xx'}(Z)$ to conditioning on an (X, Z)-measurable random variable V. Instead of unbiasedness of $E^{X=x}(Y|Z)$ and $E^{X=x'}(Y|Z)$, in Theorem 6.32 we assume

$$E(\tau_x|X, Z) \underset{P}{=} E(\tau_x|Z) \quad \text{and} \quad E(\tau_{x'}|X, Z) \underset{P}{=} E(\tau_{x'}|Z). \tag{6.45}$$

If X is dichotomous, then this assumption is equivalent to

$$E(\tau_x|1_{X=x}, Z) \underset{P}{=} E(\tau_x|Z) \quad \text{and} \quad E(\tau_{x'}|1_{X=x'}, Z) \underset{P}{=} E(\tau_{x'}|Z), \tag{6.46}$$

because, in this case, $\sigma(X, Z) = \sigma(1_{X=x}, Z) = \sigma(1_{X=x'}, Z)$ and Proposition (6.46) is equivalent to unbiasedness of $E^{X=x}(Y|Z)$ and $E^{X=x'}(Y|Z)$ (see Th. 6.15). If X can take on more than two values, each with a positive probability, then Proposition (6.45) implies unbiasedness of $E^{X=x}(Y|Z)$ and $E^{X=x'}(Y|Z)$, but it is not equivalent to unbiasedness of these two conditional expectations. (For more details see ch. 7, in particular Table 7.6). ◁

Theorem 6.32 (Identifying a Causal V-Conditional Total Effect Function via $PFE_{Z;xx'}$) *Let the Assumptions 6.1 hold, let Z be a random variable on $(\Omega,\mathscr{A},P)$ with value space $(\Omega'_Z,\mathscr{A}'_Z)$, and let V be a random variable on $(\Omega,\mathscr{A},P)$ with value space $(\Omega'_V,\mathscr{A}'_V)$ such that $\sigma(V)\subset\sigma(X,Z)$. Furthermore, assume that Proposition (6.45) holds. Then*

$$CTE_{V;xx'}(V) \underset{P}{=} E\big(PFE_{Z;xx'}(Z)\,\big|\,V\big) \tag{6.47}$$

$$\underset{P}{=} E\big(E^{X=x}(Y|Z)\,\big|\,V\big) - E\big(E^{X=x'}(Y|Z)\,\big|\,V\big). \tag{6.48}$$

(Proof p. 150)

Remark 6.33 (Identifying a Causal V-Conditional Total Effect Function) Although the assumptions in Theorems 6.30 and 6.32 differ from each other if X has more than two values, Equations (6.43) and (6.47) are identical. Note that $CTE_{V;xx'}(V)$ is not necessarily identical (almost surely) to

$$PFE_{V;xx'}(V) \underset{P}{=} E^{X=x}(Y|V) - E^{X=x'}(Y|V),$$

which is obvious if we compare the right-hand side of this equation to the right-hand side of Equation (6.48). Instead, $CTE_{V;xx'}(V)\underset{P}{=}PFE_{V;xx'}(V)$ if $PFE_{V;xx'}(V)$ is unbiased, an assumption that neither follows from the assumptions made in Theorem 6.30 nor from those made in Theorem 6.32. Hence, Theorem 6.32 offers a way to identify $CTE_{V;xx'}(V)$ even if $PFE_{V;xx'}(V)$ is biased. ◁

6.4.3 Identification of a Causal Conditional Total Effect

The causal conditional total effect $CTE_{V;xx'}(v)$ has been defined by

$$CTE_{V;xx'}(v) = E(\tau_x - \tau_{x'}|V=v), \tag{6.49}$$

presuming $P(X=x,V=v)$, $P(X=x',V=v)>0$ and that τ_x and $\tau_{x'}$ are $P^{V=v}$-unique [see Def. 5.17 (i)]. Furthermore, unbiasedness of the conditional prima facie effect $PFE_{Z;xx'}(z)$ has been defined by

$$PFE_{Z;xx'}(z) = E(\tau_x - \tau_{x'}|Z=z), \tag{6.50}$$

assuming that Z is a covariate of X, and $P^{Z=z}$-uniqueness of τ_x and $\tau_{x'}$ [see Def. 6.18 (iii)]. For $V=Z$, these two definitions immediately imply the following corollary.

Corollary 6.34 (Identifying the Causal $(Z=z)$-Conditional Total Effect by $PFE_{Z;xx'}(z)$) *Let the Assumptions 6.1 hold, let Z be a random variable on $(\Omega,\mathscr{A},P)$ with value space $(\Omega'_Z,\mathscr{A}'_Z)$, and assume that $P(X=x,Z=z)$, $P(X=x',Z=z)>0$. If $PFE_{Z;xx'}(z)$ is unbiased, then*

$$CTE_{Z;xx'}(z) = PFE_{Z;xx'}(z) = E^{X=x}(Y|Z=z) - E^{X=x'}(Y|Z=z). \tag{6.51}$$

The assumption that $PFE_{Z;xx'}(z)$ is unbiased comprises the assumption that τ_x and $\tau_{x'}$ are $P^{Z=z}$-unique. This assumption has already been explained in more detail in Remarks 5.14 and 5.16.

According to SN-Definition 10.33, Theorem 6.30 and Theorem 6.32 immediately imply the following corollary about the identification of the causal $(V=v)$-conditional total effect $CTE_{V;xx'}(v)$. According to SN-Remark 10.35, $CTE_{V;xx'}(v)$ is a uniquely defined number.

Corollary 6.35 (Identifying the Causal $(V=v)$-Conditional Total Effect via $PFE_{Z;xx'}$) *If the Assumptions of Theorem 6.30 or of Theorem 6.32 hold and $P(X=x,V=v)>0$, $P(X=x',V=v)>0$, then*

$$\begin{aligned} CTE_{V;xx'}(v) &= E\big(PFE_{Z;xx'}(Z)\,\big|\,V=v\big) && (6.52)\\ &= E\big(E^{X=x}(Y|Z)\,\big|\,V=v\big)-E\big(E^{X=x'}(Y|Z)\,\big|\,V=v\big). && (6.53)\end{aligned}$$

Remark 6.36 (Unbiasedness of $PFE_{Z;xx'}(z)$ vs. Unbiasedness of $PFE_{V;xx'}(v)$) Again, note that $CTE_{V;xx'}(v)$ is not necessarily identical to

$$PFE_{V;xx'}(v) = E^{X=x}(Y|V=v)-E^{X=x'}(Y|V=v).$$

Comparing the right-hand side of this equation to the right-hand side of Equation (6.53) reveals the difference. However, $CTE_{V;xx'}(v)=PFE_{V;xx'}(v)$, if $PFE_{V;xx'}(v)$ is unbiased. Note that unbiasedness of $PFE_{V;xx'}(v)$ is not implied by the assumptions made in Corollary 6.35. Hence, Corollary 6.35 offers a way to identify $CTE_{V;xx'}(v)$ even if $PFE_{V;xx'}(v)$ is biased. The crucial assumption in Theorem 6.30 is that $PFE_{Z;xx'}(z)$ is unbiased. Similarly, the crucial assumption in Theorem 6.32 is that Proposition (6.45) holds, which is an assumption about Z, not about V. ◁

6.5 Three Examples

Tables 6.1 to 6.3 show parameters pertaining to fictive *random experiments* such as the single-unit trials described in chapter 2. Among these parameters are the individual expectation values $E(Y|X=x,U=u)$ given the treatment conditions and the individual treatment probabilities $P(X=1|U=u)$. The parameters presented in the tables can be used to generate sample data that would result if the random experiments to which the tables refer were conducted n times. [1]

6.5.1 Commonalities of all Examples

For simplicity, we consider single-unit trials in which no *fallible* covariate is observed and in which there is neither a second treatment variable nor any other variable that is simultaneous to the treatment variable. In this case, the set

$$\Omega = \Omega_1\times\Omega_2\times\Omega_3 = \Omega_U\times\Omega_X\times\mathbb{R}$$

[1] Although the focus of this book is on theory and not on data analysis, we also provide sample data for each table on the home page of this book: www.causal-effects.de. These and other examples of this type as well as a data sample generated by these examples can easily be created with the PC-program CausalEffectsXplorer that is also provided on www.causal-effects.de, together with an extensive help file providing the most important concepts and formulas.

suffices to describe the set of possible outcomes of the random experiment, where $\Omega_U = \{u_1, \ldots, u_6\}$ and $\Omega_X = \{\textit{treatment}, \textit{control}\}$. Furthermore, we consider the product σ-algebra $\mathscr{A} = \mathscr{P}(\Omega_U) \otimes \mathscr{P}(\Omega_X) \otimes \mathscr{B}$, where $\mathscr{B}$ denotes the Borel σ-algebra on $\mathbb{R}$ (see SN-section 1.2.3). The probability measure P on $(\Omega, \mathscr{A})$ is only partly known.

Remark 6.37 (σ-Algebra of Potential Confounders) In all examples, we consider the projection $U\colon \Omega \to \Omega_U$ (see SN-section 2.3.5) with value space $\big(\Omega_U, \mathscr{P}(\Omega_U)\big)$, that is, the person variable. Furthermore, we consider the treatment variable $X\colon \Omega \to \Omega'_X = \{0,1\}$ with value space $\big(\Omega'_X, \mathscr{P}(\Omega'_X)\big)$ and the real-valued outcome variable Y. Finally, the treatment variable X takes on the value 1 for treatment and 0 for control. According to Definition 4.4, the filtration $(\mathscr{F}_t, t \in T)$ consists of three σ-algebras:

$$\mathscr{F}_1 := \sigma(h_1), \qquad \mathscr{F}_2 := \sigma(h_1, h_2), \qquad \mathscr{F}_3 := \sigma(h_1, h_2, h_3),$$

where $h_t\colon \Omega \to \Omega_t$, $t \in T = \{1,2,3\}$, denote the projections specified in Definition 4.4. In such a simple experiment, $h_1 = U$ and therefore, $\mathscr{F}_1 = \sigma(U)$, that is, the σ-algebra $\mathscr{F}_1$ of potential confounders of X is identical to the σ-algebra generated by the observational-unit variable U. Therefore, U is a global potential confounder of X. ◁

Remark 6.38 (True Outcome Variables) In all examples of this chapter, $P(X{=}x, U{=}u) > 0$ for all pairs (x,u) of values of X and U. Therefore

$$\tau_x := E^{X=x}(Y|C_X) = E^{X=x}(Y|U), \quad x \in \{0,1\}. \tag{6.54}$$

According to Equation (4.10), the true outcome variable τ_x can also be written as a function of the observational-unit variable U. More specifically, it can be written as the composition of U and the U-measurable function $g_x\colon \Omega_U \to \mathbb{R}$ defined by

$$g_x(u) = E(Y|X{=}x, U{=}u), \quad \text{for all } u \in \Omega_U. \tag{6.55}$$

Hence, $\tau_x = g_x(U)$ is the composition of U and g_x, that is,

$$\forall\, \omega \in \Omega\colon \quad \tau_x(\omega) = g_x\big(U(\omega)\big) = g_x(u), \quad \text{if } \omega \in \{U{=}u\}. \tag{6.56}$$

This implies that the values of the conditional expectation $E^{X=x}(Y|U)$ are identical to the conditional expectation values $E(Y|X{=}x, U{=}u)$ [see Eq. (4.9)]. ◁

Remark 6.39 (Individual Treatment Probabilities) In all examples of this chapter,

$$P(X{=}1|C_X) = P(X{=}1|U). \tag{6.57}$$

Note that, by definition, $P(X{=}1|C_X) = E(X|C_X)$ and $P(X{=}1|U) = E(X|U)$ (see Rem. 3.58), provided that X is dichotomous with values 0 and 1. In all examples of this chapter, this C_X-conditional treatment probability is uniquely defined and it is identical to the U-conditional treatment probability $P(X{=}1|U)$, whose values are denoted by $P(X{=}1|U{=}u)$ and also called the *individual treatment probabilities*. ◁

Box 6.2 Conditional expectation values in the three examples

Let $x \in X(\Omega)$ denote a value of the treatment variable X, let U denote the observational-unit variable in the examples presented in Tables 6.1 to 6.3, let $P(X{=}x, U{=}u) > 0$, for all values u of U, let $\mathscr{F}_1 = \sigma(U)$, and let $\tau_x = E^{X=x}(Y|U)$. Then:

$$E(Y|X{=}x) \;=\; E(\tau_x|X{=}x) \;=\; \sum_u E(Y|X{=}x, U{=}u)\cdot P(U{=}u|X{=}x), \tag{i}$$

whereas

$$E(\tau_x) \;=\; \sum_u E(Y|X{=}x, U{=}u)\cdot P(U{=}u). \tag{ii}$$

Additionally assume that Z is measurable with respect to U. Then $P(X{=}x, Z{=}z) > 0$ and

$$E(Y|X{=}x, Z{=}z) \;=\; E(\tau_x|X{=}x, Z{=}z) \;=\; \sum_u E(Y|X{=}x, U{=}u)\cdot P(U{=}u|X{=}x, Z{=}z), \tag{iii}$$

whereas

$$E(\tau_x|Z{=}z) \;=\; \sum_u E(Y|X{=}x, U{=}u)\cdot P(U{=}u|Z{=}z). \tag{iv}$$

6.5.2 Description of the Examples

In the first example (see Table 6.1), the treatment probabilities are *different for each and every unit*, and they strongly depend on the individual expectation values of the outcomes under control, that is, they depend on the true outcome variable τ_0 and also on τ_1. The $(X{=}x)$-conditional expectation values $E(Y|X=1)$ and $E(Y|X=0)$ are biased. In fact, the prima facie effect PFE_{10} is negative, whereas the causal average total effect ATE_{10} is positive. The $(X{=}x, Z{=}z)$-conditional expectation values $E(Y|X=1, Z{=}z)$ and $E(Y|X=0, Z{=}z)$ are biased as well. Although the causal $(Z{=}z)$-conditional total effects and the causal average total effect are defined (and can be computed from the fundamental parameters displayed in the upper left part of the table), they cannot be estimated from empirically estimable parameters such as the conditional expectation values $E(Y|X=1)$ and $E(Y|X=0)$ or $E(Y|X=1, Z{=}z)$ and $E(Y|X=0, Z{=}z)$.

In the second example (see Table 6.2), the treatment probabilities are *the same for all units*, that is, X and U are independent, which has many implications that are studied in detail in chapter 8. Among these implications are that the conditional expectation $E(Y|X)$ and its values $E(Y|X{=}x)$ as well as the conditional expectation $E(Y|X, Z)$ and its values $E(Y|X{=}x, Z{=}z)$ are unbiased.

In the third example (see Table 6.3), the treatment probabilities are *different between males and females*. Furthermore, these two subpopulations (sets of units) also differ in the $(Z{=}z)$-conditional expectation values of the true outcome variable τ_0, that is, $E(\tau_0|Z{=}m) \neq E(\tau_0|Z{=}f)$. *Given* a value z of Z, however, the treatment probabilities do not differ from each other. This implies that X and U are Z-conditionally independent, which in turn implies that $E(Y|X, Z)$ is unbiased. Hence, in this example, the conditional expectation values $E(Y|X{=}x)$ are biased, whereas the conditional expectation values $E(Y|X{=}x, Z{=}z)$ are unbiased. Again, this is studied extensively in chapter 8.

Tables 6.1, 6.2, and 6.3 display the true outcomes and the individual treatment probabilities. According to Equation (6.54), the true outcomes given x, that is, the values of τ_x,

Table 6.1. Self-selection of the unit to a treatment

		Fundamental parameters				Derived parameters		
Person variable U	Sex Z	$P(U=u)$	$P(X=1\,\vert\,U)$	$\tau_0 = E^{X=0}(Y\,\vert\,U)$	$\tau_1 = E^{X=1}(Y\,\vert\,U)$	$\delta_{10} = \tau_1 - \tau_0$	$P(U=u\,\vert\,X=0)$	$P(U=u\,\vert\,X=1)$
u_1	m	1/6	6/7	68	81	13	1/21	6/21
u_2	m	1/6	5/7	78	86	8	2/21	5/21
u_3	m	1/6	4/7	88	100	12	3/21	4/21
u_4	m	1/6	3/7	98	103	5	4/21	3/21
u_5	f	1/6	2/7	106	114	8	5/21	2/21
u_6	f	1/6	1/7	116	130	14	6/21	1/21

	$x=0$	$x=1$	
$E(\tau_x)$:	92.333	102.333	$ATE_{10} = 10$
$E(Y\,\vert\,X=x)$:	100.286	94.429	$PFE_{10} = -5.857$
$E(\tau_x\,\vert\,Z=m)$:	83	92.5	$CTE_{Z;10}(m) = 9.5$
$E(Y\,\vert\,X=x, Z=m)$:	88	90.278	$PFE_{Z;10}(m) = 2.278$
$E(\tau_x\,\vert\,Z=f)$:	111	122	$CTE_{Z;10}(f) = 11$
$E(Y\,\vert\,X=x, Z=f)$:	111.455	119.333	$PFE_{Z;10}(f) = 7.879$

are also the individual conditional expectation values $E(Y\,\vert\,X=x, U=u)$, and, according to Equation (6.57), the values of the conditional probability $P(X=1\,\vert\,C_X) = P(X=1\,\vert\,U)$, are identical to the individual treatment probabilities $P(X=1\,\vert\,U=u)$. The tables also display the values of the covariate $Z :=$ *sex*.

Some of the parameters appearing in these tables are called *fundamental parameters*. Other parameters are called *derived parameters* because they can be computed from the fundamental parameters. The individual total effects and the conditional probabilities $P(U=u\,\vert\,X=1)$ of observational unit u in treatment condition x, for instance, are such derived parameters. Note, however, that one may also consider the probabilities $P(U=u\,\vert\,X=1)$ as fundamental and the treatment probabilities $P(X=1\,\vert\,U=u)$ as derived. One can be computed as soon as the other one as well as the unconditional probabilities $P(U=u)$ and $P(X=1)$ are known.

Looking at the *fundamental parameters*, the three tables differ only in the treatment probabilities $P(X=1\,\vert\,U=u)$. All other entries, such as the true outcomes are the same. However, if we look at the *derived parameters*, the three tables differ in important aspects.

Table 6.2. Randomized assignment of the unit to a treatment

		Fundamental parameters				Derived parameters		
Person variable U	Sex Z	$P(U=u)$	$P(X=1\,\vert\,U)$	$\tau_0 = E^{X=0}(Y\,\vert\,U)$	$\tau_1 = E^{X=1}(Y\,\vert\,U)$	$\delta_{10} = \tau_1 - \tau_0$	$P(U=u\,\vert\,X=0)$	$P(U=u\,\vert\,X=1)$
u_1	m	1/6	3/4	68	81	13	1/6	1/6
u_2	m	1/6	3/4	78	86	8	1/6	1/6
u_3	m	1/6	3/4	88	100	12	1/6	1/6
u_4	m	1/6	3/4	98	103	5	1/6	1/6
u_5	f	1/6	3/4	106	114	8	1/6	1/6
u_6	f	1/6	3/4	116	130	14	1/6	1/6

	$x=0$	$x=1$	
$E(\tau_x)$:	92.333	102.333	$ATE_{10} = 10$
$E(Y\,\vert\,X{=}x)$:	92.333	102.333	$PFE_{10} = 10$
$E(\tau_x\,\vert\,Z{=}m)$:	83	92.5	$CTE_{Z;10}(m) = 9.5$
$E(Y\,\vert\,X{=}x, Z{=}m)$:	83	92.5	$PFE_{Z;10}(m) = 9.5$
$E(\tau_x\,\vert\,Z{=}f)$:	111	122	$CTE_{Z;10}(f) = 11$
$E(Y\,\vert\,X{=}x, Z{=}f)$:	111	122	$PFE_{Z;10}(f) = 11$

6.5.3 $(X{=}x)$-Conditional Expectation Values

We start computing the $(X{=}x)$-conditional expectation values of Y and check whether or not they are unbiased. In the first example (see Table 6.1), $E(Y\,\vert\,X=0) = 100.286$, whereas $E(\tau_0) = 92.333$, that is, $E(Y\,\vert\,X=0)$ is much larger than $E(\tau_0)$. In contrast, $E(Y\,\vert\,X=1) = 94.429$, whereas $E(\tau_1) = 102.333$, that is, $E(Y\,\vert\,X=1)$ is much smaller than $E(\tau_0)$. Hence, according to Definition 6.2 (i), the conditional expectation values $E(Y\,\vert\,X{=}x)$ are biased, and according Definition 6.2 (ii) this is also true for the conditional expectation $E(Y\,\vert\,X)$.

These expectations and conditional expectations are easy to compute from the parameters displayed in Table 6.1. The expectations $E(\tau_x)$ of the two true outcome variables are obtained from taking the expectations of the true outcome variables using the *unconditional probabilities* $P(U{=}u)$ as weights [see Box 6.2 (ii)]. In contrast, the corresponding conditional expectations $E(Y\,\vert\,X{=}x)$ are identical to the $(X{=}x)$-conditional expectation values of the true outcome variables, using as weights the *conditional probabilities* $P(U{=}u\,\vert\,X{=}x)$ [see Box 6.2 (i)].

If used for the evaluation of the total treatment effect, the $(X{=}x)$-conditional expectation values would lead to completely wrong conclusions. Not only is the direction of the prima facie effect

$$E(Y\,\vert\,X=1) \;-\; E(Y\,\vert\,X=0) \;=\; -5.857$$

Table 6.3. Conditionally randomized assignment of the unit to a treatment

		Fundamental parameters				Derived parameters		
Person variable U	Sex Z	$P(U=u)$	$P(X=1\,\vert\,U)$	$\tau_0 = E^{X=0}(Y\,\vert\,U)$	$\tau_1 = E^{X=1}(Y\,\vert\,U)$	$\delta_{10} = \tau_1 - \tau_0$	$P(U=u\,\vert\,X=0)$	$P(U=u\,\vert\,X=1)$
u_1	m	1/6	3/4	68	81	13	1/10	3/14
u_2	m	1/6	3/4	78	86	8	1/10	3/14
u_3	m	1/6	3/4	88	100	12	1/10	3/14
u_4	m	1/6	3/4	98	103	5	1/10	3/14
u_5	f	1/6	1/4	106	114	8	3/10	1/14
u_6	f	1/6	1/4	116	130	14	3/10	1/14

	$x=0$	$x=1$	
$E(\tau_x)$:	92.333	102.333	$ATE_{10} = 10$
$E(Y\,\vert\,X{=}x)$:	99.8	96.714	$PFE_{10} = -3.086$
$E(\tau_x\,\vert\,Z{=}m)$:	83	92.5	$CTE_{Z;10}(m) = 9.5$
$E(Y\,\vert\,X{=}x, Z{=}m)$:	83	92.5	$PFE_{Z;10}(m) = 9.5$
$E(\tau_x\,\vert\,Z{=}f)$:	111	122	$CTE_{Z;10}(f) = 11$
$E(Y\,\vert\,X{=}x, Z{=}f)$:	111	122	$PFE_{Z;10}(f) = 11$

reversed as compared to the difference

$$E(\tau_1) \;-\; E(\tau_0) \;=\; 10,$$

but also as compared *to each and every individual total effect* (see the column $\tau_1 - \tau_0$ in Table 6.1). All individual total effects are positive in this example, ranging between 5 and 14. The bias in this example is due to strong inter-individual differences in the true outcomes and to the fact that the individual treatment probabilities $P(X{=}1\,\vert\,U{=}u)$ heavily depend on the true outcome variables, and, therefore, on the person variable U. For instance, unit u_1 has a true outcome under control of 68 and a treatment probability of 6/7, while unit u_6 has a true outcome under control of 116 and a treatment probability of 1/7. Such a constellation is to be expected under self-selection of subjects to treatments, if the subjects base their decisions to take treatment on the *severity of their dysfunction before treatment* and if *severity of their dysfunction after treatment* is assessed as the outcome variable.

For the second example presented in Table 6.2, the situation is completely different. Although the true outcome variables are the same as in Table 6.1, here, the conditional expectation values $E(Y\,\vert\,X{=}x)$ and the expectations $E(\tau_x)$ of the true outcome variables are identical to each other, and this applies to both values 0 and 1 of X. Hence, in this example, the conditional expectation values $E(Y\,\vert\,X{=}x)$ are unbiased and can be used for the evaluation of the treatment effect. This is due to the fact that the individual treatment

probabilities *do not depend on the units*. This constellation occurs in a perfect randomized experiment, in which the experimenter decides that each subject is in treatment 1 with probability $P(X=1)$ and in treatment 0 with probability $1-P(X=1)$. In our second example, $P(X=1)=3/4$. Note, however, that $P(X=1)$ could be *any* number between 0 and 1, exclusively. The only important point is that the individual treatment probabilities *do not differ between units*, that is, $P(X=1\,|\,U=u)=P(X=1)$ for all units $u\in\Omega_U$. Such a randomized assignment may be performed by drawing a ball from an urn with three black balls and one white ball, adopting the rule that the subject is treated if a black ball is drawn.

In the third example (see Table 6.3), the conditional expectation values $E(Y\,|\,X=x)$ are biased again. Here, $E(Y\,|\,X=0)=99.8$, whereas $E(\tau_0)=92.333$. Again, $E(Y\,|\,X=0)$ is much larger than $E(\tau_0)$. In contrast, $E(Y\,|\,X=1)=96.714$, whereas $E(\tau_1)=102.333$, that is, again $E(Y\,|\,X=1)$ is much smaller than $E(\tau_0)$. Hence, in this example, the conditional expectation values $E(Y\,|\,X=x)$ are strongly biased as well. However, in contrast to the first example, the conditional expectation values $E(Y\,|\,X=x,Z=z)$ are unbiased. In this example, the treatment probability is $3/4$ *for all male units*, while it is $1/4$ *for all female units*. In this example, the crucial point is that these probabilities are the same given a value z of the covariate Z, that is, $P(X=1\,|\,Z=z,U=u)=P(X=1\,|\,Z=z)$ for each unit u and both values z of the covariate Z. This constellation holds in a perfect conditionally randomized experiment in which we assign the sampled person to treatment with probability $P(X=1\,|\,Z=m)$ if he is male and with probability $P(X=1\,|\,Z=f)$ if the sampled person is female.

6.5.4 $(X=x,Z=z)$-Conditional Expectation Values

The conditional expectation values $E(Y\,|\,X=x,Z=z)$, can be computed from the parameters displayed in Table 6.3, applying Equation (iii) of Box 6.2. For this purpose we also need the formula

$$P(U=u\,|\,X=x,Z=z) \;=\; \frac{P(X=x\,|\,U=u)\cdot P(U=u,Z=z)}{P(X=x\,|\,Z=z)\cdot P(Z=z)}, \tag{6.58}$$

where

$$P(X=x\,|\,Z=z) \;=\; \frac{\sum\limits_{u} P(X=x\,|\,U=u)\cdot P(U=u,Z=z)}{P(Z=z)} \tag{6.59}$$

(see Exercise 6-10). Note that in the three examples $P(X=x\,|\,U=u,Z=z)=P(X=x\,|\,U=u)$, because in theses examples Z is U-measurable. Intuitively speaking, this means that Z (the sex variable) does not contain any information that is not already contained in U (the person variable). All terms on the right-hand side of Equation (6.58) are displayed in Table 6.3 or can be computed from the parameters displayed in this table. [2]

Remark 6.40 (How Realistic are These Examples?) In empirical applications, assuming $\mathscr{F}_1=\sigma(U)$ is correct if (a) there is neither a second treatment variable nor another variable that is simultaneous to X and if (b) no fallible covariate is observed. In this case, u signifies the observational unit *at the onset of treatment*. If, however, a fallible covariate of X is observed and u represents the observational unit *at the time at which the covariate is assessed*, then there may very well be covariates that are not measurable with respect to U, which affect the outcome variable Y and/or the treatment probability (see section 2.2).

[2] An alternative is using the Causal Effects Xplorer provided at www.causal-effects.de, the home page of this book.

Hence, in this case, $\mathscr{F}_1 = \sigma(U)$ would not hold. In this case $\mathscr{F}_1$ would be the σ-algebra generated by U and the fallible covariates to be assessed. ◁

6.5.5 Conditional Total Effects

Comparing the conditional prima facie effects to the conditional total effects reveals that the conditional prima facie effects are still biased with respect to total effects in the random experiment presented in Table 6.1, but not in the examples displayed in Tables 6.2 and 6.3. Hence, in the first example, the conditional prima facie effect and the conditional total effect for males are *not* identical, while they *are* identical in the second and third examples, and the same applies to the corresponding prima facie effects for the females.

The bias of the conditional prima facie effects in the example presented in Table 6.1 is no surprise, because there are still individual differences within the two sex subpopulations with respect to (a) the true outcomes under treatment and under control, as well as (b) in the individual treatment probabilities $P(X{=}1 \mid U{=}u)$. In contrast, in the second and third examples, the individual treatment probabilities are all the same *within each of the two sex subpopulations*.

6.5.6 Computing the Causal Average Total Effect From Conditional Total Effects

In all three examples, the average over the individual total effects is equal to the causal average total effect. However, only in the second and third examples, the expectation of the *sex*-conditional *prima facie effects* is equal to the causal average total effect. [Remember, the causal average total effect is defined as the expectation of true total effect variable δ_{10}.] Because this is no coincidence, this fact can be used for causal inference even in those cases in which the *unconditional* prima facie effects are biased, provided that the *conditional* prima facie effects are unbiased, that is, provided that $PFE_{Z;10}(z) = E(\delta_{10} \mid Z{=}z)$ for each value z of the covariate Z.

Whether or not a causal average total effect is meaningful if there are different conditional total effects — some of which may even be negative, while some are positive — needs judgement in the specific applications considered. In some applications it might be meaningful, in others it might not. Clearly, causal conditional total effects give more specific information than the causal average total effect. However, there are also advantages of causal average total effects. First, they give a brief summary evaluation of a treatment in *a single number* and different treatments may be compared to each other with respect to this number. Second, in samples of limited size, causal average effects can be estimated with more accuracy than the plenitude of causal conditional effects. And third, one should keep in mind that even conditional effects are only causal average effects (see, e. g., Table 6.3). Hence, it is always a matter of substantive judgement how fine-grained the analysis should be.

6.5.7 First Conclusions

The three examples show that conditioning on a covariate of X does not *necessarily* yield unbiasedness given the values of the covariate. While there is no bias at all in the second example, the third example shows that conditioning *may* remove bias. Comparing Examples 2 and 3 to each other shows that unbiasedness of the conditional expectation val-

ues $E(Y|X=x,Z=z)$ relies on specific conditions [here: $P(X=1|U) = P(X=1|Z)$, that is, equal individual treatment probabilities for units with an identical value z of Z] in a similar way as unbiasedness of $E(Y|X=x)$. Such conditions are called *causality conditions*. Note, however, that there are several of such causality conditions that do *not* involve the U-conditional treatment probabilities (see ch. 9).

6.6 Example With Accidental Unbiasedness

Now we treat an example showing that there can be *unbiasedness of the conditional expectation* $E(Y|X)$ and at the same time *bias of the conditional expectation* $E(Y|X,Z)$. This example shows that unbiasedness can be accidental, that is, there are cases in which unbiasedness is not a logical consequence of experimental design but an 'accident of numbers'. In chapter 8 we show that the experimental design technique of randomization *always* induces unbiasedness of the conditional expectations $E(Y|X)$ and and $E(Y|X,Z)$ for all covariates Z of X. [3]

Conditional Expectation Values $E(Y|X=x)$

Table 6.4 displays the relevant parameters. We assume that it is a simple experiment so that $C_X = U$. In this specific example, the causal individual total effects are the same for all units, namely 5, implying that the causal average total effect is also 5. The prima facie effect can be computed from the difference between the two conditional expectation values $E(Y|X=0)$ and $E(Y|X=1)$. In this example, Box 6.2 (ii) yields

$$\begin{aligned} E(Y|X=0) &= \sum_u E(Y|X=0,U=u)\cdot P(U=u|X=0) \\ &= 95\cdot\frac{3}{16}+65\cdot\frac{1}{16}+80\cdot\frac{7}{16}+50\cdot\frac{5}{16} = 72.5 \end{aligned}$$

and

$$\begin{aligned} E(Y|X=1) &= \sum_u E(Y|X=1,U=u)\cdot P(U=u|X=1) \\ &= 100\cdot\frac{5}{16}+70\cdot\frac{7}{16}+85\cdot\frac{1}{16}+55\cdot\frac{3}{16} = 77.5. \end{aligned}$$

Using Equation (ii) of Box 6.2, the corresponding expectations of the true outcome variables are

$$\begin{aligned} E(\tau_0) &= \sum_u E(Y|X=0,U=u)\cdot P(U=u) \\ &= 95\cdot\frac{1}{4}+65\cdot\frac{1}{4}+80\cdot\frac{1}{4}+50\cdot\frac{1}{4} = 72.5 \end{aligned}$$

and

$$\begin{aligned} E(\tau_1) &= \sum_u E(Y|X=1,U=u)\cdot P(U=u) \\ &= 100\cdot\frac{1}{4}+70\cdot\frac{1}{4}+85\cdot\frac{1}{4}+55\cdot\frac{1}{4} = 77.5. \end{aligned}$$

[3] Note that unbiasedness does not refer to a sample and that there is no (successful) randomization if there is systematic attrition.

Table 6.4. Accidental unbiasedness

		Fundamental parameters				Derived parameters		
Person variable U	Sex Z	$P(U=u)$	$P(X=1\,\vert\,U)$	$\tau_0 = E^{X=0}(Y\vert U)$	$\tau_1 = E^{X=1}(Y\vert U)$	$\delta_{10} = \tau_1 - \tau_0$	$P(U=u\vert X=0)$	$P(U=u\vert X=1)$
u_1	m	1/4	5/8	95	100	5	3/16	5/16
u_2	m	1/4	7/8	65	70	5	1/16	7/16
u_3	f	1/4	1/8	80	85	5	7/16	1/16
u_4	f	1/4	3/8	50	55	5	5/16	3/16

	$x=0$	$x=1$	
$E(\tau_x)$:	72.5	77.5	$ATE_{10} = 5$
$E(Y\vert X{=}x)$:	72.5	77.5	$PFE_{10} = 5$
$E(\tau_x\vert Z{=}m)$:	80	85	$CTE_{Z;10}(m) = 5$
$E(Y\vert X{=}x, Z{=}m)$:	87.5	82.5	$PFE_{Z;10}(m) = -5$
$E(\tau_x\vert Z{=}f)$:	65	70	$CTE_{Z;10}(f) = 5$
$E(Y\vert X{=}x, Z{=}f)$:	67.5	62.5	$PFE_{Z;10}(f) = -5$

Hence, the conditional expectation values $E(Y\vert X=0)$ and $E(Y\vert X=1)$ are unbiased, because they are identical to the corresponding expectations $E(\tau_0)$ and $E(\tau_1)$ of the true outcome variables.

Conditional Expectation Values $E(Y\vert X{=}x, Z{=}z)$

The conditional expectation values $E(Y\vert X{=}1, Z{=}z)$ and $E(Y\vert X{=}0, Z{=}z)$ can be computed from the parameters displayed in Table 6.4 using Box 6.2 (iii). This equation holds, because, in this example, the random variable Z is measurable with respect to U (see SN-Cor. 2.53). While the individual expected outcomes $E(Y\vert X{=}x, U{=}u)$ are displayed in Table 6.4, the conditional probabilities $P(U{=}u\vert X{=}x, Z{=}z)$ have to be computed via Equation (6.58) (see Exercise 6-10).

For $Z{=}m$ (males), Equation (iii) of Box 6.2 yields

$$\begin{aligned} E(Y\vert X{=}0, Z{=}m) &= \sum_u E(Y\vert X{=}0, U{=}u)\cdot P(U{=}u\vert X{=}0, Z{=}m) \\ &= 95\cdot\frac{9}{12} + 65\cdot\frac{3}{12} + 80\cdot 0 + 50\cdot 0 = 87.5 \end{aligned}$$

and

$$\begin{aligned} E(Y\vert X{=}1, Z{=}m) &= \sum_u E(Y\vert X{=}1, U{=}u)\cdot P(U{=}u\vert X{=}1, Z{=}m) \\ &= 100\cdot\frac{5}{12} + 70\cdot\frac{7}{12} + 85\cdot 0 + 55\cdot 0 = 82.5. \end{aligned}$$

In contrast, using Equation (iv) of Box 6.2, the $(Z=m)$-conditional expectation values of the true outcome variables are

$$\begin{aligned} E(\tau_0|Z=m) &= \sum_u E(Y|X=0,U=u)\cdot P(U=u|Z=m) \\ &= 95\cdot\frac{1}{2}+65\cdot\frac{1}{2} + 80\cdot 0+50\cdot 0 = 80 \end{aligned}$$

and

$$\begin{aligned} E(\tau_1|Z=m) &= \sum_u E(Y|X=1,U=u)\cdot P(U=u|Z=m) \\ &= 100\cdot\frac{1}{2}+70\cdot\frac{1}{2} + 85\cdot 0 + 55\cdot 0 = 85. \end{aligned}$$

For $Z=f$ (females), Equation (iii) of Box 6.2 yields

$$\begin{aligned} E(Y|X=0,Z=f) &= \sum_u E(Y|X=0,U=u)\cdot P(U=u|X=0,Z=f) \\ &= 95\cdot 0+65\cdot 0 + 80\cdot 7/12+50\cdot 5/12 = 67.5 \end{aligned}$$

and

$$\begin{aligned} E(Y|X=1,Z=f) &= \sum_u E(Y|X=1,U=u)\cdot P(U=u|X=1,Z=m) \\ &= 100\cdot 0+70\cdot 0 + 85\cdot 3/12+55\cdot 9/12 = 62.5. \end{aligned}$$

In contrast, using Equation (iv) of Box 6.2, the $(Z=f)$-conditional expectation values of the true outcome variables are

$$\begin{aligned} E(\tau_0|Z=f) &= \sum_u E(Y|X=0,U=u)\cdot P(U=u|Z=f) \\ &= 95\cdot 0+65\cdot 0 + 80\cdot\frac{1}{2}+50\cdot\frac{1}{2} = 65 \end{aligned}$$

and

$$\begin{aligned} E(\tau_1|Z=f) &= \sum_u E(Y|X=1,U=u)\cdot P(U=u|Z=f) \\ &= 100\cdot 0+70\cdot 0 + 85\cdot\frac{1}{2} + 55\cdot\frac{1}{2} = 70. \end{aligned}$$

Obviously, the conditional expectation values $E(Y|X=x,Z=z)$ of the outcome variable Y are not identical to the corresponding conditional expectation values $E(\tau_x|Z=z)$ of the true outcome variables. Hence, the $E(Y|X=x,Z=z)$ are biased although the conditional expectation values $E(Y|X=x)$ are unbiased.

Remark 6.41 (Methodological Implications) This example shows that unbiasedness of the conditional expectation values $E(Y|X=x)$ does not imply unbiasedness of the conditional expectation values $E(Y|X=x,Z=z)$, even if Z is a covariate of X. Hence, this example shows that *unbiasedness can be accidental*, that is, it may be a fortunate coincidence, an 'accident of numbers', not a logical consequence of experimental design. In chapter 8, however, we will show that experimental design techniques such as randomized assignment of the unit to one of the treatment conditions *always* leads to unbiasedness and to $(Z=z)$-conditional unbiasedness if Z denotes a covariate of X. If Z is measurable with respect to the observational-unit variable, then this implies that randomization always leads to unbiasedness of the conditional expectation values $E(Y|X=x)$ and

$E(Y|X=x,Z=z)$. Note, however, that this beneficial implication of randomization *only applies to unbiasedness with respect to total effects*. Unfortunately, it does not apply to unbiasedness with respect to direct effects (see, e. g., Mayer, Thoemmes, Rose, Steyer, & West, 2014). ◁

6.7 Summary and Conclusions

Unbiasedness is a first kind of causality conditions, which, together with the additional structural components listed in a causality space, distinguishes a causally interpretable conditional expectation from an ordinary conditional expectation. Several kinds of conditional expectations and their values as well as their differences can be unbiased (see Box 6.1). The general insight of this chapter is that comparing conditional expectation values (true means) does not allow to draw any conclusions on the effects of a treatment or intervention *unless they are unbiased*. In terms of the metaphor discussed in the preface, conditional expectation values and their differences are like the shadow of the invisible man. The length of this shadow is identical to the height of the invisible man only under very specific conditions. The unbiasedness conditions are the weakest assumptions under which we can identify causal average total effects and causal conditional total effects. All other causality conditions imply unbiasedness, some of which, in contrast to unbiasedness itself, *are* empirically testable, at least in the sense of falsifiability.

Hence, a first limitation of unbiasedness is that it cannot be tested empirically. Another drawback of unbiasedness has been exemplified by the numerical example displayed in Table 6.4. This example shows that, even if Z is a covariate of X, the conditional expectation values $E(Y|X=x,Z=z)$ and the $(Z=z)$-conditional prima facie effects can be biased even in cases in which the conditional expectation values $E(Y|X=x)$ are unbiased (see also Greenland & Robins, 1986). In contrast, the sufficient conditions of unbiasedness treated in chapters 8 and 9 are less volatile, that is, they generalize to conditioning on a covariate Z of X. *Generalizability* and *empirical testability* are two important virtues of these alternative causality conditions.

6.8 Proofs

Proof of Theorem 6.9

Equation (6.7).

$$\begin{aligned} E(Y|X=x) &= E^{X=x}(Y) && [(3.28)] \\ &= E^{X=x}\big(E^{X=x}(Y|C_X)\big) && [\text{SN-Box 10.2 (iv)}] \\ &= E^{X=x}(\tau_x). && [\tau_x \in \mathscr{E}^{X=x}(Y|C_X)] \end{aligned}$$

Hence, Equations (6.2) and (6.7) are equivalent to each other.

Equation (6.8). First, it is shown that Equation (6.8) implies Equation (6.7). Our assumptions include positive probabilities $P(X=x)$ for all values x of X and this implies $0 < P(1_{X=x}=1), P(1_{X=x}=0) < 1$. Hence, according to SN-Corollary 10.39 (i), Equation (6.8) is equivalent to

$$E(\tau_x|1_{X=x}=0) \;=\; E(\tau_x|1_{X=x}=1) \;=\; E(\tau_x). \tag{6.60}$$

Using this result yields

$$\begin{aligned} E(\tau_x) &= E(\tau_x \mid \mathbb{1}_{X=x}=1) && [(6.60)] \\ &= E(\tau_x \mid X=x) && [\{\mathbb{1}_{X=x}=1\} = \{X=x\}] \\ &= E^{X=x}(\tau_x). && [(3.28)] \end{aligned}$$

Now we show that Equation (6.8) follows from Equation (6.7).

$$\begin{aligned} & E(\tau_x) \\ &= E(\tau_x \mid \mathbb{1}_{X=x}=1) \cdot P(\mathbb{1}_{X=x}=1) \;+\; E(\tau_x \mid \mathbb{1}_{X=x}=0) \cdot P(\mathbb{1}_{X=x}=0) && [\text{SN-(9.23)}] \\ &= E(\tau_x) \cdot P(\mathbb{1}_{X=x}=1) \;+\; E(\tau_x \mid \mathbb{1}_{X=x}=0) \cdot P(\mathbb{1}_{X=x}=0) && [(6.7)] \\ &= E(\tau_x) \cdot \big(1 - P(\mathbb{1}_{X=x}=0)\big) \;+\; E(\tau_x \mid \mathbb{1}_{X=x}=0) \cdot P(\mathbb{1}_{X=x}=0) && [P(\mathbb{1}_{X=x}=1) = 1 - P(\mathbb{1}_{X=x}=0)] \\ &= E(\tau_x) \;+\; \big(E(\tau_x \mid \mathbb{1}_{X=x}=0) - E(\tau_x)\big) \cdot P(\mathbb{1}_{X=x}=0). \end{aligned}$$

Because we assume $P(\mathbb{1}_{X=x}=0) > 0$, this equation can only hold if $E(\tau_x \mid \mathbb{1}_{X=x}=0) = E(\tau_x)$. Because Equation (6.7) implies $E(\tau_x \mid \mathbb{1}_{X=x}=1) = E(\tau_x)$, this yields Equation (6.60), which, according to Corollary 10.39 (i) is equivalent to Equation (6.8).

Equation (6.9). The definition $\varepsilon_x = \tau_x - E^{X=x}(Y)$ yields

$$E^{X=x}(\varepsilon_x) \;=\; 0, \tag{6.61}$$

which can be shown as follows:

$$\begin{aligned} E^{X=x}(\varepsilon_x) &= E^{X=x}\big(\tau_x - E^{X=x}(Y)\big) && [(6.9)] \\ &= E^{X=x}(\tau_x) - E^{X=x}\big(E^{X=x}(Y)\big) && [\text{SN-Box 6.1 (vi)}] \\ &= E^{X=x}\big(E^{X=x}(Y \mid C_X)\big) - E^{X=x}\big(E^{X=x}(Y)\big) && [\tau_x \in \mathscr{E}^{X=x}(Y \mid C_X)] \\ &= E^{X=x}(Y) - E^{X=x}(Y) && [\text{SN-Box 10.2 (iv), SN-Box 6.1 (i)}] \\ &= 0. \end{aligned}$$

According to Equation (6.9), $\tau_x \underset{P^{X=x}}{=} E^{X=x}(Y) + \varepsilon_x$. Hence,

$$\begin{aligned} E^{X=x}(\tau_x) &= E^{X=x}\big(E^{X=x}(Y) + \varepsilon_x\big) && [(6.9)] \\ &= E^{X=x}\big(E^{X=x}(Y)\big) + E^{X=x}\big(\varepsilon_x\big) && [\text{SN-Box 6.1 (vi)}] \\ &= E^{X=x}(Y) && [\text{SN-Box 6.1 (i)}, (6.61)] \\ &= E^{X=x}(Y) + E(\varepsilon_x) && [(6.9)] \\ &= E\big(E^{X=x}(Y) + \varepsilon_x\big) && [\text{SN-Box 6.1 (i)}] \\ &= E(\tau_x). && [(6.9)] \end{aligned}$$

This proves that (6.9) implies (6.7).

Now we show that Equation (6.7) implies (6.9). According to (6.7), $E^{X=x}(\tau_x) = E(\tau_x)$. Hence,

$$\begin{aligned} 0 &= E(\tau_x) - E(\tau_x) \\ &= E(\tau_x) - E^{X=x}(\tau_x) && [(6.7)] \\ &= E(\tau_x) - E\big(E^{X=x}(\tau_x)\big) && [\text{SN-Box 6.1 (i)}] \\ &= E\big(\tau_x - E^{X=x}(\tau_x)\big) && [\text{SN-Box 6.1 (vi)}] \\ &= E\big(\tau_x - E^{X=x}(Y)\big) && [\text{SN-Box 10.2 (iv)}] \\ &= E(\varepsilon_x). && [\text{def. } \varepsilon_x] \end{aligned}$$

Proof of Lemma 6.14

$$\begin{aligned}
\varepsilon_x &\underset{P^{X=x}}{=} E^{X=x}(Y|C_X) - E^{X=x}(Y|Z) && [(6.17)]\\
&\underset{P^{X=x}}{=} E^{X=x}(Y|C_X) - E^{X=x}\big(E^{X=x}(Y|C_X)\,\big|\,Z\big) && [\text{SN-Box 10.1 (v)}]\\
&\underset{P^{X=x}}{=} \tau_x - E^{X=x}(\tau_x|Z). && [\tau_x \in \mathscr{E}^{X=x}(Y|C_X)]
\end{aligned}$$

These equations show that ε_x is a residual with respect to a Z-conditional expectation of τ_x with respect to the measure $P^{X=x}$. According to SN-Box 11.1 (vi), this implies Equation (6.19).

Proof of Theorem 6.15

Because $\sigma(Z) \subset \sigma(C_X)$, according to SN-Box 14.1 (iv), P-uniqueness of $\tau_x = E^{X=x}(Y|C_X)$ implies that $E^{X=x}(Y|Z)$ is P-unique as well. Therefore, according to SN-Corollary 14.48 (a) and (c), this is equivalent to $P(X{=}x\,|\,Z) \gtrsim_P 0$, and to P-uniqueness of $E^{X=x}(\tau_x|Z)$, and $E^{X=x}(\varepsilon_x|Z)$.

Equation (6.20). Because $\sigma(Z) \subset \sigma(C_X)$, according to SN-Box 10.1 (v),

$$E^{X=x}(\tau_x|Z) \underset{P^{X=x}}{=} E^{X=x}\big(E^{X=x}(Y|C_X)\,\big|\,Z\big) \underset{P^{X=x}}{=} E^{X=x}(Y|Z). \tag{6.62}$$

Therefore, P-uniqueness of $E^{X=x}(Y|Z)$ implies

$$E^{X=x}(Y|Z) \underset{P}{=} E^{X=x}(\tau_x|Z).$$

Hence, Equations (6.16) and (6.20) are equivalent to each other.

Equation (6.21). First, it is shown that Equation (6.21) implies Equation (6.20). According to SN-Corollary 10.39 (i), Equation (6.21) is equivalent to

$$E(\tau_x|1_{X=x}{=}v, Z{=}z) \;=\; E(\tau_x|Z{=}z), \quad \text{for } P_{1_{X=x},Z}\text{-a.a. } (v,z) \in \{0,1\}\times\Omega'_Z\,.$$

Because $E(\tau_x|1_{X=x}{=}1, Z{=}z) = E(\tau_x|X{=}x, Z{=}z)$, this implies

$$E(\tau_x|X{=}x, Z{=}z) \;=\; E(\tau_x|Z{=}z), \quad \text{for } P_{1_{X=x},Z}\text{-a.a. } (1,z) \in \{1\}\times\Omega'_Z,$$

and according to SN-Corollary 14.58, this equation in turn implies

$$E^{X=x}(\tau_x|Z{=}z) \;=\; E(\tau_x|Z{=}z), \quad \text{for } P_Z\text{-a.a. } z \in \Omega'_Z.$$

According to SN-Corollary 10.39 (i), this is equivalent to Equation (6.20).

Now we show that Equation (6.21) follows from Equation (6.20).

$$\begin{aligned}
&E(\tau_x|Z)\\
&\underset{P}{=} E^{1_{X=x}=1}(\tau_x|Z)\cdot 1_{1_{X=x}=1} + E^{1_{X=x}=0}(\tau_x|Z)\cdot 1_{1_{X=x}=0} && [\text{SN-Rem. 14.34}]\\
&\underset{P}{=} E^{X=x}(\tau_x|Z)\cdot 1_{1_{X=x}=1} + E^{1_{X=x}=0}(\tau_x|Z)\cdot 1_{1_{X=x}=0} && [\{X{=}x\} = \{1_{X=x}{=}1\}]\\
&\underset{P}{=} E(\tau_x|Z)\cdot 1_{1_{X=x}=1} + E^{1_{X=x}=0}(\tau_x|Z)\cdot 1_{1_{X=x}=0} && [(6.20)]\\
&\underset{P}{=} E(\tau_x|Z)\cdot(1-1_{1_{X=x}=0}) + E^{1_{X=x}=0}(\tau_x|Z)\cdot 1_{1_{X=x}=0}\\
&\underset{P}{=} E(\tau_x|Z) + \big(E^{1_{X=x}=0}(\tau_x|Z) - E(\tau_x|Z)\big)\cdot 1_{1_{X=x}=0}.
\end{aligned}$$

Because we assume $P(1_{X=x}=0) > 0$, this equation can only hold if

$$E^{1_{X=x}=0}(\tau_x \mid Z) \underset{P}{=} E(\tau_x \mid Z). \tag{6.63}$$

Note, we assume $E^{X=x}(\tau_x \mid Z) \underset{P}{=} E(\tau_x \mid Z)$. Furthermore, $E^{X=x}(\tau_x \mid Z) \underset{P}{=} E^{1_{X=x}=1}(\tau_x \mid Z)$. Therefore,

$$E^{1_{X=x}=1}(\tau_x \mid Z) \underset{P}{=} E(\tau_x \mid Z). \tag{6.64}$$

Using (6.63) and (6.64), Corollary 10.39 (i) implies Equation (6.21).

Equation (6.22). Because $\sigma(Z) \subset \sigma(C_X)$, according to SN-Box 14.1 (iv), assuming P-uniqueness of τ_x implies P-uniqueness of the conditional expectations $E^{X=x}(Y \mid Z)$, $E^{X=x}(\tau_x \mid Z)$, and $E^{X=x}(\varepsilon_x \mid Z)$. Because (the constant) 0 is Z-measurable, Equation (6.19) yields $0 \in \mathcal{E}^{X=x}(\varepsilon_x \mid Z)$. Hence, P-uniqueness of $E^{X=x}(\varepsilon_x \mid Z)$ implies

$$E^{X=x}(\varepsilon_x \mid Z) \underset{P}{=} 0. \tag{6.65}$$

Therefore,

$$\begin{aligned}
& E^{X=x}(\tau_x \mid Z) \\
\underset{P}{=}\ & E^{X=x}\big(E^{X=x}(Y \mid Z) + \varepsilon_x \,\big|\, Z\big) && \text{[(6.22)]} \\
\underset{P}{=}\ & E^{X=x}\big(E^{X=x}(Y \mid Z) \,\big|\, Z\big) + E^{X=x}(\varepsilon_x \mid Z) && \text{[SN-Box 10.2 (xv)]} \\
\underset{P}{=}\ & E^{X=x}(Y \mid Z) && \text{[SN-Box 10.2 (vii), (6.65)]} \\
\underset{P}{=}\ & E^{X=x}(Y \mid Z) + E(\varepsilon_x \mid Z) && \text{[(6.22)]} \\
\underset{P}{=}\ & E\big(E^{X=x}(Y \mid Z) + \varepsilon_x \,\big|\, Z\big) && \text{[SN-Box 10.2 (vii), (xv)]} \\
\underset{P}{=}\ & E(\tau_x \mid Z). && \text{[(6.22)]}
\end{aligned}$$

This proves that Equation (6.20) implies (6.22).

Now we show that (6.22) implies (6.20).

$$\begin{aligned}
0 \underset{P}{=}\ & E(\varepsilon_x \mid Z) \underset{P}{=} E\big(\tau_x - E^{X=x}(Y \mid Z) \,\big|\, Z\big) && \text{[(6.22)]} \\
\underset{P}{=}\ & E(\tau_x \mid Z) - E\big(E^{X=x}(Y \mid Z) \,\big|\, Z\big) && \text{[SN-Box 10.2 (xv)]} \\
\underset{P}{=}\ & E(\tau_x \mid Z) - E\Big(E^{X=x}\big(E^{X=x}(Y \mid C_X) \,\big|\, Z\big) \,\Big|\, Z\Big) && \text{[SN-Box 10.2 (v)]} \\
\underset{P}{=}\ & E(\tau_x \mid Z) - E\big(E^{X=x}(\tau_x \mid Z) \,\big|\, Z\big) && [\tau_x := E^{X=x}(Y \mid C_X)] \\
\underset{P}{=}\ & E(\tau_x \mid Z) - E^{X=x}(\tau_x \mid Z) && \text{[SN-Box 10.2 (vii)]}
\end{aligned}$$

However, the equation

$$0 \underset{P}{=} E(\tau_x \mid Z) - E^{X=x}(\tau_x \mid Z)$$

can only hold if $E^{X=x}(\tau_x \mid Z) \underset{P}{=} E(\tau_x \mid Z)$.

Proof of Theorem 6.17

If τ_x is $P^{Z=z}$-unique, then it is also $P^{X=x,Z=z}$-unique [see SN-Box 14.1 (v)]. According to SN-Corollary 14.48 (e), this implies that the conditional expectation values on both sides of Equation (6.26) are uniquely defined. Note that we assume that Z is a covariate of X. According to Definition 4.4 (i) and Remark 4.9 this implies that there is a C_X-measurable mapping $g\colon \Omega'_{C_X} \to \Omega'_Z$ such that $Z = g(C_X)$.

First we show that unbiasedness of $E(Y \mid X{=}x, Z{=}z)$ implies Equation (6.26).

$$E^{X=x}(\tau_x \mid Z{=}z)$$

$$
\begin{aligned}
&= E^{X=x}\big(E^{X=x}(Y|C_X)\,\big|\,Z{=}z\big) && [\tau_x = E^{X=x}(Y|C_X)]\\
&= E^{X=x}(Y\,|\,Z{=}z) && [Z = g(C_X), \text{SN-(10.39)}]\\
&= E(Y|X{=}x, Z{=}z) && [(6.24)]\\
&= E(\tau_x|Z{=}z). && [(6.15)]
\end{aligned}
$$

Now we show that Equation (6.26) implies unbiasedness of $E(Y|X{=}x, Z{=}z)$.

$$
\begin{aligned}
&E(\tau_x|Z{=}z)\\
&= E^{X=x}(\tau_x|Z{=}z) && [(6.26)]\\
&= E^{X=x}\big(E^{X=x}(Y|C_X)\,\big|\,Z{=}z\big) && [\tau_x = E^{X=x}(Y|C_X)]\\
&= E^{X=x}(Y\,|\,Z{=}z) && [Z = g(C_X), \text{SN-(10.39)}]\\
&= E(Y|X{=}x, Z{=}z) && [(6.24)]
\end{aligned}
$$

Proof of Theorem 6.26

$$
\begin{aligned}
E\big(PFE_{Z;xx'}(Z)\big) &= E\big(E^{X=x}(Y|Z) - E^{X=x'}(Y|Z)\big) && [(6.29)]\\
&= E\big(E(\tau_x - \tau_{x'}|Z)\big) && [(6.32)]\\
&= E(\tau_x - \tau_{x'}) && [\text{SN-Box 10.2 (iv)}]\\
&= ATE_{xx'}. && [(5.6), (5.7)]
\end{aligned}
$$

Proof of Theorem 6.30

$$
\begin{aligned}
&E\big(PFE_{Z;xx'}(Z)\,\big|\,V\big)\\
\overset{=}{P}\ &E\big(E^{X=x}(Y|Z) - E^{X=x'}(Y|Z)\,\big|\,V\big) && [(6.29)]\\
\overset{=}{P}\ &E\big(E^{X=x}(Y|Z)\,\big|\,V\big) - E\big(E^{X=x'}(Y|Z)\,\big|\,V\big) && [\text{SN-Box 10.2 (xv)}]\\
\overset{=}{P}\ &E\big(E(\tau_x|Z)\,\big|\,V\big) - E\big(E(\tau_{x'}|Z)\,\big|\,V\big) && [(6.16)]\\
\overset{=}{P}\ &E\big(E(\tau_x|1_{X=x}, Z)\,\big|\,V\big) - E\big(E(\tau_{x'}|1_{X=x'}, Z)\,\big|\,V\big) && [(6.21)]\\
\overset{=}{P}\ &E(\tau_x|V) - E(\tau_{x'}|V) && [\sigma(V) \subset \sigma(1_{X=x}, Z),\ \sigma(V) \subset \sigma(1_{X=x'}, Z), \text{SN-Box 10.2 (v)}]\\
\overset{=}{P}\ &E(\tau_x - \tau_{x'}|V) && [\text{SN-Box 10.2 (xv)}]\\
\overset{=}{P}\ &CTE_{V;xx'}(V). && [(5.16)]
\end{aligned}
$$

Proof of Theorem 6.32

The proof is analog to that of Theorem 6.30. In this proof, we only have to replace $E(\tau_x|1_{X=x}, Z)$ by $E(\tau_x|X, Z)$ and $E(\tau_{x'}|1_{X=x'}, Z)$ by $E(\tau_{x'}|X, Z)$.

6.9 Exercises

▷ **Exercise 6-1** What is the difference between the two terms $E(\tau_x)$ and $E(Y|X{=}x)$?

▷ **Exercise 6-2** Compute the probabilities $P(Z{=}z)$ occurring in Equation (6.59) for both values of Z in the example displayed in Table 6.1.

▷ **Exercise 6-3** Which are the probabilities $P(U{=}u_1, Z{=}m)$ and $P(U{=}u_5, Z{=}m)$ occurring in Equation (6.59) for the example displayed in Table 6.1.

▷ **Exercise 6-4** Compute the two conditional probabilities $P(U{=}u_1 \,|\, Z{=}m)$ and $P(U{=}u_5 \,|\, Z{=}m)$ displayed in Table 6.1.

▷ **Exercise 6-5** Use SN-Theorem 4.25 to compute the probability $P(X{=}1)$ for the example displayed in Table 6.1.

▷ **Exercise 6-6** Compute the probabilities $P(U{=}u \,|\, X{=}0)$ and $P(U{=}u \,|\, X{=}1)$ for all six units in the example of Table 6.1.

▷ **Exercise 6-7** Compute the conditional probabilities $P(U{=}u \,|\, X{=}1, Z{=}m)$ occurring in Equation (iii) of Box 6.2 in the example of Table 6.1.

▷ **Exercise 6-8** Compute the conditional expectation values $E(Y|X=0)$ and $E(Y|X=1)$ for the example in Table 6.3.

▷ **Exercise 6-9** Compute the conditional expectation values $E(\tau_1 \,|\, Z{=}f)$ and $E(\tau_0 \,|\, Z{=}f)$ displayed in Table 6.1.

▷ **Exercise 6-10** Show that Equation (6.58) holds.

Solutions

▷ **Solution 6-1** The term $E(\tau_x)$ denotes the expectation of a true outcome variable τ_x. It is these true outcome variables that are of interest in the empirical sciences, because they describe how the outcome variable Y depends on the values x of X controlling for all potential confounders of X. This implies that the true outcome variables cannot be biased, and this also applies to their expectations $E(\tau_x)$. In empirical applications, we often aim at estimating the expectations $E(\tau_x)$ and their differences. In contrast, the conditional expectation values $E(Y|X{=}x)$ of the outcomes are often not of interest in the empirical sciences, because they do not have a causal interpretation unless $E(Y|X{=}x) = E(\tau_x)$.

▷ **Solution 6-2** The events that U takes on the value u_i and that U takes on the value u_j, $i \neq j$, are disjoint. Therefore, we can use the theorem of total probability (see SN-Th. 4.25):

$$\begin{aligned} P(Z{=}m) &= P(Z{=}m, U{=}u_1) + \ldots + P(Z{=}m, U{=}u_6) \\ &= \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + 0 + 0 = \frac{4}{6}. \end{aligned}$$

$$\begin{aligned} P(Z{=}f) &= P(Z{=}f, U{=}u_1) + \ldots + P(Z{=}f, U{=}u_6) \\ &= 0 + 0 + 0 + 0 + \frac{1}{6} + \frac{1}{6} = \frac{2}{6}. \end{aligned}$$

▷ **Solution 6-3** $P(U{=}u_1, Z{=}m) = \frac{1}{6}$ and $P(U{=}u_5, Z{=}m) = 0$.

▷ **Solution 6-4**

$$P(U{=}u_1 \,|\, Z{=}m) \;=\; \frac{P(U{=}u_1, Z{=}m)}{P(Z{=}m)} \;=\; \frac{1/6}{4/6} \;=\; \frac{1}{4}.$$
$$P(U{=}u_5 \,|\, Z{=}m) \;=\; \frac{P(U{=}u_5, Z{=}m)}{P(Z{=}m)} \;=\; \frac{0}{4/6} \;=\; 0.$$

▷ **Solution 6-5** The events $\{U{=}u_1\},\ldots,\{U{=}u_6\}$ are disjoint and all these events have positive probabilities. Hence we can apply the theorem of total probability (see SN-Th. 4.25):

$$\begin{aligned} P(X{=}1) \;&=\; P(X{=}1 \,|\, U{=}u_1) \cdot P(U{=}u_1) + \ldots + P(X{=}1 \,|\, U{=}u_6) \cdot P(U{=}u_6) \\ &=\; \frac{6}{7}\cdot\frac{1}{6} + \frac{5}{7}\cdot\frac{1}{6} + \frac{4}{7}\cdot\frac{1}{6} + \frac{3}{7}\cdot\frac{1}{6} + \frac{2}{7}\cdot\frac{1}{6} + \frac{1}{7}\cdot\frac{1}{6} \\ &=\; \frac{21}{42} \;=\; \frac{1}{2}. \end{aligned}$$

▷ **Solution 6-6** We have to use the equation

$$P(U{=}u \,|\, X{=}x) \;=\; \frac{P(X{=}x \,|\, U{=}u) \cdot P(U{=}u)}{P(X{=}x)}.$$

For $U{=}u_1$ and $X{=}1$ this equation yields:

$$\begin{aligned} P(U{=}u_1 \,|\, X{=}1) \;&=\; \frac{P(X{=}1 \,|\, U{=}u_1) \cdot P(U{=}u_1)}{P(X{=}1)} \\ &=\; \frac{6/7 \cdot 1/6}{1/2} \;=\; \frac{6}{21}. \qquad \text{[Exercise 6 – 5]} \end{aligned}$$

Using the same procedure, we obtain $5/21, 4/21, \ldots, 1/21$, the corresponding probabilities for the units $u_2, u_3 \ldots, u_6$, respectively. For $U{=}u_1$ and $X{=}0$, we obtain

$$\begin{aligned} P(U{=}u_1 \,|\, X{=}0) \;&=\; \frac{P(X{=}0 \,|\, U{=}u_1) \cdot P(U{=}u_1)}{P(X{=}0)} \\ &=\; \frac{1/7 \cdot 1/6}{1/2} \;=\; \frac{1}{21}. \qquad \text{[Exercise 6 – 5]} \end{aligned}$$

Using the same procedure, we obtain $2/21,\ 3/21, \ldots, 6/21$ for the units $u_2, u_3 \ldots, u_6$, respectively.

▷ **Solution 6-7** According to Equation (6.58) we need the conditional probabilities $P(X{=}1 \,|\, U{=}u)$ displayed in Table 6.1. The other probabilities occurring in this equation can be computed from the probabilities displayed in the table.

One of these other probabilities that needs some computation is $P(X{=}1 \,|\, Z{=}m)$. For $X{=}1$ and $Z{=}m$ in the example of Table 6.1, Equation (6.59) results in:

$$\begin{aligned} P(X{=}1 \,|\, Z{=}m) \;&=\; \frac{\sum_u P(X{=}1 \,|\, U{=}u) \cdot P(U{=}u, Z{=}m)}{P(Z{=}m)} \\ &=\; \frac{(6/7)\cdot(1/6) + \ldots + (3/7)\cdot(1/6) + (2/7)\cdot 0 + (1/7)\cdot 0}{4/6} \;=\; \frac{27}{42}. \end{aligned}$$

Using this result, Equation (6.58) yields:

$$\begin{aligned} P(U{=}u_1 \,|\, X{=}1, Z{=}m) \;&=\; \frac{P(X{=}1 \,|\, U{=}u_1) \cdot P(U{=}u_1, Z{=}m)}{P(X{=}1 \,|\, Z{=}m) \cdot P(Z{=}m)} \\ &=\; \frac{(6/7)\cdot(1/6)}{(27/42)\cdot(4/6)} \;=\; \frac{6/7}{(27/42)\cdot 4} \;=\; \frac{6}{18}, \end{aligned}$$

as well as $5/18,\ 4/18$, and $3/18$ for the corresponding conditional probabilities for u_2, u_3, and u_4. The conditional probabilities $P(U{=}u_5 \,|\, X{=}1, Z{=}m)$ and $P(U{=}u_6 \,|\, X{=}1, Z{=}m)$ are zero.

▷ **Solution 6-8** According to Equation (i) of Box 6.2,

$$\begin{aligned} E(Y \mid X=0) &= \sum_u E(Y \mid X=0, U=u) \cdot P(U=u \mid X=0) \\ &= (68+78+88+98) \cdot \frac{1}{10} + (106+116) \cdot \frac{3}{10} \\ &= 33.2 + 66.6 \;=\; 99.8, \end{aligned}$$

and

$$\begin{aligned} E(Y \mid X=1) &= \sum_u E(Y \mid X=1, U=u) \cdot P(U=u \mid X=1) \\ &= (81+86+100+103) \cdot \frac{3}{14} + (114+130) \cdot \frac{1}{14} \\ &\approx 79.286 + 17.429 \approx 96.715. \end{aligned}$$

▷ **Solution 6-9** Remember again, in this example, $E(Y \mid X, C_X) = E(Y \mid X, U)$. Hence,

$$\begin{aligned} E(\tau_0 \mid Z=f) &= \sum_u E(Y \mid X=0, U=u) \cdot P(U=u \mid Z=f) \\ &= 68 \cdot 0 + \ldots + 98 \cdot 0 + 106 \cdot \frac{1}{2} + 116 \cdot \frac{1}{2} \;=\; 111. \\ E(\tau_1 \mid Z=f) &= \sum_u E(Y \mid X=1, U=u) \cdot P(U=u \mid Z=f) \\ &= 81 \cdot 0 + \ldots + 103 \cdot 0 + 114 \cdot \frac{1}{2} + 130 \cdot \frac{1}{2} \;=\; 122. \end{aligned}$$

▷ **Solution 6-10** In the examples presented in Tables 6.1 to 6.3, Z (sex) is U-measurable. According to SN-Corollary 2.53, there is a mapping $g\colon \Omega_U \to \{m, f\}$ such that Z is the composite function of U and g, that is, $Z = g(U)$. Therefore,

$$\{U=u, Z=z\} \;=\; \begin{cases} \{U=u\}, & \text{if } g(u) = z, \\ \emptyset, & \text{otherwise.} \end{cases} \tag{6.66}$$

According to Equation (6.66), the event to sample person u and that the sampled person is male is identical to the event to sample person u, if that person is male [i. e., $g(u) = m$]. Correspondingly, the event to sample person u and that the sampled person is female [i. e., $g(u) = f$] is identical to the event to sample person u, if that person is female. In contrast, the event to sample a male person u and to observe $Z(\omega) = g\big(U(\omega)\big) = f$, and the event to sample a female person u and to observe $Z(\omega) = g\big(U(\omega)\big) = m$ are impossible. Equation (6.66) implies

$$P(U=u, Z=z) \;=\; \begin{cases} P(U=u), & \text{if } g(u) = z, \\ 0, & \text{otherwise,} \end{cases} \tag{6.67}$$

and, because $g(u) = z$ implies $P(U=u, Z=z) > 0$, we can conclude, that in these examples,

$$\begin{aligned} P(X=x \mid U=u, Z=z) &= \frac{P(X=x, U=u, Z=z)}{P(U=u, Z=z)} \\ &= \frac{P(X=x, U=u)}{P(U=u)} \\ &= P(X=x \mid U=u), \quad \text{if } P(U=u, Z=z) > 0. \end{aligned} \tag{6.68}$$

Furthermore, in our examples, $P(X=x, Z=z) > 0$. Therefore, if $P(U=u, Z=z) > 0$, then

$$
\begin{aligned}
P(U{=}u \,|\, X{=}x, Z{=}z) \;&=\; \frac{P(X{=}x, U{=}u, Z{=}z)}{P(X{=}x, Z{=}z)} \\
&= \frac{P(X{=}x \,|\, U{=}u, Z{=}z) \cdot P(U{=}u, Z{=}z)}{P(X{=}x \,|\, Z{=}z) \cdot P(Z{=}z)} \\
&= \frac{P(X{=}x \,|\, U{=}u) \cdot P(U{=}u, Z{=}z)}{P(X{=}x \,|\, Z{=}z) \cdot P(Z{=}z)}, \qquad \text{[Eq. (6.68)]}
\end{aligned}
$$

which is Equation (6.58). If $P(U{=}u, Z{=}z) = 0$, then $P(X{=}x, U{=}u, Z{=}z) = 0$. Hence, if $P(U{=}u, Z{=}z) = 0$, then

$$
P(U{=}u \,|\, X{=}x, Z{=}z) \;=\; \frac{P(X{=}x, U{=}u, Z{=}z)}{P(X{=}x, Z{=}z)} \;=\; 0.
$$

The same result is also obtained applying Equation (6.58).

Chapter 7
Rosenbaum-Rubin Conditions

In chapter 6, we introduced *unbiasedness* of various conditional expectations, conditional expectation values, prima facie effects, and prima facie effect functions. We also studied how the various causal effects and effect functions can be identified by empirically estimable parameters if we can assume that certain terms are unbiased. Those *unbiasedness conditions* are a first of several kinds of causality conditions, which, together with the structural components listed in a causality space, distinguish *causal total effects* from differences between conditional expectation values that have no causal meaning.

In this chapter we introduce some other causality conditions, including the Rosenbaum-Rubin condition of *strong ignorability*. It will be shown that this is the most restrictive of all causality conditions treated in this chapter. Note, however, that we adapted the original Rosenbaum-Rubin condition by replacing their (deterministic) potential outcome variables by the (probabilistic) true outcome variables introduced in chapter 4.

After a brief section on independence and conditional independence of random variables, we start with causality conditions that also imply unbiasedness, but are less restrictive than strong ignorability. Unlike the causality conditions treated in chapters 8 to **??**, all causality conditions treated in this chapter deal with the true outcome variables, and they are empirically untestable. Nevertheless, they are meaningful from a theoretical perspective and distinguish conditional expectation values and conditional expectations that have a causal meaning from ordinary conditional expectation values and conditional expectations that do not have any causal meaning at all.

7.1 Independence and Conditional Independence of Random Variables

For the sections to come it is useful to define independence and conditional independence of two random variables and to review some properties. The concepts on which the following definition is based have been introduced in chapter 3 [see, e. g., Def. 3.31 for the σ-algebra $Y^{-1}(\mathscr{A}'_Y)$ generated by Y and Rem. 3.58 for the Z-conditional probability $P(A\,|\,Z)$ of the event A].

Definition 7.1 (Independence and Conditional Independence of Random Variables) *Let $X\colon\Omega\to\Omega'_X$ and $Y\colon\Omega\to\Omega'_Y$ be random variables on the probability space $(\Omega,\mathscr{A},P)$ with value spaces $(\Omega'_X,\mathscr{A}'_X)$ and $(\Omega'_Y,\mathscr{A}'_Y)$, respectively.*

(i) Then X and Y are called independent (with respect to P), denoted by $X \perp\!\!\!\perp Y$, if

$$P(A\cap B) \;=\; P(A)\cdot P(B), \quad \forall (A,B)\in X^{-1}(\mathscr{A}'_X)\times Y^{-1}(\mathscr{A}'_Y). \tag{7.1}$$

(ii) Additionally, let $Z\colon \Omega \to \Omega'_Z$ be a random variable on $(\Omega, \mathscr{A}, P)$ with value space $(\Omega'_Z, \mathscr{A}'_Z)$. Then X and Y are called ***Z-conditionally independent*** *(with respect to P), denoted by* $X \perp\!\!\!\perp Y \,|\, Z$*, if*

$$P(A \cap B \,|\, Z) \underset{P}{=} P(A \,|\, Z) \cdot P(B \,|\, Z), \quad \forall (A,B) \in X^{-1}(\mathscr{A}'_X) \times Y^{-1}(\mathscr{A}'_Y). \tag{7.2}$$

Independence of two random variables is treated in more detail, for example, in SN-section 5.4. If there is ambiguity with respect to the probability measure, then we also may add the reference to the measure and use the notation $X \underset{P}{\perp\!\!\!\perp} Y$ instead of $X \perp\!\!\!\perp Y$. The more general concept of *conditional independence* (with respect to P) of two random variables X and Y given a random variable Z is presented in SN-chapter 16, including many important theorems and propositions. A number of useful properties of conditional independence are summarized in SN-Boxes 16.2 and 16.3. Again we may use the notation $X \underset{P}{\perp\!\!\!\perp} Y \,|\, Z$ instead of $X \perp\!\!\!\perp Y \,|\, Z$ if there is ambiguity with respect to the measure P. (For an overview of conditional independence in statistics see Dawid, 1979.)

Remark 7.2 (Independence Implies Mean-Independence) Under the assumptions of Definition 7.1, if Y is a numerical random variable that is nonnegative or with a finite expectation, then

$$Y \perp\!\!\!\perp X \quad \Rightarrow \quad Y \vdash X \tag{7.3}$$

(see SN-Rem. 16.36 and Rem. 3.71 for the notation $Y \vdash X$). Similarly,

$$Y \perp\!\!\!\perp X \,|\, Z \quad \Rightarrow \quad Y \vdash X \,|\, Z \tag{7.4}$$

(see SN-Rem. 16.35). That is, Z-conditional independence of Y and X implies Z-conditional mean-independence of Y from X. The only assumption is that Y satisfies the requirements mentioned above, which imply that the conditional expectation $E(Y \,|\, X, Z)$ exists (see SN-Th. 10.9). ◁

Remark 7.3 (Conditional Independence of an Indicator and a Random Variable) Let W, X, and Z denote random variables on the same probability space $(\Omega, \mathscr{A}, P)$. Then

$$P(X{=}x \,|\, W) \underset{P}{=} P(X{=}x) \quad \Leftrightarrow \quad 1_{X=x} \perp\!\!\!\perp W, \tag{7.5}$$

and, more generally,

$$P(X{=}x \,|\, W, Z) \underset{P}{=} P(X{=}x \,|\, Z) \quad \Leftrightarrow \quad 1_{X=x} \perp\!\!\!\perp W \,|\, Z. \tag{7.6}$$

According to this proposition, $1_{X=x} \perp\!\!\!\perp W \,|\, Z$ means that the conditional probability of the event $\{X{=}x\}$ does not depend on W, once we condition on Z. Also note that Proposition (7.5) is a special case of (7.6) for Z being a constant, that is, for $Z = \alpha$, $\alpha \in \mathbb{R}$. Hence, it suffices to prove Proposition (7.6) (see Lemma 7.4). ◁

Lemma 7.4 (Equivalent Propositions on Conditional Independence)
Let X, W, and Z be random variables on the probability space $(\Omega, \mathscr{A}, P)$ with value spaces $(\Omega'_X, \mathscr{A}'_X)$, $(\Omega'_W, \mathscr{A}'_W)$, and $(\Omega'_Z, \mathscr{A}'_Z)$, respectively. Furthermore, let $x \in \Omega'_X$ and $\{x\} \in \mathscr{A}'_X$. Then the following two propositions are equivalent to each other:

(a) $P(X{=}x \,|\, W, Z) \underset{P}{=} P(X{=}x \,|\, Z)$.

(b) $1_{X=x} \perp\!\!\!\perp W \,|\, Z$.

(Proof p. 174)

Remark 7.5 (An Implication of Independence) Another useful proposition about independence and conditional independence of two random variables is

$$\big(X \perp\!\!\!\perp W \;\wedge\; \sigma(Z) \subset \sigma(W)\big) \;\Rightarrow\; \big(X \perp\!\!\!\perp Z \;\wedge\; X \perp\!\!\!\perp W \,|\, Z\big), \tag{7.7}$$

[see SN-Box 16.2 (ix)]. According to this proposition, if X and W are independent and Z is W-measurable, then this implies independence of X and Z as well as Z-conditional independence of X and W. This proposition holds for any random variables W, X, and Z on the same probability space. None of them has to be numerical and none has to be discrete. Also note that the roles of X and W can be exchanged. ◁

Remark 7.6 (Independence of X and W if X is Discrete) If X is discrete, then, according to SN-Theorem 16.26 and SN-Remark 16.27,

$$X \perp\!\!\!\perp W \quad \Leftrightarrow \quad \forall\, x \in X(\Omega)\colon\; P(X{=}x \,|\, W) \underset{P}{=} P(X{=}x). \tag{7.8}$$

That is, independence of X and W (with respect to the measure P) is equivalent to postulating $P(X{=}x \,|\, W) \underset{P}{=} P(X{=}x)$ for all values x of X. ◁

Remark 7.7 (Z-Conditional Independence of X and W if X is Discrete) Similarly, if X is discrete, then, again according to SN-Theorem 16.26 and SN-Remark 16.27,

$$X \perp\!\!\!\perp W \,|\, Z \quad \Leftrightarrow \quad \forall\, x \in X(\Omega)\colon\; P(X{=}x \,|\, Z, W) \underset{P}{=} P(X{=}x \,|\, Z). \tag{7.9}$$

Hence, Z-conditional independence of X and W (with respect to the measure P) is equivalent to postulating $P(X{=}x \,|\, Z, W) \underset{P}{=} P(X{=}x \,|\, Z)$ for all values x of X, provided that X is discrete. ◁

7.2 Mean-Independence of τ_x From X

A first causality condition introduced in this chapter is called *mean-independence of τ_x from X*. Under the Assumptions 6.1 it is defined by

$$E(\tau_x \,|\, X) \underset{P}{=} E(\tau_x). \tag{7.10}$$

A shortcut for this equation is $\tau_x \vdash X$ (see Rem. 3.71).

Remark 7.8 (Mean-Independence of τ_x From $1_{X=x}$) According to Theorem 6.9, mean-independence of τ_x from the indicator $1_{X=x}$ is equivalent to unbiasedness of $E(Y \,|\, X{=}x)$, if

we presume P-uniqueness of τ_x (see Rem. 3.83 and Th. 3.84). That is, if the Assumptions 6.1 hold and τ_x is P-unique, then

$$\tau_x \vdash 1_{X=x} \quad \Leftrightarrow \quad E(Y|X{=}x) \text{ is unbiased.} \tag{7.11}$$

In terms of equations, Proposition (7.11) can equivalently by written

$$E(\tau_x|1_{X=x}) \underset{P}{=} E(\tau_x) \quad \Leftrightarrow \quad E(Y|X{=}x) \;=\; E(\tau_x). \tag{7.12}$$

◁

Remark 7.9 (Dichotomous X) If X is dichotomous, then

$$E(\tau_x|X) \underset{P}{=} E(\tau_x|1_{X=x}), \tag{7.13}$$

because, if X is dichotomous, then the σ-algebras generated by X and $1_{X=x}$ are identical. [Remember, the σ-algebras $\sigma(X)$ and $\sigma(1_{X=x})$ play a crucial role in the definition of the conditional expectations $E(\tau_x|X)$ and $E(\tau_x|1_{X=x})$ (see Def. 3.54)]. Hence, if X is dichotomous and the true outcome variables τ_x are P-unique for both values x and x' of X, then

$$\tau_x \vdash X \quad \Leftrightarrow \quad \tau_x \vdash 1_{X=x} \quad \Leftrightarrow \quad E(Y|X) \text{ is unbiased,} \tag{7.14}$$

and this is equivalent to unbiasedness of $E(Y|X{=}x)$ and $E(Y|X{=}x')$. ◁

Remark 7.10 (Mean-Independence of τ_x From X) If X has more than two different values and τ_x is P-unique for all $x \in X(\Omega)$, then, according to Theorem 7.11 (i), mean-independence of τ_x from X still implies unbiasedness of $E(Y|X{=}x)$, but is not implied by unbiasedness. Hence, if $\tau_x \vdash X$, that is, if

$$E(\tau_x|X) \underset{P}{=} E(\tau_x), \tag{7.15}$$

then $\tau_x \vdash 1_{X=x}$. Using the shortcut

$$\tau \vdash X \quad :\Leftrightarrow \quad \forall\, x \in X(\Omega) :\; \tau_x \vdash X, \tag{7.16}$$

where $\tau := (\tau_0, \tau_1, \ldots, \tau_J)$, then this result implies the proposition on unbiasedness of $E(Y|X)$ stated in Theorem 7.11 (ii). ◁

Theorem 7.11 (Implications of Mean-Independence of τ_x From X)
Let the Assumptions 6.1 hold.

(i) If τ_x is P-unique, then

$$\tau_x \vdash X \quad \Rightarrow \quad E(Y|X{=}x) \textit{ is unbiased.} \tag{7.17}$$

(ii) If τ_x is P-unique for all $x \in X(\Omega)$, then

$$\tau \vdash X \quad \Rightarrow \quad E(Y|X) \textit{ is unbiased.} \tag{7.18}$$

(Proof p. 175)

Note that Proposition (i) of Theorem 7.11 holds for any value x of X for which $P(X{=}x) > 0$. Hence, if, for all $x \in X(\Omega)$, τ_x is P-unique, then Proposition (7.17) holds for all $x \in X(\Omega)$ (see Exercise 7-1).

Remark 7.12 (Implications of $\tau \vdash X$) Under the Assumptions 6.1 and presuming that τ_x is P-unique for all $x \in X(\Omega)$, the propositions of Theorem 7.11 can be summarized as follows:

$$\begin{aligned} \tau \vdash X \quad &\Leftrightarrow \quad \forall\, x \in X(\Omega):\ \tau_x \vdash X \\ &\Rightarrow \quad \forall\, x \in X(\Omega):\ \tau_x \vdash 1_{X=x} \\ &\Leftrightarrow \quad \forall\, x \in X(\Omega):\ E(Y\,|\,X{=}x) \text{ is unbiased} \\ &\Leftrightarrow \quad E(Y\,|\,X) \text{ is unbiased.} \end{aligned} \tag{7.19}$$

The last two equivalence propositions have already been established in Theorem 6.9 and Definition 6.2 (ii), respectively. ◁

7.3 Conditional Mean-Independence of τ_x From X

Under the Assumptions 6.1 and assuming that Z is a random variable on $(\Omega, \mathscr{A}, P)$, we define *Z-conditional mean-independence of τ_x from X* by

$$E(\tau_x\,|\,X, Z) \underset{P}{=} E(\tau_x\,|\,Z). \tag{7.20}$$

A shortcut of this equation is $\tau_x \vdash X\,|\,Z$.

Remark 7.13 (Z-Conditional Mean-Independence of τ_x From $1_{X=x}$) The propositions of Theorem 7.11 can be generalized to conditioning *on a covariate Z of X*. According to Theorem 6.15, Z-conditional mean-independence of τ_x from the indicator $1_{X=x}$ is equivalent to unbiasedness of $E^{X=x}(Y\,|\,Z)$, if we presume that τ_x is P-unique. That is, if the Assumptions 6.1 hold and τ_x is P-unique, then

$$\tau_x \vdash 1_{X=x}\,|\,Z \quad \Leftrightarrow \quad E^{X=x}(Y\,|\,Z) \text{ is unbiased.} \tag{7.21}$$

An equivalent statement of Proposition (7.21) in terms of equations is

$$E(\tau_x\,|\,1_{X=x}, Z) \underset{P}{=} E(\tau_x\,|\,Z) \quad \Leftrightarrow \quad E^{X=x}(Y\,|\,Z) \underset{P}{=} E(\tau_x\,|\,Z). \tag{7.22}$$

Note again that in Propositions (7.21) and (7.22) we presume that Z is a *covariate of X* (see Def. 4.4 and 4.9). ◁

Example 7.14 (Z-Conditional Mean-Independence of τ_x From $1_{X=x}$) Table 7.1 presents an example in which Z-conditional mean-independence of τ_x from $1_{X=x}$ holds for each value x of X, but neither Z-conditional independence of τ_x from X (see section 7.17) nor Z-conditional independence of τ from X (see section 7.4). Again, in this example, the person variable U is a global potential confounder of X. Table 7.1 displays the true outcomes under treatments 0, 1, and 2 as well as the probabilities $P(U{=}u)$ for each observational unit to be sampled, the conditional probabilities to be assigned to treatment 1, $P(X{=}1\,|\,U{=}u)$, and to treatment 2, $P(X{=}2\,|\,U{=}u)$. These are the *fundamental parameters*;

Table 7.1. Z-conditional mean-independence of τ_x from $1_{X=x}$

		Fundamental parameters						Derived parameters				
Person variable U	Sex Z	$P(U=u)$	$P(X=1\|U)$	$P(X=2\|U)$	$\tau_0 = E^{X=0}(Y\|U)$	$\tau_1 = E^{X=1}(Y\|U)$	$\tau_2 = E^{X=2}(Y\|U)$	$\delta_{10} = \tau_1 - \tau_0$	$\delta_{20} = \tau_2 - \tau_0$	$P(U=u\|X=0)$	$P(U=u\|X=1)$	$P(U=u\|X=2)$
u_1	m	1/6	3/5	1/3	75	87	97	12	22	4/63	12/59	1/6
u_2	m	1/6	1/2	1/3	70	80	92	10	22	10/63	10/59	1/6
u_3	m	1/6	3/5	1/3	65	73	107	8	42	4/63	12/59	1/6
u_4	f	1/6	1/2	1/3	106	114	90	8	−16	10/63	10/59	1/6
u_5	f	1/6	1/4	1/3	116	130	120	14	4	25/63	5/59	1/6
u_6	f	1/6	1/2	1/3	126	146	126	20	0	10/63	10/59	1/6

	$x=0$	$x=1$	$x=2$		$x=1$	$x=2$
$E(\tau_x)$:	93	105	105.333	ATE_{x0}:	12	12.333
$E(Y\|X=x)$:	102.857	101.186	105.333	PFE_{x0}:	−1.671	2.476
$E(\tau_x\|Z=m)$:	70	80	98.667	$CTE_{Z;x0}(m)$:	10	28.667
$E(Y\|X=x,Z=m)$:	70	80	98.667	$PFE_{Z;x0}(m)$:	10	28.667
$E(\tau_x\|Z=f)$:	116	130	112	$CTE_{Z;x0}(f)$:	14	−4
$E(Y\|X=x,Z=f)$:	116	130	112	$PFE_{Z;x0}(f)$:	14	−4

all other parameters, such as associated individual causal effects, the conditional probabilities $P(U=u|X=x)$, etc. can be computed from these fundamental parameters. The table also displays the values of the potential confounder $Z = sex$. Note again that the table does not contain sample data; It displays the parameters describing the laws of the single-unit trial.

Verifying Conditional Mean-Independence of τ_x From $1_{X=x}$. In this random experiment, the true treatment probabilities and true outcomes are such that $\tau_x \vdash 1_{X=x}|Z$ holds for each of the three values of X, that is,

$$E(\tau_0|1_{X=0},Z) = E(\tau_0|Z), \quad E(\tau_1|1_{X=1},Z) = E(\tau_1|Z), \quad E(\tau_2|1_{X=2},Z) = E(\tau_2|Z). \tag{7.23}$$

For example, in order to verify that $\tau_1 \vdash 1_{X=1}|Z$, that is, $E(\tau_1|1_{X=1},Z) = E(\tau_1|Z)$, we have to show that

$$\begin{aligned}
E(\tau_1|1_{X=1}{=}0, Z{=}m) &= E(\tau_1|Z{=}m),\\
E(\tau_1|1_{X=1}{=}0, Z{=}f) &= E(\tau_1|Z{=}f),\\
E(\tau_1|1_{X=1}{=}1, Z{=}m) &= E(\tau_1|Z{=}m),\\
E(\tau_1|1_{X=1}{=}1, Z{=}f) &= E(\tau_1|Z{=}f)
\end{aligned}$$

Table 7.2. Conditional probabilities $P(U=u\,|\,X=x, Z=z)$ for Table 7.1

Person variable U	$P(U=u\,\|\,X=0,Z=m)$	$P(U=u\,\|\,X=1,Z=m)$	$P(U=u\,\|\,X=2,Z=m)$	$P(U=u\,\|\,X=0,Z=f)$	$P(U=u\,\|\,X=1,Z=f)$	$P(U=u\,\|\,X=2,Z=f)$
u_1	2/9	6/17	1/3	0	0	0
u_2	5/9	5/17	1/3	0	0	0
u_3	2/9	6/17	1/3	0	0	0
u_4	0	0	0	2/9	2/5	1/3
u_5	0	0	0	5/9	1/5	1/3
u_6	0	0	0	2/9	2/5	1/3

hold for these four conditional expectation values.

The conditional expectation value $E(\tau_1\,|\,1_{X=1}{=}0, Z{=}m)$ can be computed from Tables 7.1 and 7.2 by

$$\begin{aligned} &E(\tau_1\,|\,1_{X=1}{=}0, Z{=}m) \\ &= E(\tau_1\,|\,X{=}0, Z{=}m) \;=\; \sum_u E(Y\,|\,X{=}1, U{=}u)\cdot P(U{=}u\,|\,X{=}0, Z{=}m) \\ &= 87\cdot\frac{2}{9} + 80\cdot\frac{5}{9} + 73\cdot\frac{2}{9} \;=\; 80 \end{aligned} \tag{7.24}$$

(see Exercise ???). This is exactly the same result as obtained for

$$E(\tau_1\,|\,Z{=}m) \;=\; \sum_u E(Y\,|\,X{=}1, U{=}u)\cdot P(U{=}u\,|\,Z{=}m) \;=\; (87+80+73)\cdot\frac{1}{3} \;=\; 80.$$

Analogously,

$$\begin{aligned} &E(\tau_1\,|\,1_{X=1}{=}1, Z{=}m) \\ &= E(\tau_1\,|\,X{=}1, Z{=}m) \;=\; \sum_u E(Y\,|\,X{=}1, U{=}u)\cdot P(U{=}u\,|\,X{=}1, Z{=}m) \\ &= 87\cdot\frac{6}{17} + 80\cdot\frac{5}{17} + 73\cdot\frac{6}{17} \;=\; 80. \end{aligned} \tag{7.25}$$

Hence,

$$E(\tau_1\,|\,X{=}0, Z{=}m) \;=\; E(\tau_1\,|\,X{=}1, Z{=}m) \;=\; E(\tau_1\,|\,Z{=}m) \;=\; 80,$$

and the correspondig identities hold for the other two combinations of values of X and Z:

$$E(\tau_1\,|\,X{=}0, Z{=}f) \;=\; E(\tau_1\,|\,X{=}1, Z{=}f) \;=\; E(\tau_1\,|\,Z{=}m) \;=\; 130.$$

In the same way it can be shown that $\tau_0 \vdash 1_{X=0}\,|\,Z$ and $\tau_2 \vdash 1_{X=2}\,|\,Z$ hold as well. ◁

Remark 7.15 (Z-Conditional Mean-Independence of τ_x From X) If X has more than two different values and we assume

$$E(\tau_x \mid X, Z) \underset{P}{=} E(\tau_x \mid Z), \tag{7.26}$$

abbreviated by $\tau_x \vdash X \mid Z$, then this implies $\tau_x \vdash 1_{X=x} \mid Z$, which is equivalent to $E^{X=x}(Y \mid Z)$ being unbiased, provided that τ_x is P-unique and Z is a covariate of X. ◁

Remark 7.16 (Dichotomous X) If X is dichotomous, then

$$E(\tau_x \mid X, Z) \underset{P}{=} E(\tau_x \mid 1_{X=x}, Z), \tag{7.27}$$

because $\sigma(X,Z) = \sigma(1_{X=x}, Z)$ (see again Def. 3.54). Hence, if X is dichotomous, τ_x is P-unique for both values x of X, and Z is a covariate of X, then

$$\tau_x \vdash X \mid Z \quad \Leftrightarrow \quad \tau_x \vdash 1_{X=x} \mid Z \quad \Leftrightarrow \quad E(Y \mid X, Z) \text{ is unbiased.} \tag{7.28}$$

◁

Remark 7.17 (Strong Mean-Ignorability) For all $x \in X(\Omega)$, assume that τ_x is P-unique, that Z is a covariate of X, and $\tau_x \vdash X \mid Z$. Then, alluding to Rosenbaum and Rubin's strong ignorability condition (see, e. g., Rosenbaum & Rubin, 1983b), we say that *strong mean-ignorability* holds. Now we define $\tau := (\tau_0, \tau_1, \ldots, \tau_J)$ and introduce the shortcut

$$\tau \vdash X \mid Z \quad :\Leftrightarrow \quad \forall\, x \in X(\Omega)\colon\ E(\tau_x \mid X, Z) \underset{P}{=} E(\tau_x \mid Z). \tag{7.29}$$

Then $\tau \vdash X \mid Z$ is a shortcut for strong mean-ignorability, presuming that each τ_x, $x \in X(\Omega)$, is P-unique and Z is a covariate of X. ◁

According to Theorem 7.18 (iii), strong mean ignorability implies that $E(Y \mid X, Z)$ is unbiased.

Theorem 7.18 (Implications of Z-Conditional Mean-Independence of τ_x From X) *Let the Assumptions 6.1 hold and let Z be a covariate of X.*

(i) If τ_x is P-unique, then

$$\tau_x \vdash X \mid Z \quad \Rightarrow \quad E^{X=x}(Y \mid Z) \text{ is unbiased.} \tag{7.30}$$

(ii) If τ_x is P-unique and $P(X{=}x, Z{=}z) > 0$, then

$$\tau_x \vdash X \mid Z \quad \Rightarrow \quad E(Y \mid X{=}x, Z{=}z) \text{ is unbiased.} \tag{7.31}$$

(iii) If, for all values $x \in X(\Omega)$, τ_x is P-unique, then

$$\tau \vdash X \mid Z \quad \Rightarrow \quad E(Y \mid X, Z) \text{ is unbiased.} \tag{7.32}$$

(i).

$$\begin{aligned} E(\tau_x \mid 1_{X=x}, Z) &\underset{P}{=} E\big(E(\tau_x \mid X, Z) \,\big|\, 1_{X=x}, Z\big) && \text{[SN-Box 10.2 (v)]} \\ &\underset{P}{=} E\big(E(\tau_x \mid Z) \,\big|\, 1_{X=x}, Z\big) && \text{[(7.20)]} \end{aligned}$$

$$\underset{P}{=} E(\tau_x|Z). \qquad \text{[SN-Box 10.2 (vii)]}$$

If we presume that τ_x is P-unique, then, according to Theorem 6.15, this equation is equivalent to unbiasedness of $E^{X=x}(Y|Z)$.

(ii). If τ_x is P-unique, then it is also $P^{Z=z}$-unique [(see SN-Box 14.1 (v)]. Furthermore, (i) implies

$$E^{X=x}(Y|Z) \underset{P}{=} E(\tau_x|Z).$$

If $P(X=x,Z=z)>0$, then $P(Z=z)>0$, and, according to SN-Equation (10.31), this equation implies

$$E^{X=x}(Y|Z=z) = E(\tau_x|Z=z).$$

Equation (3.70) then yields $E(Y|X=x,Z=z)=E(\tau_x|Z=z)$.

(iii). This proposition immediately follows from (i), (7.29), and Definition 6.11 (iii).

(Proof p. 175)

Again, note that Theorem 7.18 (i) holds for any value x of X for which $P(X=x)>0$. Hence, if τ_x is P-unique for all values $x\in X(\Omega)$, then Proposition (7.30) holds for all $x\in X(\Omega)$.

Remark 7.19 (Implications of Strong Mean-Ignorability) Presuming that τ_x is P-unique for all $x\in X(\Omega)$ and that Z is a covariate of X, the propositions of Theorem 7.18 can be summarized by

$$\begin{aligned} \tau \vdash X|Z \quad &\Leftrightarrow \quad \forall\, x\in X(\Omega):\ \tau_x \vdash X|Z \\ &\Rightarrow \quad \forall\, x\in X(\Omega):\ \tau_x \vdash 1_{X=x}|Z \\ &\Leftrightarrow \quad \forall\, x\in X(\Omega):\ E^{X=x}(Y|Z) \text{ is unbiased} \\ &\Leftrightarrow \quad E(Y|X,Z) \text{ is unbiased.} \end{aligned} \tag{7.33}$$

Again, note that the last two equivalence propositions have already been established in Theorem 6.15 and Definition 6.11 (iii), respectively. ◁

Example 7.20 (Conditional Mean-Independence of τ From X) Table 7.3 presents an example in which Z-conditional mean-independence of τ from X holds, but not Z-conditional independence of τ from X (see section 7.4). Again, in this example, the person variable U is a global potential confounder of X. This table displays the true outcomes under treatment and under control as well as the probabilities for each observational unit to be assigned to treatment condition $X=1$. These are the *fundamental parameters*; all other parameters, such as associated individual causal effects, the conditional probabilities $P(U=u|X=x)$, etc. can be computed from these fundamental parameters. The table also displays the values of the potential confounder $Z=$ *sex*. Note again that the table does not contain any sample data. It displays the parameters describing the laws of the single-unit trial.

Verifying That Conditional Mean-Independence of τ From X holds. In this random experiment, the true treatment probabilities and true outcomes are such that $\tau \vdash X|Z$ holds,

Table 7.3. Z-conditional mean-independence of τ from X

		Fundamental parameters				Derived parameters						
Person variable U	Sex Z	$P(U=u)$	$P(X=1\,\vert\,U)$	$\tau_0 = E^{X=0}(Y\,\vert\,U)$	$\tau_1 = E^{X=1}(Y\,\vert\,U)$	$\delta_{10} = \tau_1 - \tau_0$	$P(U=u\,\vert\,X=0)$	$P(U=u\,\vert\,X=1)$	$P(U=u\,\vert\,X=0, Z=m)$	$P(U=u\,\vert\,X=1, Z=m)$	$P(U=u\,\vert\,X=0, Z=f)$	$P(U=u\,\vert\,X=1, Z=f)$
u_1	m	1/6	16/20	75	87	12	4/50	16/70	2/10	8/30	0	0
u_2	m	1/6	14/20	70	80	10	6/50	14/70	3/10	7/30	0	0
u_3	m	1/6	16/20	93	97	4	4/50	16/70	2/10	8/30	0	0
u_4	m	1/6	14/20	98	104	6	6/50	14/70	3/10	7/30	0	0
u_5	f	1/6	5/20	106	114	8	15/50	5/70	0	0	1/2	1/2
u_6	f	1/6	5/20	116	130	14	15/50	5/70	0	0	1/2	1/2

	$x=0$	$x=1$	
$E(\tau_x)$:	93	102	$ATE_{10} = 9$
$E(Y\,\vert\,X{=}x)$:	102.2	96.286	$PFE_{10} = -3.914$
$E(\tau_x\,\vert\,Z{=}m)$:	84	92	$CTE_{Z;10}(m) = 8$
$E(Y\,\vert\,X{=}x, Z{=}m)$:	84	92	$PFE_{Z;10}(m) = 8$
$E(\tau_x\,\vert\,Z{=}f)$:	111	122	$CTE_{Z;10}(f) = 11$
$E(Y\,\vert\,X{=}x, Z{=}f)$:	111	122	$PFE_{Z;10}(f) = 11$

that is,

$$E(\tau_0\,\vert\,X,Z) \;=\; E(\tau_0\,\vert\,Z) \qquad \text{and} \qquad E(\tau_1\,\vert\,X,Z) \;=\; E(\tau_1\,\vert\,Z). \tag{7.34}$$

In order to verify that $\tau \vdash X\,\vert\,Z$ holds, consider

$$E(\tau_x\,\vert\,X{=}x^*, Z{=}z) \;=\; \sum_u E(Y\,\vert\,X{=}x, U{=}u)\cdot P(U{=}u\,\vert\,X{=}x^*, Z{=}z), \tag{7.35}$$

which is true if there is a mapping f such that $Z = f(U)$ and $P(X{=}x^*, Z{=}z) > 0$ (see Exercise ???). Note that the values x and x^* of X may be different in this equation. If we want to apply this equation to the example in Table 7.3, we have to use the probabilities $P(U{=}u\,\vert\,X{=}x^*, Z{=}z)$ displayed in the last four columns of Table 7.3 (see Exercise 6-7) ???.

Using Equation (7.35), we can now compute

$$\begin{aligned} E(\tau_0\,\vert\,X{=}1, Z{=}m) \;&=\; \sum_u E(Y\,\vert\,X{=}0, U{=}u)\cdot P(U{=}u\,\vert\,X{=}1, Z{=}m) \\ &=\; 75\cdot 8/30 + 70\cdot 7/30 + 93\cdot 8/30 + 98\cdot 7/30 \;=\; 84. \end{aligned}$$

This is exactly the same result as obtained for

$$E(\tau_0\,\vert\,Z{=}m) \;=\; \sum_u E(Y\,\vert\,X{=}0, U{=}u)\cdot P(U{=}u\,\vert\,Z{=}m) \;=\; (75+70+93+98)\cdot 1/4 \;=\; 84.$$

Hence, $E(\tau_0 | X{=}1, Z{=}m) = E(\tau_0 | Z{=}m)$, and the same identities can be shown for all other combinations of values of X and Z. Because the corresponding property also holds for τ_1 [see Eq. (7.34)], the true outcome variables τ_0 and τ_1 are Z-conditionally mean-independent from the treatment variable X. ◁

7.4 Z-Conditional Independence of τ and X (Strong Ignorability)

Rosenbaum and Rubin (1983b) presented the *strong ignorability condition*, which plays a crucial role in Rubin's approach to causal effects. In this section we translate this condition into probability theory, replacing Rubin's potential outcome variables by the true outcome variables.

Remark 7.21 (Strong Ignorability) Let $X(\Omega) = \{0, 1, \ldots, J\}$, and let $\tau = (\tau_0, \tau_1, \ldots, \tau_J)$ denote a $(J{+}1)$-dimensional random variable consisting of the unidimensional true outcome variables $\tau_0, \tau_1, \ldots, \tau_J$. Furthermore, assume that each τ_x, $x \in X(\Omega)$, is P-unique, and that Z is a covariate of X. Then strong ignorability given Z can be written as

$$\tau \perp\!\!\!\perp X \,|\, Z, \tag{7.36}$$

or equivalently, using Lemma 7.4,

$$\forall\, x \in X(\Omega): \quad P(X{=}x \,|\, \tau, Z) \underset{P}{=} P(X{=}x \,|\, Z). \tag{7.37}$$

Note that the additional assumption $P(X{=}x \,|\, Z) \underset{P}{>} 0$ of Rosenbaum and Rubin (1983b) follows from P-uniqueness of τ_x (see Exercise 7-2). ◁

Remark 7.22 (A Reminder About P-Uniqueness) Also remember that

$$P(X{=}x \,|\, C_X) \underset{P}{>} 0 \quad \Leftrightarrow \quad \tau_x \text{ is } P\text{-unique}, \tag{7.38}$$

and that P-uniqueness of τ_x is required in the definition of causal conditional and average total effects (see Defs. 5.17 and 5.8). The weaker assumption $P(X{=}x \,|\, Z) \underset{P}{>} 0$ is neither sufficient for the definition of the causal average total effect nor is it sufficient for $\tau \perp\!\!\!\perp X \,|\, Z$ to imply unbiasedness of $E(Y \,|\, X, Z)$. ◁

Example 7.23 (Z-Conditional Independence of τ and X) Table 7.4 displays the true outcomes under treatment and under control as well as the probabilities for each observational unit to be assigned to treatment condition 1. In the random experiment presented in this table, the true treatment probabilities and true outcomes are such that Z-conditional independence of τ and X and strong ignorability hold.

We check if Proposition (7.37) actually holds for $X{=}1$, that is, we check

$$P(X{=}1 \,|\, Z, \tau) \underset{P}{=} P(X{=}1 \,|\, Z). \tag{7.39}$$

According to Lemma 7.4, Equation (7.39) is equivalent to $\tau \perp\!\!\!\perp \mathbb{1}_{X=1} \,|\, Z$ and because, in this example, $\mathbb{1}_{X=1} = X$, this equation is equivalent to $\tau \perp\!\!\!\perp X \,|\, Z$. Because Z and τ are discrete, in this example, Equation (7.39) is also equivalent to

$$\begin{aligned} &P(X{=}1 \,|\, Z{=}z, \tau{=}t) \;=\; P(X{=}1 \,|\, Z{=}z), \\ &\text{for all pairs } (z, t) \text{ with } P(Z{=}z, \tau{=}t) > 0 \end{aligned} \tag{7.40}$$

Table 7.4. Strong ignorability

			Fundamental parameters					Derived parameters		
Person variable U	Sex Z_1	College Z_2	$P(U=u)$	$P(X=1\,\vert\,U)$	$P(X=1\,\vert\,Z_1,Z_2)$	$\tau_0=E^{X=0}(Y\,\vert\,U)$	$\tau_1=E^{X=1}(Y\,\vert\,U)$	$\delta_{10}=\tau_1-\tau_0$	$P(U=u\,\vert\,X=0)$	$P(U=u\,\vert\,X=1)$
u_1	m	*no*	1/6	7/8	6/8	72	83	11	1/22	7/26
u_2	m	*no*	1/6	5/8	6/8	72	83	11	3/22	5/26
u_3	m	*yes*	1/6	5/8	5/8	95	100	5	3/22	5/26
u_4	m	*yes*	1/6	5/8	5/8	100	105	5	3/22	5/26
u_5	f	*yes*	1/6	2/8	2/8	106	114	8	6/22	2/26
u_6	f	*yes*	1/6	2/8	2/8	116	130	14	6/22	2/26

	$x=0$	$x=1$	
$E(\tau_x)$:	93.5	102.5	$ATE_{10}=9$
$E(Y\,\vert\,X=x)$:	100.227	96.5	$PFE_{10}=-3.727$
$E\big(\tau_x\,\big\vert\,Z=(m,no)\big)$:	72	83	$CTE_{Z;10}(m,no)=11$
$E\big(Y\,\big\vert\,X=x,Z=(m,no)\big)$:	72	83	$PFE_{Z;10}(m,no)=11$
$E\big(\tau_x\,\big\vert\,Z=(m,yes)\big)$:	97.5	102.5	$CTE_{Z;10}(m,yes)=5$
$E\big(Y\,\big\vert\,X=x,Z=(m,yes)\big)$:	97.5	102.5	$PFE_{Z;10}(m,yes)=5$
$E\big(\tau_x\,\big\vert\,Z=(f,yes)\big)$:	111	122	$CTE_{Z;10}(f,yes)=11$
$E\big(Y\,\big\vert\,X=x,Z=(f,yes)\big)$:	111	122	$PFE_{Z;10}(f,yes)=11$

Note: In this table $Z=(Z_1,Z_2)$ is a two-dimensional random variable.

[see SN-Rem. 10.4 and SN-Cor. 10.39 (i)]. Note that, in this example, $Z=(Z_1,Z_2)$ and $\tau=(\tau_0,\tau_1)$ are two-dimensional random variables.

In the sequel, we use

$$P(X=x\,|\,V=v)\;=\;\sum_u P(X=x\,|\,V=v,U=u)\cdot P(U=u|V=v), \tag{7.41}$$

which is always true if $P(V=v,U=u)>0$ for all values of U [see SN-Box 9.2 (ii) for $Y=1_{X=x}$]. Using Equation (7.41) with Z taking the role of V and considering $X=1$ and $Z=(m,no)$ yields

$$\begin{aligned}
&P\big(X=1\,\big|\,Z=(m,no)\big)\\
&=\sum_u P\big(X=1\,\big|\,Z=(m,no),U=u\big)\cdot P\big(U=u\,\big|\,Z=(m,no)\big)\\
&=7/8\cdot 1/2+5/8\cdot 1/2\;=\;6/8.
\end{aligned}$$

In this case we only have to sum over the first two units displayed in Table 7.4, because, for the other four units, the probabilities $P\big(U=u\,\big|\,Z=(m,no)\big)$ are zero. Applying Equation

(7.41) to $V = (Z, \tau)$ and the combination of values $Z = (m, no)$ and $\tau = (72, 83)$ yields exactly the same probability

$$\begin{aligned} & P\big(X{=}1 \,\big|\, Z{=}(m,no), \tau{=}(72,83)\big) \\ = & \sum_u P\big(X{=}1 \,\big|\, Z{=}(m,no), \tau{=}(72,83), U{=}u\big) \cdot P\big(U{=}u \,\big|\, Z{=}(m,no), \tau{=}(72,83)\big) \\ = & \, 7/8 \cdot 1/2 + 5/8 \cdot 1/2 \;=\; 6/8 \end{aligned}$$

(see the first two rows of Table 7.4). Hence we have shown

$$P\big(X{=}1 \,\big|\, Z{=}(m,no)\big) \;=\; P\big(X{=}1 \,\big|\, Z{=}(m,no), \tau{=}(72,83)\big) \;=\; 6/8.$$

Using Equation (7.41) with Z taking the role of V and considering the case $X{=}1$ and $Z{=}(m,yes)$ yields

$$\begin{aligned} & P\big(X{=}1 \,\big|\, Z{=}(m,yes)\big) \\ = & \sum_u P\big(X{=}1 \,\big|\, Z{=}(m,yes), U{=}u\big) \cdot P\big(U{=}u \,\big|\, Z{=}(m,yes)\big) \\ = & \, 5/8 \cdot 1/2 + 5/8 \cdot 1/2 \;=\; 5/8. \end{aligned}$$

In this case we only have to sum over units three and four displayed in Table 7.4; for the other four units, the conditional probabilities $P\big(U{=}u \,\big|\, Z{=}(m,yes)\big)$ are zero.

Applying Equation (7.41) to $V = (Z, \tau)$ and the combination of values $Z = (m, yes)$ and $\tau = (95, 100)$ yields exactly the same conditional probability

$$\begin{aligned} & P\big(X{=}1 \,\big|\, Z{=}(m,yes), \tau{=}(95,100)\big) \\ = & \sum_u P\big(X{=}1 \,\big|\, Z{=}(m,yes), \tau{=}(95,100), U{=}u\big) \cdot P\big(U{=}u \,\big|\, Z{=}(m,yes), \tau{=}(95,100)\big) \\ = & \, 5/8 \cdot 1 \;=\; 5/8 \end{aligned}$$

(see the third row of Table 7.4). The same result is obtained if we apply Equation (7.41) to $V = (Z, \tau)$ and the combination of values $Z = (m, yes)$ and $\tau = (100, 105)$ (see the fourth row of Table 7.4). Hence we have shown

$$\begin{aligned} P\big(X{=}1 \,\big|\, Z{=}(m,yes)\big) \;&=\; P\big(X{=}1 \,\big|\, Z{=}(m,yes), \tau{=}(95,100)\big) \\ &=\; P\big(X{=}1 \,\big|\, Z{=}(m,yes), \tau{=}(100,105)\big) \;=\; 5/8. \end{aligned}$$

Finally, the analog procedure yields

$$\begin{aligned} P\big(X{=}1 \,\big|\, Z{=}(f,yes)\big) \;&=\; P\big(X{=}1 \,\big|\, Z{=}(f,yes), \tau{=}(106,114)\big) \\ &=\; P\big(X{=}1 \,\big|\, Z{=}(f,yes), \tau{=}(116,130)\big) \;=\; 2/8. \end{aligned}$$

This proves that Proposition (7.40), and with it, Equation (7.39) hold in this example. ◁

Remark 7.24 (Weak Ignorability and Some Weaker Causality Conditions) Assuming that τ_x is P-unique and that Z is a covariate of X, we start investigating the implications of the following four causality conditions, all of which are weaker than strong ignorability. That is, all four conditions are less restrictive than strong ignorability and all four are implied by strong ignorability.

First,

$$P(X{=}x\,|\,\tau_x, Z) \underset{P}{=} P(X{=}x\,|\,Z), \tag{7.42}$$

which we denote by $\tau_x \perp\!\!\!\perp 1_{X=x}\,|\,Z$. Second,

$$\forall\, x \in X(\Omega):\ P(X{=}x\,|\,\tau_x, Z) \underset{P}{=} P(X{=}x\,|\,Z), \tag{7.43}$$

denoted by $\forall x\colon \tau_x \perp\!\!\!\perp 1_{X=x}\,|\,Z$. Third

$$\forall\, x' \in X(\Omega):\ P(X{=}x'\,|\,\tau_x, Z) \underset{P}{=} P(X{=}x'\,|\,Z), \tag{7.44}$$

denoted by $\tau_x \perp\!\!\!\perp X\,|\,Z$. Fourth,

$$\forall\, x \in X(\Omega):\ \tau_x \perp\!\!\!\perp X\,|\,Z, \tag{7.45}$$

denoted by $\forall x\colon \tau_x \perp\!\!\!\perp X\,|\,Z$. This condition is called *weak ignorability* in Porta (2014, p. 142)

According to the following theorem, if we assume that τ_x is P-unique and that Z is a covariate of X, then $\tau_x \perp\!\!\!\perp 1_{X=x}\,|\,Z$ implies unbiasedness of $E^{X=x}(Y\,|\,Z)$, which is equivalent to $\tau_x \vdash 1_{X=x}\,|\,Z$. Using the shortcut

$$\forall x\colon \tau_x \vdash 1_{X=x}\,|\,Z \quad :\Leftrightarrow \quad \forall\, x \in X(\Omega):\ E(\tau_x\,|\,1_{X=x}, Z) \underset{P}{=} E(\tau_x\,|\,Z), \tag{7.46}$$

Theorem 7.25 (ii) states a sufficient condition for unbiasedness of $E(Y\,|\,X,Z)$. ◁

Theorem 7.25 (Implications of $\tau_x \perp\!\!\!\perp \mathbf{1}_{X=x}\,|\,Z$)
Let the Assumptions 6.1 hold and let Z be a covariate of X.

(i) If τ_x is P-unique, then

$$\tau_x \perp\!\!\!\perp 1_{X=x}\,|\,Z \quad \Rightarrow \quad E^{X=x}(Y\,|\,Z) \text{ is unbiased.} \tag{7.47}$$

(ii) If τ_x is P-unique for all $x \in X(\Omega)$, then

$$\forall x\colon \tau_x \vdash 1_{X=x}\,|\,Z \quad \Rightarrow \quad E(Y\,|\,X,Z) \text{ is unbiased.} \tag{7.48}$$

(Proof p. 175)

Again, note that proposition (i) of Theorem 7.25 holds for any value x of X for which $P(X{=}x) > 0$. Hence, if τ_x is P-unique for all $x \in X(\Omega)$ and Z is a covariate of X, then Proposition (7.47) holds for all $x \in X(\Omega)$.

Remark 7.26 (Implications of $\forall x\colon \tau_x \perp\!\!\!\perp \mathbf{1}_{X=x}\,|\,Z$) If we presume that τ_x is P-unique for all $x \in X(\Omega)$ and that Z is a covariate of X, then the propositions of Theorem 7.25 can be summarized as follows:

$$\begin{aligned} \forall x\colon \tau_x \perp\!\!\!\perp 1_{X=x}\,|\,Z \quad &\Rightarrow \quad \forall x\colon \tau_x \vdash 1_{X=x}\,|\,Z \\ &\Leftrightarrow \quad \forall\, x \in X(\Omega):\ E^{X=x}(Y\,|\,Z) \text{ is unbiased} \\ &\Leftrightarrow \quad E(Y\,|\,X,Z) \text{ is unbiased.} \end{aligned} \tag{7.49}$$

◁

For Z being a constant, Theorem 7.25 immediately implies the following corollary.

Corollary 7.27 (Implications of $\tau_x \perp\!\!\!\perp \mathbf{1}_{X=x}$)
Let the Assumptions 6.1 hold.

(i) If τ_x is P-unique, then

$$\tau_x \perp\!\!\!\perp 1_{X=x} \quad \Rightarrow \quad E(Y|X=x) \text{ is unbiased.} \tag{7.50}$$

(ii) If τ_x is P-unique for all $x \in X(\Omega)$, then

$$\forall\, x\colon \tau_x \perp\!\!\!\perp 1_{X=x} \quad \Rightarrow \quad E(Y|X) \text{ is unbiased.} \tag{7.51}$$

Of course, if τ_x is P-unique for all $x \in X(\Omega)$, then Proposition (7.50) holds for all $x \in X(\Omega)$.

Remark 7.28 (Implications of $\forall\, x\colon \tau_x \perp\!\!\!\perp \mathbf{1}_{X=x}$) If we introduce the shortcuts

$$\forall\, x\colon \tau_x \perp\!\!\!\perp 1_{X=x} \quad :\Leftrightarrow \quad \forall\, x \in X(\Omega)\colon \tau_x \perp\!\!\!\perp 1_{X=x}, \tag{7.52}$$

and

$$\forall\, x\colon \tau_x \vdash 1_{X=x} \quad :\Leftrightarrow \quad \forall\, x \in X(\Omega)\colon \tau_x \vdash 1_{X=x}, \tag{7.53}$$

then, presuming that τ_x is P-unique for all $x \in X(\Omega)$, the propositions of Corollary 7.27 can be written as follows:

$$\begin{aligned} \forall\, x\colon \tau_x \perp\!\!\!\perp 1_{X=x} \quad &\Rightarrow \quad \forall\, x\colon \tau_x \vdash 1_{X=x} \\ &\Leftrightarrow \quad \forall\, x \in X(\Omega)\colon E(Y|X=x) \text{ is unbiased} \\ &\Leftrightarrow \quad E(Y|X) \text{ is unbiased.} \end{aligned} \tag{7.54}$$

◁

In the following corollary we use the shortcuts

$$\forall\, x\colon \tau_x \perp\!\!\!\perp X|Z \quad :\Leftrightarrow \quad \forall\, x \in X(\Omega)\colon \tau_x \perp\!\!\!\perp X|Z, \tag{7.55}$$

and

$$\forall\, x\colon \tau_x \vdash X|Z \quad :\Leftrightarrow \quad \forall\, x \in X(\Omega)\colon \tau_x \vdash X|Z. \tag{7.56}$$

Corollary 7.29 (Implications of $\tau_x \perp\!\!\!\perp X|Z$)
Let the Assumptions 6.1 hold and let Z be a covariate of X.

(i) If τ_x is P-unique, then

$$\tau_x \perp\!\!\!\perp X|Z \quad \Rightarrow \quad \tau_x \perp\!\!\!\perp 1_{X=x}|Z \text{ and } \tau_x \vdash X|Z. \tag{7.57}$$

(ii) If τ_x is P-unique for all $x \in X(\Omega)$, then

$$\forall\, x\colon \tau_x \perp\!\!\!\perp X|Z \quad \Rightarrow \quad E(Y|X,Z) \text{ is unbiased.} \tag{7.58}$$

(Proof p. 175)

If τ_x is P-unique for all $x \in X(\Omega)$ and Z is a covariate of X, then Proposition (7.57) holds for all $x \in X(\Omega)$.

Remark 7.30 (Implications of $\forall x\colon \tau_x \perp\!\!\!\perp X \,|\, Z$) If we presume that τ_x is P-unique for all $x \in X(\Omega)$ and Z is a covariate of X, then the propositions of Corollary 7.29 can be summarized as follows:

$$\begin{aligned} \forall x\colon \tau_x \perp\!\!\!\perp X \,|\, Z \quad &\Rightarrow \quad \forall x\colon \tau_x \perp\!\!\!\perp 1_{X=x} \,|\, Z \\ &\Rightarrow \quad \forall x\colon \tau_x \vdash 1_{X=x} \,|\, Z \\ &\Leftrightarrow \quad \forall\, x \in X(\Omega)\colon\ E^{X=x}(Y \,|\, Z) \text{ is unbiased} \\ &\Leftrightarrow \quad E(Y \,|\, X, Z) \text{ is unbiased.} \end{aligned} \tag{7.59}$$

◁

For Z being a constant, Corollary 7.29 immediately implies the following corollary.

Corollary 7.31 (Implications of $\tau_x \perp\!\!\!\perp X$)
Let the Assumptions 6.1 hold.

(i) If τ_x is P-unique, then

$$\tau_x \perp\!\!\!\perp X \quad \Rightarrow \quad \tau_x \perp\!\!\!\perp 1_{X=x} \ \textit{and}\ \ \tau_x \vdash X. \tag{7.60}$$

(ii) If τ_x is P-unique for all $x \in X(\Omega)$, then

$$\forall x\colon \tau_x \perp\!\!\!\perp X \quad \Rightarrow \quad E(Y \,|\, X) \textit{ is unbiased.} \tag{7.61}$$

Of course, if τ_x is P-unique for all $x \in X(\Omega)$, then Proposition (7.60) holds for all $x \in X(\Omega)$,

Remark 7.32 (Implications of $\forall x\colon \tau_x \perp\!\!\!\perp X$) Presuming that τ_x is P-unique for all $x \in X(\Omega)$, the propositions of Corollary 7.29 can be summarized as follows:

$$\begin{aligned} \forall x\colon \tau_x \perp\!\!\!\perp X \quad &\Rightarrow \quad \forall x\colon \tau_x \perp\!\!\!\perp 1_{X=x} \\ &\Rightarrow \quad \forall x\colon \tau_x \vdash 1_{X=x} \\ &\Leftrightarrow \quad \forall\, x \in X(\Omega)\colon\ E(Y \,|\, X{=}x) \text{ is unbiased} \\ &\Leftrightarrow \quad E(Y \,|\, X) \text{ is unbiased.} \end{aligned} \tag{7.62}$$

◁

Remark 7.33 ($\tau \perp\!\!\!\perp X \,|\, Z$ implies $\tau_x \perp\!\!\!\perp 1_{X=x} \,|\, Z$) According to Lemma 7.4,

$$P(X{=}x \,|\, Z, \tau_x) \underset{P}{=} P(X{=}x \,|\, Z) \quad \Leftrightarrow \quad \tau_x \perp\!\!\!\perp 1_{X=x} \,|\, Z. \tag{7.63}$$

As mentioned before, even if we assume $\tau_x \perp\!\!\!\perp X \,|\, Z$ for all $x \in X(\Omega)$, then this is less restrictive than the original strong ignorability condition of Rosenbaum and Rubin [see Eq. (7.36)]. More precisely, for $\tau = (\tau_0, \tau_1, \ldots, \tau_J)$,

$$\tau \perp\!\!\!\perp X \,|\, Z \quad \Rightarrow \quad \forall\, x \in X(\Omega)\colon\ \tau_x \perp\!\!\!\perp X \,|\, Z \tag{7.64}$$

[see SN-Box 16.3 (vi)], but not vice versa. Hence, together with Theorem 7.25 this immediately implies the following corollary. ◁

Corollary 7.34 (Implications of $\tau \perp\!\!\!\perp X \mid Z$)
Let the Assumptions 6.1 hold, assume that τ_x is P-unique, for all $x \in X(\Omega)$, and let Z be a covariate of X. Then:

$$\begin{aligned}
\tau \perp\!\!\!\perp X \mid Z \quad &\Rightarrow \quad \forall\, x \in X(\Omega):\ \tau_x \perp\!\!\!\perp X \mid Z && (7.65)\\
&\Rightarrow \quad \forall\, x \in X(\Omega):\ \tau_x \perp\!\!\!\perp 1_{X=x} \mid Z && (7.66)\\
&\Rightarrow \quad \forall\, x \in X(\Omega):\ \tau_x \vdash 1_{X=x} \mid Z && (7.67)\\
&\Leftrightarrow \quad \forall\, x \in X(\Omega):\ E^{X=x}(Y \mid Z) \text{ is unbiased} && (7.68)\\
&\Rightarrow \quad E(Y \mid X, Z) \text{ is unbiased.} && (7.69)
\end{aligned}$$

For Z being a constant, Corollary 7.34 immediately yields the following corollary.

Corollary 7.35 (Implications of $\tau \perp\!\!\!\perp X$)
Let the Assumptions 6.1 hold and assume that τ_x is P-unique for all $x \in X(\Omega)$. Then

$$\begin{aligned}
\tau \perp\!\!\!\perp X \quad &\Rightarrow \quad \forall\, x \in X(\Omega):\ \tau_x \perp\!\!\!\perp X && (7.70)\\
&\Rightarrow \quad \forall\, x \in X(\Omega):\ \tau_x \perp\!\!\!\perp 1_{X=x} && (7.71)\\
&\Rightarrow \quad \forall\, x \in X(\Omega):\ \tau_x \vdash 1_{X=x} && (7.72)\\
&\Leftrightarrow \quad \forall\, x \in X(\Omega):\ E(Y \mid X{=}x) \text{ is unbiased} && (7.73)\\
&\Rightarrow \quad E(Y \mid X) \text{ is unbiased.} && (7.74)
\end{aligned}$$

Remark 7.36 (A Note on Testability) As mentioned before, unbiasedness of the conditional expectations $E(Y \mid X)$ and $E(Y \mid X, Z)$ cannot be tested in empirical applications, because this involves the true outcome variables τ_x, which cannot be estimated for more than one single value x of X (see the fundamental problem of causal inference described in the Preface). The same applies to the other causality conditions described in this section, including the most restrictive one, strong ignorability. Strong ignorability is interesting from a theoretical perspective, because it implies unbiasedness of the conditional expectations $E^{X=x}(Y \mid Z)$ and $E(Y \mid X, Z)$, provided that we presume P-uniqueness of the true outcome variables τ_x and that Z is a covariate of X. However, unlike the causality conditions treated in chapters 8 to **??**, they cannot be tested empirically without unrealistic additional assumptions, and this is why they cannot be used for covariate selection. ◁

Box 7.1 summarizes all causality conditions, their definitions and their symbols, treated in this chapter, and Tables 7.5 and 7.6 display the implication relations between all these conditions. Note that the propositions about the implications summarized in Table 7.5 are special cases of the corresponding propositions summarized in Table 7.6, which are proved in Exercise 7-3.

Box 7.1 Causality conditions involving the true outcome variables

Let $\big((\Omega,\mathscr{A},P),(\mathscr{F}_t,t\in T),X,Y\big)$ be a causality space, let Y be real-valued with $E(Y^2)<\infty$, let X be discrete with a finite number of values, $P(X{=}x)>0$ for all $x\in X(\Omega)=\{0,1,\ldots,J\}$. Furthermore, let C_X be a global potential confounder of X, let Z be a covariate of X, assume that τ_x is P-unique for all $x\in X(\Omega)$, and define $\tau:=(\tau_0,\tau_1,\ldots,\tau_J)$.

Each of the following four conditions (listed by symbol and meaning) implies that $E(Y|X{=}x)$ is unbiased:

$\tau_x \vdash 1_{X=x}$	$E(\tau_x \mid 1_{X=x}) = E(\tau_x)$.
$\tau_x \perp\!\!\!\perp 1_{X=x}$	$P(X{=}x \mid \tau_x) \underset{P}{=} P(X{=}x)$.
$\tau_x \vdash X$	$E(\tau_x \mid X) = E(\tau_x)$.
$\tau_x \perp\!\!\!\perp X$	$\forall\, x'\in X(\Omega)\colon\ P(X{=}x' \mid \tau_x) \underset{P}{=} P(X{=}x')$.

Each of the following five conditions implies that $E(Y|X)$ is unbiased:

$\forall x\colon \tau_x \vdash 1_{X=x}$	$\forall\, x\in X(\Omega)\colon\ E(\tau_x \mid 1_{X=x}) = E(\tau_x)$.
$\forall x\colon \tau_x \perp\!\!\!\perp 1_{X=x}$	$\forall\, x\in X(\Omega)\colon\ P(X{=}x \mid \tau_x) \underset{P}{=} P(X{=}x)$.
$\forall x\colon \tau_x \perp\!\!\!\perp X$	$\forall\, x,x'\in X(\Omega)\colon\ P(X{=}x' \mid \tau_x) \underset{P}{=} P(X{=}x')$.
$\tau \vdash X$	$\forall\, x\in X(\Omega)\colon\ E(\tau_x \mid X) \underset{P}{=} E(\tau_x)$.
$\tau \perp\!\!\!\perp X$	$\forall\, x\in X(\Omega)\colon\ P(X{=}x \mid \tau) \underset{P}{=} P(X{=}x)$.

Each of the following four conditions implies that $E^{X=x}(Y|Z)$ is unbiased:

$\tau_x \vdash 1_{X=x} \mid Z$	$E(\tau_x \mid 1_{X=x}, Z) \underset{P}{=} E(\tau_x \mid Z)$.
$\tau_x \perp\!\!\!\perp 1_{X=x} \mid Z$	$P(X{=}x \mid \tau_x, Z) \underset{P}{=} P(X{=}x \mid Z)$.
$\tau_x \vdash X \mid Z$	$E(\tau_x \mid X, Z) \underset{P}{=} E(\tau_x \mid Z)$.
$\tau_x \perp\!\!\!\perp X \mid Z$	$P(X{=}x' \mid \tau_x, Z) \underset{P}{=} P(X{=}x' \mid Z), \quad \forall\, x'\in X(\Omega)$.

Each of the following five conditions implies that $E(Y|X,Z)$ is unbiased:

$\forall x\colon\ \tau_x \vdash 1_{X=x} \mid Z$	$\forall\, x\in X(\Omega)\colon\ E(\tau_x \mid 1_{X=x}, Z) \underset{P}{=} E(\tau_x \mid Z)$.
$\forall x\colon\ \tau_x \perp\!\!\!\perp 1_{X=x} \mid Z$	$\forall\, x\in X(\Omega)\colon\ P(X{=}x \mid \tau_x, Z) \underset{P}{=} P(X{=}x \mid Z)$.
$\forall x\colon\ \tau_x \perp\!\!\!\perp X \mid Z$	$\forall\, x,x'\in X(\Omega)\colon\ P(X{=}x' \mid \tau_x, Z) \underset{P}{=} P(X{=}x' \mid Z)$.
$\tau \vdash X \mid Z$	$\forall\, x\in X(\Omega)\colon\ E(\tau_x \mid X, Z) \underset{P}{=} E(\tau_x \mid Z)$.
$\tau \perp\!\!\!\perp X \mid Z$	$\forall\, x\in X(\Omega)\colon\ P(X{=}x \mid \tau, Z) \underset{P}{=} P(X{=}x \mid Z)$.

Table 7.5. Implication relations among causality conditions for $E(Y|X{=}x)$ and $E(Y|X)$

	$\tau_x \vdash 1_{X=x}$	$\tau_x \perp\!\!\!\perp 1_{X=x}$	$\tau_x \vdash X$	$\tau_x \perp\!\!\!\perp X$	$\forall x{:}\ \tau_x \vdash 1_{X=x}$	$\forall x{:}\ \tau_x \perp\!\!\!\perp 1_{X=x}$	$\tau \vdash X$	$\forall x{:}\ \tau_x \perp\!\!\!\perp X$	$\tau \perp\!\!\!\perp X$
$\tau_x \vdash 1_{X=x}$	$\Leftrightarrow$								
$\tau_x \perp\!\!\!\perp 1_{X=x}$	$\Rightarrow$	$\Leftrightarrow$							
$\tau_x \vdash X$	$\Rightarrow$		$\Leftrightarrow$						
$\tau_x \perp\!\!\!\perp X$	$\Rightarrow$	$\Rightarrow$	$\Rightarrow$	$\Leftrightarrow$					
$\forall x{:}\ \tau_x \vdash 1_{X=x}$	$\Rightarrow$				$\Leftrightarrow$				
$\forall x{:}\ \tau_x \perp\!\!\!\perp 1_{X=x}$	$\Rightarrow$	$\Rightarrow$			$\Rightarrow$	$\Leftrightarrow$			
$\tau \vdash X$	$\Rightarrow$		$\Rightarrow$		$\Rightarrow$		$\Leftrightarrow$		
$\forall x{:}\ \tau_x \perp\!\!\!\perp X$	$\Rightarrow$	$\Rightarrow$	$\Rightarrow$	$\Rightarrow$	$\Rightarrow$	$\Rightarrow$	$\Rightarrow$	$\Leftrightarrow$	
$\tau \perp\!\!\!\perp X$	$\Rightarrow$	$\Rightarrow$	$\Rightarrow$	$\Rightarrow$	$\Rightarrow$	$\Rightarrow$	$\Rightarrow$	$\Rightarrow$	$\Leftrightarrow$

Note: An entry $\Rightarrow$ ($\Leftrightarrow$) means that the condition in the row implies (is equivalent to) the condition in the column, provided that the Assumptions 6.1 hold. The symbols involving $\vdash$ and $\perp\!\!\!\perp$ are explained in Box 7.1. Presuming P-uniqueness of τ_x, the first four conditions imply unbiasedness of $E(Y|X{=}x)$ and the last five imply unbiasedness of $E(Y|X)$, provided that P-uniqueness of τ_x holds for all $x \in X(\Omega)$.

7.5 Summary and Conclusions

In this chapter, we treated some causality conditions, all of which are listed in Box 7.1. The implication relations among these causality conditions are listed in Tables 7.5 and 7.6. According to the last row of Table 7.6, Rosenbaum and Rubin's strong ignorability condition is the strongest (i. e., the most restrictive) condition; it implies all other causality conditions treated in this chapter. Note that there are no implications between the causality conditions summarized in Table 7.5 and those listed in Table 7.6. For example, $\tau \perp\!\!\!\perp X$, does not imply any of the causality conditions listed in Table 7.6. This implies, for example, that $\forall x{:}\ \tau_x \vdash 1_{X=x}$, which is equivalent to unbiasedness of $E(Y|X)$, does not imply unbiasedness of $E(Y|X,Z)$ (see also the counter example described in section 6.6).

Unfortunately, just like unbiasedness, all causality conditions treated in this chapter, including strong ignorabiliy, cannot be tested empirically without unrealistic additional assumptions, because, unlike in the examples treated in this chapter, in empirical applications, the values of the true outcome variables τ_x are unknown and cannot be estimated. Therefore, these causality conditions cannot be used for covariate selection (see also Rem. 7.36). Nevertheless, these causality conditions are of theoretical interest, because they are implied by other causality conditions that are empirically testable (see chs. 8 to **??**).

Table 7.6. Implication relations among causality conditions for $E^{X=x}(Y|Z)$ and $E(Y|X,Z)$

	$Z \mid \tau_x \vdash 1_{X=x}$	$Z \mid \tau_x \perp\!\!\!\perp 1_{X=x}$	$Z \mid \tau_x \vdash X$	$Z \mid \tau_x \perp\!\!\!\perp X$	$Z \mid \tau_x \vdash 1_{X=x}$:$\forall x$	$Z \mid \tau_x \perp\!\!\!\perp 1_{X=x}$:$\forall x$	$Z \mid \tau \vdash X$	$Z \mid \tau_x \perp\!\!\!\perp X$:$\forall x$	$Z \mid \tau \perp\!\!\!\perp X$
$\tau_x \vdash 1_{X=x} \mid Z$	$\Leftrightarrow$								
$\tau_x \perp\!\!\!\perp 1_{X=x} \mid Z$	$\Rightarrow$	$\Leftrightarrow$							
$\tau_x \vdash X \mid Z$	$\Rightarrow$		$\Leftrightarrow$						
$\tau_x \perp\!\!\!\perp X \mid Z$	$\Rightarrow$	$\Rightarrow$	$\Rightarrow$	$\Leftrightarrow$					
$\forall x\colon \tau_x \vdash 1_{X=x} \mid Z$	$\Rightarrow$				$\Leftrightarrow$				
$\forall x\colon \tau_x \perp\!\!\!\perp 1_{X=x} \mid Z$	$\Rightarrow$	$\Rightarrow$			$\Rightarrow$	$\Leftrightarrow$			
$\tau \vdash X \mid Z$	$\Rightarrow$		$\Rightarrow$		$\Rightarrow$		$\Leftrightarrow$		
$\forall x\colon \tau_x \perp\!\!\!\perp X \mid Z$	$\Rightarrow$	$\Rightarrow$	$\Rightarrow$	$\Rightarrow$	$\Rightarrow$	$\Rightarrow$	$\Rightarrow$	$\Leftrightarrow$	
$\tau \perp\!\!\!\perp X \mid Z$	$\Rightarrow$	$\Rightarrow$	$\Rightarrow$	$\Rightarrow$	$\Rightarrow$	$\Rightarrow$	$\Rightarrow$	$\Rightarrow$	$\Leftrightarrow$

Note: An entry $\Rightarrow$ ($\Leftrightarrow$) means that the condition in the row implies (is equivalent to) the condition in the column, provided that the Assumptions 6.1 hold. The symbols involving $\vdash$ and $\perp\!\!\!\perp$ are explained in Box 7.1. Presuming P-uniqueness of τ_x and that Z is a covariate of X, the first four conditions imply unbiasedness of $E^{X=x}(Y|Z)$ and the last five imply unbiasedness of $E(Y|X,Z)$, provided that P-uniqueness of τ_x holds for all $x \in X(\Omega)$.

7.6 Proofs

Proof of Lemma 7.4

For any σ-algebra $\mathscr{C} \subset \mathscr{A}$,

$$1 \underset{P}{=} P(X \in \Omega'_X | \mathscr{C}) \underset{P}{=} P(X{=}x | \mathscr{C}) + P(X{\neq}x | \mathscr{C}).$$

Therefore,

$$P(X{\neq}x | \mathscr{C}) \underset{P}{=} 1 - P(X{=}x | \mathscr{C}). \tag{7.75}$$

Hence,

$$\begin{aligned} P(X{\neq}x | W, Z) &\underset{P}{=} 1 - P(X{=}x \mid W, Z) && [(7.75) \text{ with } \mathscr{C} = \sigma(W,Z)] \\ &\underset{P}{=} 1 - P(X{=}x \mid Z) && [(a)] \\ &\underset{P}{=} P(X{\neq}x | Z). && [(7.75) \text{ with } \mathscr{C} = \sigma(Z)] \end{aligned}$$

According to SN-Remark 16.27, in conjunction with (a), this equation is equivalent to (b), because

$$\begin{aligned} P(X{=}x | W, Z) &\underset{P}{=} P(1_{X=x}{=}1 | W, Z) \\ P(X{\neq}x | W, Z) &\underset{P}{=} P(1_{X=x}{=}0 | W, Z) \\ P(X{=}x | Z) &\underset{P}{=} P(1_{X=x}{=}1 | Z) \end{aligned}$$

$$P(X{\neq}x\,|\,Z) \underset{P}{=} P(1_{X=x}{=}0\,|\,Z).$$

These equations follow from the definitions of these Z-conditional probabilities as the Z-conditional probabilities of the events

$$A = \{X{=}x\} = \{1_{X=x}{=}1\} = \{\omega \in \Omega\colon X(\omega) = x\}$$

and

$$A^c = \{X{\neq}x\} = \{1_{X=x}{=}0\} = \{\omega \in \Omega\colon X(\omega) \neq x\}$$

(see SN-Rem. 10.4).

Proof of Theorem 7.11

(i)

$$\begin{aligned}
E(\tau_x\,|\,1_{X=x}) &\underset{P}{=} E\big(E(\tau_x\,|\,X)\,\big|\,1_{X=x}\big) && \text{[SN-Box 10.2 (v)]}\\
&\underset{P}{=} E\big(E(\tau_x)\,\big|\,1_{X=x}\big) && \text{[(7.10)]}\\
&\underset{P}{=} E(\tau_x). && \text{[SN-Box 10.2 (i)]}
\end{aligned}$$

If we presume that τ_x is P-unique, then, according to Theorem 6.9, this equation is equivalent to unbiasedness of $E(Y\,|\,X{=}x)$

(ii) This proposition immediately follows from (i), (7.16), and Definition 6.11 (iii).

Proof of Theorem 7.18

aa

Proof of Theorem 7.25

(i)

$$\begin{aligned}
& P(X{=}x\,|\,Z,\tau_x) \underset{P}{=} P(X{=}x\,|\,Z) && \\
\Leftrightarrow\ & 1_{X=x} \perp\!\!\!\perp \tau_x\,|\,Z && \text{[Lemma 7.4]}\\
\Rightarrow\ & E(\tau_x\,|\,1_{X=x},Z) \underset{P}{=} E(\tau_x\,|\,Z) && \text{[τ_x is P-unique, SN-Rem. 16.35]}\\
\Leftrightarrow\ & \tau_x \vdash 1_{X=x}\,|\,Z && \text{[(6.23)]}\\
\Leftrightarrow\ & E^{X=x}(Y\,|\,Z) \text{ is unbiased.} && \text{[τ_x is P-unique, Th. 6.15]}
\end{aligned}$$

Note that assuming P-uniqueness of τ_x in the third line is necessary if this proposition is assumed to hold for all versions $\tau_x \in \mathscr{E}^{X=x}(Y\,|\,C_X)$ and not only for a fixed version.

(ii) This proposition immediately follows from (i), (7.46), and Definition 6.11 (iii).

Proof of Corollary 7.29

(i) If τ_x is P-unique, then $\tau_x \perp\!\!\!\perp X\,|\,Z \Rightarrow \tau_x \perp\!\!\!\perp 1_{X=x}\,|\,Z$ follows from SN-Box 16.3 (vi) and $\tau_x \perp\!\!\!\perp X\,|\,Z \Rightarrow \tau_x \vdash X\,|\,Z$ follows from SN-Rem. 16.35.

(ii) This proposition immediately follows from (i), (7.55), and Definition 6.11 (iii).

7.7 Exercises

▷ **Exercise 7-1** . Show that $P(X{=}x) > 0$ follows from P-uniqueness of τ_x.

▷ **Exercise 7-2** Show that $P(X{=}x \,|\, C_X) \gtrsim_P 0$ implies $P(X{=}x \,|\, Z) \gtrsim_P 0$, if Z is a covariate of X.

▷ **Exercise 7-3** Prove the implications listed in Table 7.6.

Solutions

▷ **Solution 7-1** In the definition of $\tau_x \underset{P}{=} E^{X=x}(Y|C_X)$ we assume $P(X{=}x) > 0$ (see Defs. 3.79 and 4.16).

▷ **Solution 7-2** Consider

$$\begin{aligned} P(X{=}x \,|\, Z) &\underset{P}{=} E(1_{X=x} \,|\, Z) && \text{[SN-(10.4)]} \\ &\underset{P}{=} E\big(E(1_{X=x} \,|\, C_X) \,\big|\, Z\big) && \text{[SN-Box 10.2 (v)]} \\ &\underset{P}{=} E\big(P(X{=}x \,|\, C_X) \,\big|\, Z\big). && \text{[SN-(10.4)]} \end{aligned} \tag{7.76}$$

Hence,

$$\begin{aligned} P(X{=}x \,|\, C_X) \gtrsim_P 0 \quad &\Rightarrow \quad E\big(P(X{=}x \,|\, C_X) \,\big|\, Z\big) \gtrsim_P 0 && \text{[(3.48)]} \\ &\Rightarrow \quad P(X{=}x \,|\, Z) \gtrsim_P 0. && \text{[SN-(2.40), (7.76)]} \end{aligned}$$

▷ **Solution 7-3**

(1) $\tau_x \perp\!\!\!\perp 1_{X=x} \,|\, Z \;\Rightarrow\; \tau_x \vdash 1_{X=x} \,|\, Z$. This is the proposition of Theorem 7.25 (i).

(2) $\tau_x \vdash X \,|\, Z \;\Rightarrow\; \tau_x \vdash 1_{X=x} \,|\, Z$. This is the proposition of Theorem 7.18 (i).

(3) $\tau_x \perp\!\!\!\perp X \,|\, Z \;\Rightarrow\; \tau_x \vdash 1_{X=x} \,|\, Z$. This proposition follows from the conjunction of Corollary 7.29 and Theorem 7.18 (i).

(4) $(\forall x\colon \tau_x \vdash 1_{X=x} \,|\, Z) \;\Rightarrow\; (\tau_x \vdash 1_{X=x} \,|\, Z)$. This proposition is trivial.

(5) $(\forall x\colon \tau_x \perp\!\!\!\perp 1_{X=x} \,|\, Z) \;\Rightarrow\; (\tau_x \vdash 1_{X=x} \,|\, Z)$. This proposition immediately follows from (1).

(6) $(\forall x\colon \tau_x \perp\!\!\!\perp X \,|\, Z) \;\Rightarrow\; (\tau_x \vdash 1_{X=x} \,|\, Z)$. This proposition immediately follows from (3).

(7) $\tau \vdash X \,|\, Z \;\Rightarrow\; \tau_x \vdash 1_{X=x} \,|\, Z$. This is the proposition of Theorem 7.18 (i).

(8) $\tau \perp\!\!\!\perp X \,|\, Z \;\Rightarrow\; \tau_x \vdash 1_{X=x} \,|\, Z$. This proposition is contained in Corollary 7.34.

(9) $\tau_x \perp\!\!\!\perp X \,|\, Z \;\Rightarrow\; \tau_x \perp\!\!\!\perp 1_{X=x} \,|\, Z$. This proposition is contained in the proposition of Corollary 7.29.

(10) $(\forall x\colon \tau_x \perp\!\!\!\perp 1_{X=x} \,|\, Z) \;\Rightarrow\; (\tau_x \perp\!\!\!\perp 1_{X=x} \,|\, Z)$. This proposition is trivial.

(11) $(\forall x\colon \tau_x \perp\!\!\!\perp X \,|\, Z) \;\Rightarrow\; (\tau_x \perp\!\!\!\perp 1_{X=x} \,|\, Z)$. This proposition immediately follows from (9).

(12) $\tau \perp\!\!\!\perp X \,|\, Z \;\Rightarrow\; (\tau_x \perp\!\!\!\perp 1_{X=x} \,|\, Z)$. This proposition is contained in Corollary 7.34.

(13) $\tau_x \perp\!\!\!\perp X \,|\, Z \;\Rightarrow\; \tau_x \vdash X \,|\, Z$. This proposition is contained in Corollary 7.29.

(14) $(\forall x\colon \tau_x \perp\!\!\!\perp X \,|\, Z) \;\Rightarrow\; \tau_x \vdash X \,|\, Z$. This proposition immediately follows from (13).

(15) $\tau \vdash X|Z \;\Rightarrow\; \tau_x \vdash X|Z$. This proposition immediately follows from the definition of $\tau \vdash X|Z$.

(16) $(\forall x\colon \tau_x \perp\!\!\!\perp X \,|\, Z) \;\Rightarrow\; \tau_x \vdash X \,|\, Z$. This proposition immediately follows from (13).

(17) $\tau \perp\!\!\!\perp X \,|\, Z \;\Rightarrow\; \tau_x \vdash X \,|\, Z$. This proposition from Corollaries 7.34 and 7.29.

(18) $(\forall x\colon \tau_x \perp\!\!\!\perp X \,|\, Z) \;\Rightarrow\; \tau_x \perp\!\!\!\perp X \,|\, Z$. This proposition is trivial.

(19) $\tau \perp\!\!\!\perp X \,|\, Z \;\Rightarrow\; \tau_x \perp\!\!\!\perp X \,|\, Z$. This proposition is contained in Corollary 7.34.

(20) $(\forall x\colon \tau_x \perp\!\!\!\perp 1_{X=x} \,|\, Z) \;\Rightarrow\; (\forall x\colon \tau_x \vdash 1_{X=x} \,|\, Z)$. This is the proposition of Theorem 7.25.

(21) $\tau \vdash X \mid Z \Rightarrow (\forall x\colon \tau_x \vdash 1_{X=x} \mid Z)$. This proposition is contained in Proposition (7.33).

(22) $(\forall x\colon \tau_x \perp\!\!\!\perp X \mid Z) \Rightarrow (\forall x\colon \tau_x \vdash 1_{X=x} \mid Z)$. This is the proposition of Corollary 7.29 (ii).

(23) $\tau \perp\!\!\!\perp X \mid Z \Rightarrow (\forall x\colon \tau_x \vdash 1_{X=x} \mid Z)$. This proposition follows from Corollaries 7.34 and 7.29 (ii).

(24) $(\forall x\colon \tau_x \perp\!\!\!\perp X \mid Z) \Rightarrow (\forall x\colon \tau_x \perp\!\!\!\perp 1_{X=x} \mid Z)$. This is the proposition of Corollary 7.29 (ii).

(25) $\tau \perp\!\!\!\perp X \mid Z \Rightarrow (\forall x\colon \tau_x \perp\!\!\!\perp 1_{X=x} \mid Z)$. This proposition is contained in Corollary 7.34.

(26) $(\forall x\colon \tau_x \perp\!\!\!\perp X \mid Z) \Rightarrow \tau \vdash X \mid Z$. This proposition follows from Corollary 7.29 (ii) and Proposition (7.33).

(27) $\tau \perp\!\!\!\perp X \mid Z \Rightarrow \tau \vdash X \mid Z$. This proposition follows from Corollaries 7.34 and 7.29 (ii).

(28) $\tau \perp\!\!\!\perp X \mid Z \Rightarrow (\forall x\colon \tau_x \perp\!\!\!\perp X \mid Z)$. This proposition is contained in Corollary 7.34.

Chapter 8

Fisher Conditions

In chapter 6, we defined *unbiasedness* of the conditional expectation values $E(Y|X{=}x)$ and $E(Y|X{=}x,Z{=}z)$, as well as unbiasedness of the conditional expectations $E(Y|X)$, $E^{X=x}(Y|Z)$, and $E(Y|X,Z)$. These kinds of unbiasedness constitute a first kind of causality conditions. For example, if $P(X{=}x), P(X{=}x') > 0$ and $E(Y|X{=}x)$ as well as $E(Y|X{=}x')$ are unbiased, then the prima facie effect $PFE_{xx'} = E(Y|X{=}x) - E(Y|X{=}x')$ is unbiased and identical to the *causal average total effect* $ATE_{xx'}$ (see Cor. 6.23). Therefore, under these assumptions, an estimate of $PFE_{xx'}$ is also an estimate of the $ATE_{xx'}$. Similarly, if $P(X{=}x,Z{=}z)$, $P(X{=}x',Z{=}z) > 0$, and $E(Y|X{=}x,Z{=}z)$ as well as $E(Y|X{=}x',Z{=}z)$ are unbiased, then the $(Z{=}z)$-conditional prima facie effect $PFE_{Z;xx'}(z) = E(Y|X{=}x,Z{=}z) - E(Y|X{=}x',Z{=}z)$ is unbiased, and $PFE_{Z;xx'}(z)$ is identical to $CTE_{Z;xx'}(z)$, the *causal* $(Z{=}z)$*-conditional total effect* (see Cor. 6.34).

In chapter 7, we treated some other conditions that imply unbiasedness of the conditional expectation values $E(Y|X{=}x)$ and $E(Y|X{=}x,Z{=}z)$, and of the conditional expectations $E(Y|X)$, $E^{X=x}(Y|Z)$ and $E(Y|X,Z)$. However, these conditions as well as unbiasedness itself cannot be tested empirically. This implies that none of these conditions can be used for covariate selection. Furthermore, unbiasedness can be *accidental* in the sense that it may hold for $E(Y|X)$ but not for $E(Y|X,Z)$, where Z is a covariate of X. This also applies to Rosenbaum and Rubin's strong ignorability. This is one of the reasons why, in this chapter, we study other causality conditions that are less volatile.

8.1 Fisher Conditions

The conditions to be treated in this section [see Eqs. (8.1) to (8.4)] can be created by randomized assignment of the unit to one of several treatment conditions that are represented by the values x of X. This class of conditions will be referred to as the *Fisher conditions* (for total effects). This name is chosen to acknowledge the contributions of Sir R. A. Fisher to understanding the relevance of the design technique of randomization for causal inference (see, e. g., Fisher, 1925/1946). In contrast to the causality conditions treated in the previous chapters, the Fisher conditions are empirically testable. Furthermore, they are not accidental in the sense described above. Finally, they may also hold if X is a continuous random variable, whereas all causality conditions treated so far and even the true outcome variables are not defined if X is continuous.

8.1.1 Independence of C_X and X

In this section and in other parts of the chapter and the book, we will often refer to the following assumptions and notation.

Notation and Assumptions 8.1
Let $\big((\Omega,\mathscr{A},P),(\mathscr{F}_t,t\in T),X,Y\big)$ be a causality space, let C_X be a global potential confounder of X, and let $x\in\Omega'_X$ be a value of X.

Note that in this set of assumptions we neither assume that X is discrete nor that Y is numerical. Later, we will add these assumptions when they are necessary to prove some propositions.

Remark 8.2 (Independence of C_X and X if X is Discrete) If the Assumptions 8.1 hold, then, according to Lemma 7.4, $P(X{=}x\,|\,C_X)\underset{P}{=}P(X{=}x)$ is equivalent to independence of C_X and $1_{X=x}$, that is,

$$C_X \perp\!\!\!\perp 1_{X=x} \quad\Leftrightarrow\quad P(X{=}x\,|\,C_X)\underset{P}{=}P(X{=}x). \tag{8.1}$$

Furthermore, if we additionally assume that X is discrete, then, according to SN-Theorem 16.26 and SN-Remark 16.27,

$$C_X \perp\!\!\!\perp X \quad\Leftrightarrow\quad \forall\, x\in X(\Omega)\colon\ P(X{=}x\,|\,C_X)\underset{P}{=}P(X{=}x). \tag{8.2}$$

If X is discrete with $P(X{=}x)>0$, for all $x\in X(\Omega)$, then $C_X\perp\!\!\!\perp X$ implies $\tau\perp\!\!\!\perp X$ (see Th. 8.14), and with it all causality conditions that follow from $\tau\perp\!\!\!\perp X$ (see Table 7.5). ◁

Remark 8.3 (Continuous X) Note, however, that $C_X\perp\!\!\!\perp X$ is also defined and may hold if X is not discrete [see Def. 7.1 (i)]. In this case, under the Assumptions 8.1, $C_X\perp\!\!\!\perp X$ is still a causality condition. It guarantees that $E(Y|X)$ describes a *causal* dependence of Y on X, because $C_X\perp\!\!\!\perp X$ implies that the distributions of all potential confounders do not depend on X. For example, if W is a discrete potential confounder of X, then $C_X\perp\!\!\!\perp X$ implies $W\perp\!\!\!\perp X$, which is equivalent to

$$P(W{=}w\,|\,X)\underset{P}{=}P(W{=}w),\quad \forall\, w\in W(\Omega)$$

(see again SN-Th. 16.26 and SN-Rem. 16.27). More general and in more formal terms, $C_X\perp\!\!\!\perp X$ is equivalent to $P_W\in\mathscr{P}_{W|X}$, for all C_X-measurable random variables W. This means, for each potential confounder W, the distribution P_W of W is also a version of the X-conditional distribution of W (see SN-Th. 17.45 and SN-Rem. 17.46). Hence, we may call $C_X\perp\!\!\!\perp X$ the *ceteris distributionibus paribus* clause. ◁

8.1.2 Z-Conditional Independence of C_X and X

In this section and in other parts of the chapter and the book, we will also often refer to the following assumptions and notation. Compared to the Assumptions 8.1 we just add that Z denotes a (possibly multivariate) covariate of X.

Notation and Assumptions 8.4
Let $\big((\Omega,\mathscr{A},P),(\mathscr{F}_t,t\in T),X,Y\big)$ *be a causality space, let* C_X *be a global potential confounder of* X*, let* $x\in\Omega'_X$ *be a value of* X*, and let* Z *be a covariate of* X *with value space* $(\Omega'_Z,\mathscr{A}'_Z)$*.*

We denote Z-conditional independence of C_X and X by $C_X \perp\!\!\!\perp X \,|\, Z$ [see Def. 7.1 (ii)]. For details on conditional independence see SN-chapter 16. Here, we review some properties of conditional independence in the case in which X is discrete.

Remark 8.5 (Z-Conditional Independence of C_X and a Discrete X) If X is discrete, then, under the Assumptions 8.4, the conditional versions of Propositions (8.1) and (8.2) are

$$C_X \perp\!\!\!\perp 1_{X=x} \,|\, Z \quad\Leftrightarrow\quad P(X{=}x\,|\,C_X,Z) \underset{P}{=} P(X{=}x\,|\,Z) \tag{8.3}$$

(see Lemma 7.4) and

$$C_X \perp\!\!\!\perp X \,|\, Z \quad\Leftrightarrow\quad \forall\, x\in X(\Omega)\colon\ P(X{=}x\,|\,C_X,Z) \underset{P}{=} P(X{=}x\,|\,Z) \tag{8.4}$$

(see SN-Theorem 16.26 and SN-Remark 16.27), respectively. Note that (8.1) is a special case of (8.3) for Z being a constant mapping, that is, for $Z(\omega) = const$, for all $\omega\in\Omega$. For the same reason (8.2) is a special case of (8.4).

According to Assumptions 8.4, Z is a covariate of X. This implies that $\sigma(Z)\subset\sigma(C_X)$ holds for the σ-algebras generated by these two random variables (see Def. 4.4 and Rem. 4.9), and that $\sigma(C_X)=\sigma(C_X,Z)$. Therefore,

$$P(X{=}x\,|\,C_X,Z) \underset{P}{=} P(X{=}x\,|\,C_X) \tag{8.5}$$

(see SN-sections 2.3.2 and 10.1). Hence, if X is discrete, then

$$C_X \perp\!\!\!\perp 1_{X=x} \,|\, Z \quad\Leftrightarrow\quad P(X{=}x\,|\,C_X) \underset{P}{=} P(X{=}x\,|\,Z) \tag{8.6}$$

and

$$C_X \perp\!\!\!\perp X \,|\, Z \quad\Leftrightarrow\quad \forall\, x\in X(\Omega)\colon\ P(X{=}x\,|\,C_X) \underset{P}{=} P(X{=}x\,|\,Z). \tag{8.7}$$

As will be shown in Theorem 8.21, $C_X \perp\!\!\!\perp X \,|\, Z$ implies $\tau \perp\!\!\!\perp X \,|\, Z$ (strong ignorability), and with it all other causality conditions that follow from $\tau \perp\!\!\!\perp X \,|\, Z$ (see Table 7.6). ◁

Remark 8.6 (Randomized Assignment of a Unit to a Treatment) Because we assume that Z is C_X-measurable, $C_X \perp\!\!\!\perp X$ implies $C_X \perp\!\!\!\perp X \,|\, Z$ [see Prop. (7.7) and SN-Box 16.2 (ix)]. As already mentioned, $C_X \perp\!\!\!\perp X$ is created by randomized assignment (e. g., via a coin flip) of a unit to a treatment, because, (a) by definition of randomized assignment, X deterministically depends on the coin flip whose outcome does not depend on C_X, and (b) a global potential confounder C_X is defined such that it is a random variable that cannot be affected by the treatment variable X (see section 4.1). ◁

Remark 8.7 (Z-Conditional Randomization) Also note that we can conduct experiments in which we create $C_X \perp\!\!\!\perp X \,|\, Z$ but not $C_X \perp\!\!\!\perp X$. For example, Z-conditional independence of X and the global potential confounder C_X may be created via Z-conditional randomization, where assignment to treatment is arranged such that the conditional probabilities $P(X{=}x\,|\,Z{=}z)$ differ for different values z of Z, but $C_X \perp\!\!\!\perp X \,|\, Z$ still holds. A numerical example has already been presented in Table 6.3. ◁

Remark 8.8 (Falsifiability of the Fisher Conditions) The Fisher conditions are empirically falsifiable. More specifically, if W is measurable with respect to C_X, then $C_X \perp\!\!\!\perp X$ implies $W \perp\!\!\!\perp X$ (see Exercise 8-1), which is a shortcut for

$$P(X{=}x \mid W) \underset{P}{=} P(X{=}x) \quad \forall\, x \in X(\Omega)\,. \tag{8.8}$$

Hence, in order to test $C_X \perp\!\!\!\perp X$, we simply have to select a C_X-measurable random variable W and see if Equation (8.8) holds. If Equation (8.8) does not hold, then we can conclude that $C_X \perp\!\!\!\perp X$ does not hold as well.

Similarly, if W is measurable with respect to C_X, then $C_X \perp\!\!\!\perp X \mid Z$ implies $W \perp\!\!\!\perp X \mid Z$, which is a shortcut for

$$P(X{=}x \mid Z, W) \underset{P}{=} P(X{=}x \mid Z), \quad \forall\, x \in X(\Omega)\,. \tag{8.9}$$

Therefore, in order to test $C_X \perp\!\!\!\perp X \mid Z$, we can select a C_X-measurable random variable W and see if Equation (8.9) holds. If it does not hold, then we can conclude that $C_X \perp\!\!\!\perp X \mid Z$ does not hold as well. In empirical applications such tests of independence or conditional independence can be conducted, for example, via logistic regression analysis. ◁

Example 8.9 (First Examples) A first example of independence of the cause X and a global potential confounder C_X has already been presented in Table 6.2 (p. 139). Furthermore, Table 6.3 (p. 140) displays an example for Z-conditional independence of C_X and X. In both examples, U takes the role of a global potential confounder C_X. In section 8.3 we will treat several examples in more detail. ◁

Remark 8.10 (Continuous X) Again, note that $C_X \perp\!\!\!\perp X \mid Z$ is also defined and may hold if X is not discrete. In this case, $C_X \perp\!\!\!\perp X \mid Z$ is still a causality condition. It guarantees that $E(Y \mid X, Z)$ describes a *causal* dependence of Y on X, for reasons that are similar to those mentioned for $C_X \perp\!\!\!\perp X$. For example, if W is a discrete potential confounder of X, then $C_X \perp\!\!\!\perp X \mid Z$ implies $W \perp\!\!\!\perp X \mid Z$, which is equivalent to

$$P(W{=}w \mid X, Z) \underset{P}{=} P(W{=}w \mid Z), \quad \forall\, w \in W(\Omega)$$

(see again SN-Th. 16.26 and SN-Rem. 16.27). ◁

8.2 Implications of the Fisher Conditions

8.2.1 Implications of Independence of C_X and X

Now we return to the assumptions that we used in the previous chapters. In particular we presume that X is discrete with $P(X{=}x) > 0$ for all $x \in X(\Omega)$. In the following theorem we summarize the implications of $C_X \perp\!\!\!\perp 1_{X=x}$, which is equivalent to

$$P(X{=}x \mid C_X) \underset{P}{=} P(X{=}x) \tag{8.10}$$

(see Lemma 7.4).

Theorem 8.11 (Implications of $C_X \perp\!\!\!\perp \mathbf{1}_{X=x}$)
Let the Assumptions 5.1 hold. Then $C_X \perp\!\!\!\perp 1_{X=x}$ implies:

(i) τ_x *is P-unique.*
(ii) $\tau_x \perp\!\!\!\perp 1_{X=x}$.
(iii) $\tau_x \vdash 1_{X=x}$.
(iv) $E(Y|X=x)$ *is unbiased.*

If the Assumptions 5.1 hold and Z is a covariate of X, then $C_X \perp\!\!\!\perp 1_{X=x}$ implies:

(v) $C_X \perp\!\!\!\perp 1_{X=x} | Z$.
(vi) $\tau_x \perp\!\!\!\perp 1_{X=x} | Z$.
(vii) $\tau_x \vdash 1_{X=x} | Z$.
(viii) $E^{X=x}(Y|Z)$ *is unbiased.*
(ix) $E(Y|X=x, Z=z)$ *is unbiased, if* $P(X=x, Z=z) > 0$.

(Proof p. 195)

Remark 8.12 ($C_X \perp\!\!\!\perp \mathbf{1}_{X=x}$ Implies That τ_x is P-Unique) Note that $C_X \perp\!\!\!\perp 1_{X=x}$ is the first causality condition implying that τ_x is P-unique. ◁

Remark 8.13 ($C_X \perp\!\!\!\perp \mathbf{1}_{X=x}$ Implies $C_X \perp\!\!\!\perp \mathbf{1}_{X=x} | Z$) Also note that $C_X \perp\!\!\!\perp 1_{X=x}$ is the first causality condition implying $C_X \perp\!\!\!\perp 1_{X=x} | Z$. Hence, $C_X \perp\!\!\!\perp 1_{X=x}$ does not only imply that $E(Y|X=x)$ is unbiased, but it also implies unbiasedness of $E^{X=x}(Y|Z)$, if Z is a covariate of X (and therefore is C_X-measurable). ◁

In Theorem 8.11 we considered the implications of independence of C_X and the indicator $1_{X=x}$ for a single value x of X. In the following theorem, we consider the implications of independence of C_X and X, which, if X is discrete, is equivalent to

$$\forall\, x \in X(\Omega) :\ C_X \perp\!\!\!\perp 1_{X=x}.$$

This in turn implies that also all propositions listed in Theorem 8.11 hold for all values x of X. In the following theorem we list some additional propositions that are not already included in Theorem 8.11.

Theorem 8.14 (Implications of $C_X \perp\!\!\!\perp X$)
Let the Assumptions 6.1 hold. Then $C_X \perp\!\!\!\perp X$ implies:

(i) *For all* $x \in X(\Omega)$: $C_X \perp\!\!\!\perp 1_{X=x}$.
(ii) *For all* $x \in X(\Omega)$: $\tau_x \perp\!\!\!\perp X$.
(iii) $\tau \perp\!\!\!\perp X$, *where* $\tau := (\tau_0, \tau_1, \ldots, \tau_J)$.
(iv) *For all* $x \in X(\Omega)$: $\tau_x \vdash X$.
(v) $E(Y|X)$ *is unbiased.*

If the Assumptions 6.1 hold and Z is a covariate of X, then $C_X \perp\!\!\!\perp X$ implies:

(vi) $C_X \perp\!\!\!\perp X | Z$.
(vii) *For all* $x \in X(\Omega)$: $\tau_x \perp\!\!\!\perp X | Z$.
(viii) $\tau \perp\!\!\!\perp X | Z$.
(ix) *For all* $x \in X(\Omega)$: $\tau_x \vdash X | Z$.

(x) $E(Y|X,Z)$ *is unbiased.*

(Proof p. 196)

Remark 8.15 ($C_X \perp\!\!\!\perp X$ Implies That all τ_x are P-Unique) Note that $C_X \perp\!\!\!\perp X$ implies that all τ_x, $x \in X(\Omega)$, are P-unique. ◁

Remark 8.16 ($C_X \perp\!\!\!\perp X$ Implies $C_X \perp\!\!\!\perp X|Z$) Also note that $C_X \perp\!\!\!\perp X$ implies $C_X \perp\!\!\!\perp X|Z$, because Z is C_X-measurable. Hence, $C_X \perp\!\!\!\perp X$ does not only imply that $E(Y|X)$ is unbiased, but it also implies unbiasedness of $E(Y|X,Z)$. ◁

8.2.2 Unbiasedness of the Prima Facie Effects

The following corollary is an immediate implication of Theorem 8.14. In this corollary we specify a condition under which the prima facie effect

$$PFE_{xx'} = E(Y|X{=}x) - E(Y|X{=}x') \tag{8.11}$$

is unbiased. Remember, if τ_x and $\tau_{x'}$ are P-unique, then $\delta_{xx'} = \tau_x - \tau_{x'}$ is a true total effect variable [see Def. 5.2 (iii)] and

$$ATE_{xx'} = E(\delta_{xx'}) = E(\tau_x) - E(\tau_{x'}) \tag{8.12}$$

is the causal average total effect (see Def. 5.8).

Corollary 8.17 (Unbiasedness of Prima Facie Effects)
Let the Assumptions 5.1 hold. Then $C_X \perp\!\!\!\perp X$ implies

$$PFE_{xx'} = ATE_{xx'}. \tag{8.13}$$

Remark 8.18 (Identification of $ATE_{xx'}$) Hence, if $C_X \perp\!\!\!\perp X$, then Equations (8.11) and (8.13) imply that the causal average total effect of x compared to x' is identical to the difference between the two conditional expectation values $E(Y|X{=}x)$ and $E(Y|X{=}x')$, that is, if $C_X \perp\!\!\!\perp X$, then

$$ATE_{xx'} = E(Y|X{=}x) - E(Y|X{=}x'). \tag{8.14}$$

In this context we also say that the $ATE_{xx'}$ *is identified* by the difference between $E(Y|X{=}x)$ and $E(Y|X{=}x')$. Note that, in a t-test for independent observations, we test the hypothesis $E(Y|X{=}x) - E(Y|X{=}x') = 0$. Therefore, if $C_X \perp\!\!\!\perp X$, then we also test the hypothesis $ATE_{xx'} = 0$. ◁

8.2.3 Implications of Z-Conditional Independence of C_X and X

In the last subsection we studied the implications of $C_X \perp\!\!\!\perp X$. Now we turn to the implications of $C_X \perp\!\!\!\perp X|Z$, assuming that Z is covariate of X, which implies that Z is C_X-measurable. Remember, if Z is C_X-measurable, then $C_X \perp\!\!\!\perp X$ implies $C_X \perp\!\!\!\perp X|Z$, but not vice versa [see Prop. (7.7)]. Hence, if $C_X \perp\!\!\!\perp X|Z$ holds, then $C_X \perp\!\!\!\perp X$ does not necessarily

hold. In fact, we may conduct an experiment in such a way that $C_X \perp\!\!\!\perp X \mid Z$ holds but not $C_X \perp\!\!\!\perp X$ (see Table 6.3 for a simple example).

Theorem 8.19 summarizes the implications of $C_X \perp\!\!\!\perp 1_{X=x} \mid Z$, that is, of

$$P(X{=}x \mid C_X) \underset{P}{=} P(X{=}x \mid Z), \tag{8.15}$$

where we presume that Z is C_X-measurable, so that $P(X{=}x \mid C_X) \underset{P}{=} P(X{=}x \mid C_X, Z)$ (see SN-Rem. 10.4).

Theorem 8.19 (Implications of $C_X \perp\!\!\!\perp 1_{X=x} \mid Z$)
Let the Assumptions 5.1 hold and let Z be a covariate of X. Then $C_X \perp\!\!\!\perp 1_{X=x} \mid Z$ implies:

(i) $\tau_x \perp\!\!\!\perp 1_{X=x} \mid Z$.
(ii) $\tau_x \vdash 1_{X=x} \mid Z$.

(Proof p. 196)

Hence, under the assumptions of Theorem 8.19,

$$C_X \perp\!\!\!\perp 1_{X=x} \mid Z \quad \Rightarrow \quad \tau_x \perp\!\!\!\perp 1_{X=x} \mid Z \quad \Rightarrow \quad \tau_x \vdash 1_{X=x} \mid Z. \tag{8.16}$$

The following corollary is an immediate implication of Theorem 8.19 (i) and Theorem 7.25 (i).

Corollary 8.20 (Unbiasedness of $E^{X=x}(Y \mid Z)$)
Let the Assumptions 5.1 hold, let Z be a covariate of X, and assume $C_X \perp\!\!\!\perp 1_{X=x} \mid Z$. Then:

(i) If τ_x is P-unique, then $E^{X=x}(Y \mid Z)$ is unbiased.
(ii) If $z \in \Omega_Z'$, $P(X{=}x, Z{=}z) > 0$, and τ_x is $P^{Z=z}$-unique, then $E(Y \mid X{=}x, Z{=}z)$ is unbiased.

In Theorem 8.19 we considered the implications of Z-conditional independence of C_X and the indicator $1_{X=x}$ for a single value x of X. In the following theorem, we turn to the implications of Z-conditional independence of C_X and X, which is equivalent to

$$\forall\, x \in X(\Omega) : \; C_X \perp\!\!\!\perp 1_{X=x} \mid Z.$$

This in turn implies that also all propositions listed in Theorem 8.19 hold for all values x of X. In the following theorem we list some additional propositions that are not already included in Theorem 8.19.

Theorem 8.21 (Implications of $C_X \perp\!\!\!\perp X \mid Z$)
Let the Assumptions 6.1 hold, let Z be a covariate of X, and define $\tau := (\tau_0, \tau_1, \ldots, \tau_J)$. Then $C_X \perp\!\!\!\perp X \mid Z$ implies:

(i) For all $x \in X(\Omega)$: $C_X \perp\!\!\!\perp 1_{X=x} \mid Z$.
(ii) For all $x \in X(\Omega)$: $\tau_x \perp\!\!\!\perp X \mid Z$.
(iii) $\tau \perp\!\!\!\perp X \mid Z$.
(iv) $\tau \vdash X \mid Z$.

(Proof p. 196)

The following corollary immediately follows from Theorem 8.21 (i), Corollary 8.20, and Definition 6.11 (iii).

Corollary 8.22 (Unbiasedness of $E(Y|X,Z)$)
Let the Assumptions 6.1 hold, let Z be a covariate of X, and assume $C_X \perp\!\!\!\perp X \mid Z$. Then:

(i) If τ_x is P-unique, then $E^{X=x}(Y|Z)$ is unbiased.
(ii) If, for all $x \in X(\Omega)$, τ_x is P-unique, then $E(Y|X,Z)$ is unbiased.

8.2.4 Unbiasedness of a Z-Conditional Prima Facie Effect Function

The following corollary is an immediate implication of Theorem 8.21. In this corollary we specify a condition under which a Z-conditional prima facie effect variable

$$PFE_{Z;xx'}(Z) \underset{P}{=} E^{X=x}(Y|Z) - E^{X=x'}(Y|Z) \tag{8.17}$$

is unbiased. Remember that $CTE_{Z;xx'}(Z)$ denotes $E(\delta_{xx'}|Z) \underset{P}{=} E(\tau_x|Z) - E(\tau_{x'}|Z)$, a causal Z-conditional total effect variable [see Def. 5.17 (iii)].

Corollary 8.23 (Unbiasedness of a Z-Conditional Prima Facie Effect Function)
Let the Assumptions 5.1 hold and let Z be a covariate of X.

(i) If τ_x and $\tau_{x'}$ are P-unique, then $C_X \perp\!\!\!\perp X \mid Z$ implies

$$PFE_{Z;xx'}(Z) \underset{P}{=} CTE_{Z;xx'}(Z). \tag{8.18}$$

(ii) If $z \in \Omega_Z'$ such that $P(X{=}x, Z{=}z), P(X{=}x', Z{=}z) > 0$ and $\tau_x, \tau_{x'}$ are $P^{Z=z}$-unique, then $C_X \perp\!\!\!\perp X \mid Z$ implies

$$PFE_{Z;xx'}(z) = CTE_{Z;xx'}(z). \tag{8.19}$$

(Proof p. 197)

Remark 8.24 (Identification of $CTE_{Z;xx'}(Z)$) Hence, if $C_X \perp\!\!\!\perp X|Z$ and the true outcome variables τ_x and $\tau_{x'}$ are P-unique, then Equations (8.17) and (8.18) imply that the causal Z-conditional total effect variable is almost surely identical to the difference between the two conditional expectations $E^{X=x}(Y|Z)$ and $E^{X=x'}(Y|Z)$, that is,

$$CTE_{Z;xx'}(Z) \underset{P}{=} E^{X=x}(Y|Z) - E^{X=x'}(Y|Z). \tag{8.20}$$

In this context we also say that the effect variable $CTE_{Z;xx'}(Z)$ is *identified* by the difference between $E^{X=x}(Y|Z)$ and $E^{X=x'}(Y|Z)$. ◁

Remark 8.25 (Identification of $CTE_{Z;xx'}(z)$) Under the assumptions mentioned in Corollary 8.23 (ii), the causal $(Z{=}z)$-conditional total effect is identical to the difference between the two conditional expectation values $E(Y|X{=}x, Z{=}z)$ and $E(Y|X{=}x', Z{=}z)$, that is,

$$CTE_{Z;xx'}(z) = E(Y|X{=}x, Z{=}z) - E(Y|X{=}x', Z{=}z). \tag{8.21}$$

In this context we also say that the conditional effect $CTE_{Z;xx'}(z)$ is *identified* by the difference between $E(Y|X{=}x, Z{=}z)$ and $E(Y|X{=}x', Z{=}z)$. ◁

Remark 8.26 (Conditional Randomization) The condition $C_X \perp\!\!\!\perp X \,|\, Z$ that, together with P-uniqueness of the true outcome variables τ_x, implies unbiasedness of $E(Y|X,Z)$, can be created in an experiment either by

(a) randomized assignment of the sampled unit to a treatment condition (represented by a value x of X) [see SN-Box 16.3 (ix)]
(b) randomized assignment of the unit to a treatment condition *conditional on the values z of a covariate Z*.

However, $C_X \perp\!\!\!\perp X \,|\, Z$ may also hold if the (possibly multivariate) covariate Z is selected in such a way that $C_X \perp\!\!\!\perp X \,|\, Z$ holds. ◁

Remark 8.27 (Continuous X) It should be emphasized that $C_X \perp\!\!\!\perp X$ and $C_X \perp\!\!\!\perp X \,|\, Z$ may also hold if X is a continuous random variable. If $C_X \perp\!\!\!\perp X$ holds, then $E(Y|X)$ and $E(Y|X,Z)$ describe the *causal* total dependency of Y on X and of Y on X given Z, respectively. If X is continuous, then the theory based on true outcome variables does not apply any more, because it rests on the assumption that the values x of X have a nonzero probability. This assumption does not hold if X is a continuous random variable. Nevertheless, we can meaningfully talk about causal effects. For example, if

$$E(Y|X) \underset{P}{=} \alpha_0 + \alpha_1 \cdot X, \quad \alpha_0, \alpha_1 \in \mathbb{R}, \tag{8.22}$$

and $C_X \perp\!\!\!\perp X$ hold, then we may define α_1 to be the *causal total effect of X on Y*.

Similarly, if $C_X \perp\!\!\!\perp X \,|\, Z$ holds for a covariate Z with value space $(\Omega'_Z, \mathscr{A}'_Z)$ and there are measurable functions $g_0, g_1 : \Omega'_Z \to \mathscr{B}$ with $g_0^{-1}(\mathscr{B}) \subset \mathscr{A}'_Z$ and $g_1^{-1}(\mathscr{B}) \subset \mathscr{A}'_Z$ such that

$$E(Y|X,Z) \underset{P}{=} g_0(Z) + g_1(Z) \cdot X, \tag{8.23}$$

then we may define g_1 to be the *causal Z-conditional total effect function of X on Y* and the composition $g_1(Z)$ to be the corresponding *causal Z-conditional total effect variable*. Hence, the causality conditions $C_X \perp\!\!\!\perp X$ and $C_X \perp\!\!\!\perp X \,|\, Z$ are relevant beyond the true outcome theory of causal effects. ◁

8.3 Examples

Now we illustrate the causality conditions treated in this chapter by some examples. In chapter 3 we introduced a first example with independence of X and a global potential confounder C_X (see Table 3.1, p. 44). A second example has already been presented in chapter 6 (see Table 6.2, p. 139), and in the same chapter , there is also an example with Z-conditional independence of C_X and X (see Table 6.3, p. 140). In all these examples, the observational-unit variable U is a global potential confounder, that is, $C_X = U$, implying

$$E(Y|X,C_X) \underset{P}{=} E(Y|X,U) \tag{8.24}$$

and

$$P(X{=}1\,|\,C_X) \underset{P}{=} P(X{=}1\,|\,U). \tag{8.25}$$

In this equation, $P(X{=}1\,|\,U)$ denotes the individual treatment probability function, whose values are the individual treatment probabilities $P(X{=}1\,|\,U{=}u)$ (see Remarks 3.58 and

3.67). If the values u of U represent the observational units *at the onset of treatment*, then $C_X = U$ will hold in empirical applications if (a) no fallible covariate is assessed and (b) there is no other variable that is simultaneous to the treatment variable X (such as a second treatment variable).

Example 8.28 (Independence of X and U) Table 6.2 (p. 139) displays an example in which $C_X \perp\!\!\!\perp X$ holds, where $C_X = U$. In this table it is easily seen that the individual treatment probabilities are the same for all units, that is,

$$P(X{=}1\,|\,U) \;=\; P(X{=}1) \;=\; \frac{3}{4}.$$

$U \perp\!\!\!\perp X$ implies that X and *all* random variables that are measurable with respect to U are independent.

In the same example, $C_X \perp\!\!\!\perp X\,|\,Z$ holds as well, where Z = *sex*, because the $(Z{=}z)$-conditional treatment probabilities also do not depend on the units, that is,

$$P(X{=}1\,|\,U,Z) \;=\; P(X{=}1\,|\,Z) \;=\; P(X{=}1) \;=\; \frac{3}{4}.$$

$U \perp\!\!\!\perp X\,|\,Z$ implies that X and all random variables that are measurable with respect to U are also Z-conditionally independent. This is no coincidence but an implication of the fact that $U \perp\!\!\!\perp X$ and that measurability of Z with respect to U implies $U \perp\!\!\!\perp X\,|\,Z$ [see Prop. (7.7) or SN-Box 16.3 (ix)], which in turn implies Z-conditional independence of X and all random variables that are measurable with respect to U [see SN-Box 16.3 (vi)].

If $C_X = U$, then, according to Theorem 8.14 (v), independence of U and X implies unbiasedness of the conditional expectation $E(Y\,|\,X)$. Furthermore, because in this example $P(X=0) > 0$ and $P(X=1) > 0$, independence of X and U also implies unbiasedness of the conditional expectation values $E(Y\,|\,X=0)$ and $E(Y\,|\,X=1)$ [see Th. 8.11 (iv)] as well as unbiasedness of the prima facie effect

$$PFE_{10} \;:=\; E(Y\,|\,X{=}1) - E(Y\,|\,X{=}0) \;\approx\; 102.333 - 92.333 \;\approx\; 10$$

(see Cor. 8.17). In other words, $PFE_{10} = ATE_{10}$.

As already stated above, $U \perp\!\!\!\perp X\,|\,Z$ holds as well, where Z = *sex*. Because $C_X = U$ and Z is U-measurable, according to Theorem 8.14 (x), the conditional expectation $E(Y\,|\,X,Z)$ is unbiased. Furthermore, because $P(X=0,Z{=}z),\, P(X=1,Z{=}z) > 0$ for both values z of Z, this implies unbiasedness of all conditional expectation values $E(Y\,|\,X{=}x,Z{=}z)$ [see Th. 8.11 (ix)], and that the $(Z{=}z)$-conditional prima facie effects

$$PFE_{Z;10}(m) \;=\; E(Y\,|\,X{=}1,Z{=}m) - E(Y\,|\,X{=}0,Z{=}m) \;\approx\; 92.50 - 83.00 \;\approx\; 9.50$$

and

$$PFE_{Z;10}(f) \;=\; E(Y\,|\,X{=}1,Z{=}f) - E(Y\,|\,X{=}0,Z{=}f) \;=\; 122 - 111 \;=\; 11$$

are unbiased [see Cor. 8.23 (ii)].

Also note that

$$PFE_{10} \;=\; E[PFE_{Z;10}(Z)] \;=\; 9.50 \cdot \frac{4}{6} + 11 \cdot \frac{2}{6} \;=\; 10.$$

This means that the prima facie effect is the expectation of the corresponding conditional prima facie effects. This property, is no coincidence. Instead it follows from $C_X \perp\!\!\!\perp X$ (see SN-Th. 15.14 and SN-Rem. 15.17). ◁

Example 8.29 (Z-Conditional Independence of X and U) In Table 6.3 (see p. 140) we already treated an example in which $C_X \perp\!\!\!\perp X \mid Z$ holds, whereas $C_X \perp\!\!\!\perp X$ does not. Again, $C_X = U$. As is easily seen in this table, the individual treatment probabilities are the same for all units *within each of the two subsets of males and females*, that is,

$$P(X{=}1 \mid U{=}u, Z{=}m) \;=\; P(X{=}1 \mid Z{=}m) \;=\; \frac{3}{4}, \quad \text{for each male unit } u$$

and

$$P(X{=}1 \mid U{=}u, Z{=}f) \;=\; P(X{=}1 \mid Z{=}f) \;=\; \frac{1}{4}, \quad \text{for each female unit } u.$$

Hence, the individual treatment probabilities are 3/4 for males and 1/4 for females. These individual treatment probabilities differ for different values of the covariate Z, but they are invariant *given a value of the covariate Z*. Note that the conditional treatment probability $P(X{=}1|Z{=}m) = 3/4$ is also the individual treatment probability $P(X{=}1|U{=}u)$ for the male units u_1 to u_4, and the conditional treatment probability $P(X{=}1 \mid Z{=}f) = 1/4$ is also the individual treatment probability $P(X{=}1|U{=}u)$ for the female units u_5 and u_6. This follows from

$$\begin{aligned} P(X{=}1 \mid U) \;&=\; E[P(X{=}1 \mid U, Z) \mid U] && \text{[SN-Box 10.2 (v)]} \\ &=\; E[P(X{=}1 \mid Z) \mid U] && [P(X{=}1 \mid U, Z) = P(X{=}1 \mid Z)] \\ &=\; P(X{=}1 \mid Z). && \text{[SN-Box 10.2 (vii)]} \end{aligned}$$

In Table 6.3 (p. 140) there is only *Z-conditional* independence of U and the treatment variable X, that is, $U \perp\!\!\!\perp X \mid Z$. Therefore, in this table, the prima facie effect

$$PFE_{10} \;=\; E(Y \mid X{=}1) - E(Y \mid X{=}0) \;\approx\; 96.715 - 99.800 \;\approx\; -3.085$$

is biased, because the average total effect in this example is $ATE_{10} = 10$ (see Exercise 8-2). However, the *$(Z{=}z)$-conditional* prima facie effects are unbiased. In fact, they are the same as in Table 6.2 (see p. 139 and Example 8.28). Hence, we can use the conditional prima facie effects to compute the average total effect. Remember, if the conditional prima facie effects are unbiased, that is, if they are equal to the causal conditional total effects, then the expectation of the conditional prima facie effect variable is equal to the causal average total effect [see Eq. (6.38)]. In our example, this expectation is,

$$\begin{aligned} ATE_{10} \;=\; E[PFE_{Z;10}(Z)] \;&=\; PFE_{Z;10}(m) \cdot \frac{4}{6} + PFE_{Z;10}(f) \cdot \frac{2}{6} \\ &=\; 9.50 \cdot \frac{4}{6} + 11 \cdot \frac{2}{6} \;=\; 10.00. \end{aligned}$$

◁

8.4 Methodological Conclusions

Now we discuss the conclusions from the theory treated in this chapter for the design and analysis of experiments and quasi-experiments. Theorems 8.11 and 8.14 are the theoretical foundation of the design technique of *randomization* and of the analysis of causal

conditional and causal average total treatment effects in experiments by comparing (unadjusted) means between treatment conditions. Theorems 8.19 and 8.21 may also be used for conditionally randomized experiments for the analysis of causal $(Z=z)$-conditional total treatment effects. Together with Equation (6.38) these theorems can also be used for the analysis of causal average total effects. Finally, Theorems 8.19 and 8.21 can also be used in (nonrandomized) quasi-experiments for covariate selection and the analysis of causal conditional and causal average total effects. This will now be explained in more detail.

Remark 8.30 (Randomization) It is well-known at least since (Fisher, 1925/1946) that *randomization* plays a crucial role in ruling out biases in comparisons of means. In a randomized experiment, there is always an observational-unit variable U whose value u denotes the observational unit that is sampled if the random experiment considered is actually conducted. The definition of a global potential confounder C_X [see Def. 4.4 (ii)] guarantees that U is measurable with respect to a global potential confounder C_X. In fact, in a simple experiment, in which we do not observe any fallible pretests and in which there is only one single treatment variable X, the random variable U itself is a global potential confounder of X. In any case, if X represents a discrete treatment variable, then $C_X \perp\!\!\!\perp X$ implies

$$P(X=x \mid C_X) \underset{P}{=} P(X=x \mid U) \underset{P}{=} P(X=x), \quad \forall\, x \in X(\Omega)\,. \tag{8.26}$$

The last one of these two equations implies that each unit u has the same probability $P(X=x \mid U=u) = P(X=x)$ to be assigned to treatment x.

Remember that we are talking about a *single-unit trial*. A simple example of such a single-unit trial consists of sampling a single unit from a set of units, assessing a number of covariates, assigning the unit (or observing its assignment) to one of the treatment conditions and observing the outcome variable (see ch. 2 for more details and other kinds of single-unit trials).

Note that Equation (8.26) does *not* imply that the probabilities $P(X=x)$ are the same for all treatment conditions x. If, for example, there are two treatment conditions, say 0 and 1, then the two treatment probabilities might be $P(X=1) = 1/4$ and $P(X=0) = 3/4$. However, $C_X \perp\!\!\!\perp X$ implies $P(X=1 \mid U) = P(X=1)$, provided, of course, that we consider a random experiment in which U is C_X-measurable. In other words, $C_X \perp\!\!\!\perp X$ implies that the individual treatment probabilities $P(X=x \mid U=u)$ are identical between units, and they are equal to the unconditional treatment probability $P(X=x)$. Hence, in a perfect randomized experiment we ensure $C_X \perp\!\!\!\perp X$. This condition implies that the conditional expectation values $E(Y \mid X=x)$, $x \in X(\Omega)$, are unbiased (see Ths. 8.11 and 8.14) and that a prima facie effect $E(Y \mid X=x) - E(Y \mid X=x')$ is identical to the causal average total effect $ATE_{xx'}$ (see Cor. 8.17).

It is important to note that in a perfect randomized experiment we do not only create $C_X \perp\!\!\!\perp X$ but also $C_X \perp\!\!\!\perp X \mid Z$ for each covariate Z of X (see Th. 8.14). This in turn implies that each conditional expectation $E^{X=x}(Y \mid Z)$, $x \in X(\Omega)$, the conditional expectation $E(Y \mid X, Z)$, and each prima facie effect variable $PFE_{Z;xx'} = E^{X=x}(Y \mid Z) - E^{X=x'}(Y \mid Z)$, $x, x' \in X(\Omega)$, are unbiased (see Cor. 8.23).

◁

Remark 8.31 (Conditional Randomization) In the single-unit trial of an experiment or quasi-experiment, in which a unit u is sampled and a value z of the covariate Z is assessed

prior to treatment, *conditional randomization* refers to randomized assignment of the unit u to treatment condition x with probability $P(X{=}x \mid Z{=}z)$, where $x = 0, 1, \ldots, J$. In this procedure we have to make sure that treatment assignment only depends on the values z of Z, but not on any other attribute of the units or any other covariate. In more formal terms, we create $P(X{=}x \mid C_X) \underset{P}{=} P(X{=}x \mid Z)$ for all values x of X.

We distinguish two cases. *First*, if Z is measured without measurement error, each value z of Z will represent a subset of observational units. Table 6.3 (p. 140) contains an example with conditional independence of units and treatments given males and given females. *Second*, if the covariate Z is measured *with* error, that is, $Z{=}f(U)+\varepsilon$, then a value z of Z does *not* represent a subset of units, because z is not only determined by u but also by other factors and/or chance. Nevertheless, treatment can be assigned such that the treatment probability depends on the *observed score* z, and not on the attribute $f(U)$ of the unit itself. In both cases, namely $Z{=}f(U)$ and $Z{=}f(U)+\varepsilon$, conditional randomized assignment of a unit u to one of the treatment conditions given a value z of Z secures that a global potential confounder C_X satisfies $C_X \perp\!\!\!\perp X \mid Z$, that is, Z-conditional independence of C_X and X.

Conditional random assignment allows that units with different values z_1 and z_2 of Z have different probabilities to be assigned to treatment x. Thus it is possible to assign a unit for which we observe a covariate value z_1 (e. g., high need) to treatment x with a higher probability than a unit with covariate value z_2 (e. g., low need). In this way we may respect ethical and/or other requirements without compromising the validity of causal inference. Remember, unconditional randomized assignment means that each unit has the *same* probability of being assigned to a given treatment, regardless of his or her need (and any other attribute of the unit).

For simplicity, suppose there are just two treatment conditions, $X{=}0$ (control) and $X{=}1$ (treatment). Conditional randomization consists of:

(a) fixing the conditional treatment probabilities $P(X{=}1 \mid Z{=}z)$ for all values z of the covariate Z,
(b) sampling a unit u and assessing the value z of the covariate, and
(c) randomized assignment of the unit with probability $P(X{=}1 \mid Z{=}z)$ to treatment 1.

If, for example, the covariate has three values, say *high*, *medium*, and *low need*, we may toss a dice and assign a unit with *high need* to treatment if the dice shows less than six dots, and assign it to control otherwise. A unit with *medium need* might be assigned to treatment if the dice shows less than four dots, and to control otherwise. Finally, a unit with *low need* might be assigned to treatment if the dice shows one dot, and to control otherwise.

If the covariate Z is discrete, then the conditional treatment probabilities $P(X{=}1 \mid Z{=}z)$ may be fixed in a table by assigning a treatment probability to each value z that seems appropriate with respect to ethical and other requirements. In the example above, these were the values 5/6 for *high need*, 3/6 for *medium need* and 1/6 for *low need*. If the covariate is univariate continuous, we may also use a function, such as

$$P(X{=}1 \mid Z) \;=\; \frac{\exp(\lambda_0 + \lambda_1 \cdot Z)}{1 + \exp(\lambda_0 + \lambda_1 \cdot Z)}$$

with real-valued coefficients λ_0 and λ_1 that seem appropriate for the experiment considered.

If Z is discrete and $P(X{=}x, Z{=}z), P(X{=}x', Z{=}z) > 0$, then Z-conditional randomization implies that the $(Z{=}z)$-conditional prima facie effects $PFE_{Z;xx'}(z) = E(Y|X{=}x, Z{=}z) - E(Y|X{=}x', Z{=}z)$ are identical to the causal conditional total effects given a value z of the covariate Z. Hence, studying conditional prima facie effects in a perfect conditionally randomized experiment is tantamount to studying causal conditional total effects. As has been shown in chapter 5 [see, e. g., Eq. (5.22)], once we know the causal conditional total effects given the values z of a covariate Z, we can also compute the causal average total effect. ◁

Remark 8.32 (Covariate Selection in Quasi-Experiments) Covariate selection is also the first step in a number of techniques for the analysis of causal conditional and causal average total effects in *quasi-experiments*, in which by definition of the quasi-experiment, the experimenter cannot fix the true treatment probabilities himself. The steps to follow distinguish different techniques of analysis. By definition, there is no randomization and no conditional randomization in a quasi-experiment. However, even under initial randomization, systematic attrition of subjects may invalidate the condition $X \perp\!\!\!\perp C_X$ (see, e. g., Abraham & Russell, 2004; Fichman & Cummings, 2003; Graham & Donaldson, 1993; Shadish et al., 2002). In this case, we will say that randomization failed and the initially randomized experiment turned into a quasi-experiment.

In quasi-experiments, selecting the covariates in the covariate vector $Z := (Z_1, \ldots, Z_m)$ for which we can hope that $C_X \perp\!\!\!\perp X \,|\, Z$ holds is a useful strategy in the analysis of causal conditional and causal average total treatment effects. However, note that there might be many covariates determining the treatment probabilities. For instance, the *severity of the disorder*, *knowing about the treatment*, and *availability of the treatment* are candidates for such covariates. Also note that there is no guarantee that $C_X \perp\!\!\!\perp X \,|\, Z$ holds for a specified (univariate or multivariate) covariate Z. ◁

Remark 8.33 (Falsifiability) As already mentioned before, in contrast to the causality conditions treated in chapters 6 and 7, the causality conditions $C_X \perp\!\!\!\perp X$ and $C_X \perp\!\!\!\perp X \,|\, Z$ can be tested in empirical applications, at least in the sense that some consequences of these conditions can be checked. Falsifiability is important, because otherwise we would not have any criterion for covariate selection, that is, for deciding whether or not a specific covariate should be included in the vector $Z := (Z_1, \ldots, Z_m)$ of covariates with respect to which we hope that $C_X \perp\!\!\!\perp X \,|\, Z$ holds.

Let us briefly outline how we can falsify the assumption that $C_X \perp\!\!\!\perp X$ holds. Remember, if X is discrete, then $C_X \perp\!\!\!\perp X$, is equivalent to

$$P(X{=}x \,|\, C_X) \underset{P}{=} P(X{=}x), \quad \forall\, x \in X(\Omega)\,. \tag{8.27}$$

This implies that

$$P(X{=}x \,|\, W) \underset{P}{=} P(X{=}x), \quad \forall\, x \in X(\Omega) \tag{8.28}$$

holds for each random variable W that is measurable with respect to C_X. Examples for such random variables W are sex, educational status, but also a fallible pre-test.

Similarly, if X is discrete, then $C_X \perp\!\!\!\perp X \,|\, Z$ is equivalent to

$$P(X{=}x \,|\, C_X) \underset{P}{=} P(X{=}x \,|\, Z), \quad \forall\, x \in X(\Omega)\,. \tag{8.29}$$

(Remember that Z denotes a covariate of X, which, by definition, is C_X-measurable.) This equation implies

$$P(X{=}x \mid W, Z) \underset{P}{=} P(X{=}x \mid Z), \quad \forall\, x \in X(\Omega), \tag{8.30}$$

for each random variable W that is measurable with respect to C_X. Equations (8.28) and (8.30) are easily tested using standard procedures for the analysis of conditional probabilities such as logistic regression (see, e. g., Agresti, 2007; Bonney, 1987; Green, 2003; Hosmer & Lemeshow, 2000) or probit regression (see, e. g., McCullagh & Nelder, 1989; Borooah, 2001; Liao, 1994). ◁

Remark 8.34 (The Covariate Selection Process) If Equation (8.30) does not hold for specified C_X-measurable random variables W and Z, then $C_X \perp\!\!\!\perp X \mid Z$ cannot hold as well. In this case we may define $Z^* := (Z, W)$ and select a new C_X-measurable random variable W^*, and check if

$$P(X{=}x \mid W^*, Z^*) \underset{P}{=} P(X{=}x \mid Z^*), \quad \forall\, x \in X(\Omega), \tag{8.31}$$

holds. This process can be continued as long as there is doubt that

$$P(X{=}x \mid C_X) \underset{P}{=} P(X{=}x \mid Z^*), \quad \forall\, x \in X(\Omega), \tag{8.32}$$

holds. Of course, such a procedure does not guarantee that we find a (possibly multivariate) covariate Z^* such that Equation (8.32) holds. Instead, in a quasi-experiment, Equation (8.32) always remains an assumption on which causal inference relies. However, this assumption can always be falsified, which has a positive and a negative side. The negative side is that we can never be sure that this assumption holds. The positive side is that this assumption is empirically testable and, in this sense, it is not just a matter of belief. ◁

8.5 Summary and Conclusions

In chapter 6 we showed that unbiasedness of the conditional or unconditional prima facie effects is crucial for computing causal conditional and unconditional total effects from (the empircally estimable) conditional or unconditional prima facie effects. In chapter 7 we treated several causality conditions that imply unbiasedness, which cannot be tested in empirical applications. The most restrictive of these conditions is Rosenbaum and Rubin's strong ignorability. In this chapter, we introduced a first kind of causality conditions that are empirically falsifiable. These causality conditions are summarized in Box 8.1.

The implication relations among the causality conditions treated in this chapter are listed in Table 8.1. This table also includes $\tau \perp\!\!\!\perp X$ and the strong ignorability condition $\tau \perp\!\!\!\perp X \mid Z$ that were treated in chapter 7. The implications of these two conditions on other causality conditions are summarized in the last rows of Table 7.5 and Table 7.6, respectively. Hence, putting these tables together yields long chains of implications. Obviously, $C_X \perp\!\!\!\perp X$ is the most powerful causality condition. It implies not only unbiasedness of all $E(Y \mid X{=}x)$, $E^{X=x}(Y \mid Z)$, $x \in X(\Omega)$, as well as unbiasedness of $E(Y \mid X)$ and $E(Y \mid X, Z)$, but it also implies all other conditions treated until now (including those dealt with in ch. 7), and it even implies that all true outcomes variables are P-unique. In contrast, assuming $C_X \perp\!\!\!\perp X \mid Z$, we need the additional assumption that all true outcomes variables are P-unique in order to derive unbiasedness of all $E^{X=x}(Y \mid Z)$, $x \in X(\Omega)$, as well as unbiasedness of $E(Y \mid X, Z)$.

Box 8.1 Fisher conditions

Let the Assumptions 8.1 hold.

$C_X \perp\!\!\!\perp X$ *Independence of C_X and X.* If X is discrete, then $C_X \perp\!\!\!\perp X$ is equivalent to

$$\forall\, x \in X(\Omega)\colon P(X{=}x \,|\, C_X) \underset{P}{=} P(X{=}x).$$

It can be created by randomized assignment of the unit sampled to one of the treatments x. If $P(X{=}x) > 0$ for all values x, then $C_X \perp\!\!\!\perp X$ implies P-uniqueness of all true outcome variables τ_x, which is equivalent to $P(X{=}x \,|\, C_X) \underset{P}{>} 0$ for all values x. Furthermore, $C_X \perp\!\!\!\perp X$ does not only imply unbiasedness of $E(Y|X)$ and all $E(Y|X{=}x)$, $x \in X(\Omega)$, but it also implies $C_X \perp\!\!\!\perp X \,|\, Z$, provided that Z is a covariate of X.

Let the Assumptions 8.4 hold and let Z be a covariate of X.

$C_X \perp\!\!\!\perp X|Z$ *Z-conditional independence of C_X and X.* If X is discrete, then $C_X \perp\!\!\!\perp X|Z$ is equivalent to

$$\forall\, x \in X(\Omega)\colon P(X{=}x \,|\, C_X) \underset{P}{=} P(X{=}x \,|\, Z).$$

$C_X \perp\!\!\!\perp X|Z$ follows from $C_X \perp\!\!\!\perp X$. However, it can also be created by conditional randomized assignment of the unit sampled to treatment condition x based on the values z of Z. If $P(X{=}x \,|\, C_X) \underset{P}{>} 0$ holds for all $x \in X(\Omega)$, then $C_X \perp\!\!\!\perp X|Z$ implies that $E(Y|X,Z)$ and all $E^{X=x}(Y|Z)$, $x \in X(\Omega)$, are unbiased. Furthermore, we can try to select the (possibly multivariate) covariate $Z = (Z_1, \ldots, Z_m)$ such that $C_X \perp\!\!\!\perp X|Z$ is satisfied.

Randomization and Conditional Randomization in Experiments

In experiments, the condition $C_X \perp\!\!\!\perp X$ can be created by randomized assignment of the observational unit to one of the treatment conditions represented by a value x of X. Similarly, $C_X \perp\!\!\!\perp X|Z$ can be created by conditional randomized assignment based on a value z of covariate Z. Outside the randomized experiment, and even in cases in which X does not denote a treatment variable that is manipulated by an experimenter, $C_X \perp\!\!\!\perp X|Z$ may also be valid if the covariate Z is carefully selected. In this case, however, there is no *guarantee* that this condition will hold.

Falsifiability

Unlike unbiasedness and the other causality conditions treated in chapter 7, independence of X and potential confounders as well as Z-conditional independence of X and potential confounders, are empirically falsifiable. Suppose that X is discrete. In order to test $C_X \perp\!\!\!\perp X$, we simply have to select a covariate W and investigate if the conditional probability $P(X{=}x|W)$ is in fact identical to $P(X{=}x)$ for all values x of X. [Remember, a covariate of X has been defined such that it is measurable with respect to C_X (see Rem. 4.9).] Similarly, if X is discrete and Z is a covariate of X, then we can test if $C_X \perp\!\!\!\perp X \mid Z$ holds by

Table 8.1. Implications among causality conditions

	$\tau \perp\!\!\!\perp X$	$C_X \perp\!\!\!\perp X$	$Z \mid X \perp\!\!\!\perp \tau$	$C_X \mid X \perp\!\!\!\perp Z$
$\tau \perp\!\!\!\perp X$	$\Leftrightarrow$			
$C_X \perp\!\!\!\perp X$	$\Rightarrow$	$\Leftrightarrow$	$\Rightarrow$	$\Rightarrow$
$\tau \perp\!\!\!\perp X \mid Z$			$\Leftrightarrow$	
$C_X \perp\!\!\!\perp X \mid Z$			$\Rightarrow$	$\Leftrightarrow$

Note: An entry $\Rightarrow$ ($\Leftrightarrow$) means that the condition in the row implies (is equivalent to) the condition in the column, provided that the Assumptions 6.1 hold and Z is a covariate of X.

selecting another covariate W and see if $P(X{=}x \mid Z, W) \underset{P}{=} P(X{=}x \mid Z)$ holds for all values x of X.

Continuous X

It should be emphasized again that $C_X \perp\!\!\!\perp X$ and $C_X \perp\!\!\!\perp X \mid Z$ may also hold if X is a continuous random variable. If X is continuous, then the theory of true outcome variables does not apply any more, because it rests on the assumption that the values x of X have a nonzero probability. This assumption does not hold if X is continuous. Nevertheless, we can meaningfully talk about causal dependencies. Hence, the causality conditions $C_X \perp\!\!\!\perp X$ and $C_X \perp\!\!\!\perp X \mid Z$ are relevant beyond the true outcome theory of causal effects (see Rem. 8.27 for more details).

8.6 Proofs

Proof of Theorem 8.11

(i). According to SN-Corollary 14.48 (a) and (c), P-uniqueness of τ_x is equivalent to $P(X{=}x \mid C_X) \underset{P}{>} 0$. Therefore, Equation (8.10), which is equivalent to $C_X \perp\!\!\!\perp 1_{X=x}$, in conjunction with our assumption $P(X{=}x) > 0$ implies that τ_x is P-unique.

(ii).

$$
\begin{aligned}
P(X{=}x \mid \tau_x) &\underset{P}{=} E(1_{X=x} \mid \tau_x) && \text{[SN-(10.4)]} \\
&\underset{P}{=} E\big(E(1_{X=x} \mid C_X) \,\big|\, \tau_x\big) && \text{[}\sigma(\tau_x) \subset \sigma(C_X)\text{, SN-Box 10.2 (v)]} \\
&\underset{P}{=} E\big(P(X{=}x \mid C_X) \,\big|\, \tau_x\big) && \text{[SN-(10.4)]} \\
&\underset{P}{=} E\big(P(X{=}x) \,\big|\, \tau_x\big) && \text{[(8.10)]} \\
&\underset{P}{=} P(X{=}x). && \text{[SN-Box 10.2 (i)]}
\end{aligned}
$$

According to Lemma 7.4, this equation is equivalent to independence of $1_{X=x}$ and τ_x.

(iii). $\tau_x \vdash 1_{X=x}$ is equivalent to $E(\tau_x|1_{X=x}) \underset{P}{=} E(\tau_x)$. According to SN-Box 10.2 (vi), this equation follows from $\tau_x \perp\!\!\!\perp 1_{X=x}$.

(iv). Unbiasedness of $E(Y|X{=}x)$ immediately follows from (iii) and Theorem 6.9.

(v). This proposition immediately follows from C_X-measurability of Z and Proposition (7.7).

(vi). This proposition immediately follows from (v) and C_X-measurability of τ_x.

(vii). $\tau_x \vdash 1_{X=x}|Z$ is equivalent to $E(\tau_x|1_{X=x},Z) \underset{P}{=} E(\tau_x|Z)$. According to SN-Proposition (16.37), this equation follows from (vi).

(viii). Unbiasedness of $E^{X=x}(Y|Z)$ immediately follows from (vii) and Theorem 6.15.

(ix). According to (viii),

$$E^{X=x}(Y|Z) \underset{P}{=} E(\tau_x|Z).$$

If $P(X{=}x, Z{=}z) > 0$, then, according to SN-Corollary 10.39 (i), this equation implies

$$E(Y|X{=}x, Z{=}z) = E^{X=x}(Y|Z{=}z) = E(\tau_x|Z{=}z).$$

Proof of Theorem 8.14

(i). This proposition follows from $\sigma(1_{X=x}) \subset \sigma(X)$ and SN-Box 16.3 (vi) for Z being a constant.

(ii). This proposition follows from $\sigma(\tau_x) \subset \sigma(C_X)$ and SN-Box 16.3 (vi) for Z being a constant.

(iii). This proposition follows from $\sigma(\tau) \subset \sigma(C_X)$ and SN-Box 16.3 (vi) for Z being a constant.

(iv). $\tau_x \vdash X$ denotes $E(\tau_x|X) \underset{P}{=} E(\tau_x)$. Therefore, this proposition follows from (ii) and SN-Box 10.2 (vi).

(v). This proposition follows from (i), Theorem 8.11 (i), (iv), and Definition 6.2 (ii).

(vi). This proposition follows from $\sigma(Z) \subset \sigma(C_X)$, $C_X \perp\!\!\!\perp X \Leftrightarrow (C_X, Z) \perp\!\!\!\perp X$, and SN-Box 16.3 (ix) with Z taking the role of Y in that rule.

(vii). This proposition follows from (vi), $\sigma(\tau_x) \subset \sigma(C_X)$, and SN-Box 16.3 (vi).

(viii). This proposition immediately follows from (vi) and $\sigma(\tau) \subset \sigma(C_X)$, and SN-Box 16.3 (vi).

(ix). $\tau_x \vdash X|Z$ denotes $E(\tau_x|X,Z) \underset{P}{=} E(\tau_x|Z)$. Hence, this proposition follows from (vii) and SN-Remark 16.35.

(x). This proposition follows from (i), Theorem 8.11 (i), (viii), and Definition 6.11 (iii).

Proof of Theorem 8.19

(i). This proposition immediately follows from C_X-measurability of τ_x and Z and Proposition (7.7).

(ii). $\tau_x \vdash 1_{X=x} \mid Z$ is equivalent to $E(\tau_x|1_{X=x},Z) \underset{P}{=} E(\tau_x|Z)$. According to SN-Proposition (16.37), this equation follows from (i).

Proof of Theorem 8.21

(i) This proposition follows from $\sigma(1_{X=x}) \subset \sigma(X)$ and SN-Box 16.3 (vi).

(ii) This proposition follows from $\sigma(\tau_x) \subset \sigma(C_X)$ and SN-Box 16.3 (vi).

(iii) This proposition follows from $\sigma(\tau) \subset \sigma(C_X)$ and SN-Box 16.3 (vi).

(iv) $\tau \vdash X|Z$ denotes $\forall\, x \in X(\Omega)\colon E(\tau_x|X,Z) \underset{P}{=} E(\tau_x|Z)$. Hence, this proposition follows from (ii) and SN-Remark 16.35.

Proof of Corollary 8.23

(i). This proposition is an immediate implication of Corollary 8.22 (i) and Equation (8.17).

(ii). If $P(X{=}x, Z{=}z) > 0$, $P(X{=}x', Z{=}z) > 0$ and τ_x, $\tau_{x'}$ are $P^{Z=z}$-unique, then, according to Corollary 8.20 (ii), the conditional expectation values $E(Y \mid X{=}x, Z{=}z)$ and $E(Y \mid X{=}x', Z{=}z)$ are unbiased. Equations (5.2), (5.14), and (6.30) then yield the proposition.

8.7 Exercises

▷ **Exercise 8-1** Show that $C_X \perp\!\!\!\perp X$ implies:

$$P(X{=}x \mid W) \underset{P}{=} P(X{=}x), \quad \text{for each } x = 0, 1, \ldots, J,$$

if the random variable W is measurable with respect to C_X. Assume that X is discrete with a finite number of values $x = 0, 1, \ldots, J$ and that $P(X{=}x) > 0$ for all values x of X.

▷ **Exercise 8-2** Compute the conditional expectation value $E(Y \mid X = 0)$ and the expectation $E(\tau_0)$ in the example presented in Table 6.3 (p. 140) and compare these numbers to each other.

▷ **Exercise 8-3** Assume that X is discrete with a finite number of values $x = 0, 1, \ldots, J$. Which terms are unbiased if $C_X \perp\!\!\!\perp X$ holds and which ones are unbiased under $C_X \perp\!\!\!\perp X \mid Z$? (For $C_X \perp\!\!\!\perp X \mid Z$ additionally assume that all true outcome variables are P-unique and that Z is a covariate of X.)

▷ **Exercise 8-4** Describe randomized assignment of a unit to one of two treatment conditions!

▷ **Exercise 8-5** Describe *conditionally* randomized assignment of a unit to one of two treatment conditions given a covariate Z ! For simplicity, assume that Z is discrete with $P(Z{=}z) > 0$ for all its values $z \in Z(\Omega)$ and that X is dichotomous with values 0 and 1.

▷ **Exercise 8-6** Show that $E[P(X{=}x \mid U)] = P(X{=}x)$.

▷ **Exercise 8-7** Show that $U \perp\!\!\!\perp X$ implies $U \perp\!\!\!\perp X \mid Z$, provided that Z is measurable with respect to U. Assume that X is discrete with a finite number of values $x = 0, 1, \ldots, J$ and that $P(X{=}x) > 0$ for all values x of X.

▷ **Exercise 8-8** Check that all implications listed in Table 8.1 have been proven in this or the previous chapter. Use the Assumptions 6.1, that Z is a covariate of X, and the additional assumption that the true outcome variables τ_x are P-unique where necessary.

Solutions

▷ **Solution 8-1**

$$\begin{aligned}
P(X{=}x \mid W) &\underset{P}{=} E(\mathbb{1}_{X=x} \mid W) && \text{[SN-(10.4)]}\\
&\underset{P}{=} E\big(E(\mathbb{1}_{X=x} \mid C_X) \,\big|\, W\big) && \text{[SN-Box 10.2 (v)]}\\
&\underset{P}{=} E\big(P(X{=}x \mid C_X) \,\big|\, W\big) && \text{[SN-(10.4)]}\\
&\underset{P}{=} E[P(X{=}x) \mid W] && \text{[(8.2)]}\\
&\underset{P}{=} P(X{=}x), && \text{[SN-Box 10.2 (i)]}
\end{aligned}$$

for each $x = 0, 1, \ldots, J$.

▷ **Solution 8-2** According to SN-Box 9.2 (iv), the conditional expectation value $E(Y \mid X=0)$ can be computed via

$$\begin{aligned} E(Y \mid X=0) &= \sum_u E(Y \mid X=0, U=u) \cdot P(U=u \mid X=0) \\ &= (68+78+88+98) \cdot \frac{1}{10} + (106+116) \cdot \frac{3}{10} = 99.8. \end{aligned}$$

In contrast, according to SN-Theorem 6.15 as well as Equations (6.55) and (6.56), the expectation $E(\tau_0)$ can be computed via

$$\begin{aligned} E(\tau_0) &= E\big(g_x(U)\big) = \sum_u g_x(u) \cdot P(U=u) \\ &= \sum_u E(Y \mid X=0, U=u) \cdot P(U=u) \\ &= (68+78+88+98+106+116) \cdot \frac{1}{6} = 92.3333. \end{aligned}$$

Comparing $E(Y \mid X=0) = 99.8$ to $E(\tau_0) = 92.3333$ shows that $E(Y \mid X=0)$ is strongly biased [see Def. 6.2 (i)].

▷ **Solution 8-3** $C_X \perp\!\!\!\perp X$ implies that $E(Y \mid X)$, $E(Y \mid X, Z)$, all $E(Y \mid X{=}x)$, and all $E^{X=x}(Y \mid Z)$, $x \in X(\Omega)$, are unbiased. In contrast, $C_X \perp\!\!\!\perp X \mid Z$ only implies that $E(Y \mid X, Z)$ and all $E^{X=x}(Y \mid Z)$, $x \in X(\Omega)$, are unbiased, provided that we assume that all true outcome variables τ_x, $x \in X(\Omega)$, are P-unique. However, note that the causal average total effects $ATE_{xx'}$ can also be computed from $E(Y \mid X, Z)$, provided that it is unbiased.

▷ **Solution 8-4** In a random experiment, in which a unit u is sampled and assigned to one of two treatment conditions, we may assign the unit by coin toss, for instance. This ensures

$$P(X{=}1 \mid C_X) \underset{P}{=} P(X{=}1),$$

that is, the treatment probabilities do not depend on the global potential confounder C_X and therefore not on any potential confounder. If C_X is specified such that U is measurable with respect to C_X, then $P(X{=}1 \mid C_X) \underset{P}{=} P(X{=}1)$ implies that each unit has the same probability $P(X{=}1 \mid U{=}u) = P(X{=}1)$ of being assigned to treatment 1. If we assume that X is dichotomous with values 0 and 1, then this implies that each unit u has the same probability $P(X{=}0)$ to be assigned to treatment 0 as well, because

$$P(X{=}0 \mid U{=}u) = 1 - P(X{=}1 \mid U{=}u) = 1 - P(X{=}1) = P(X{=}0).$$

▷ **Solution 8-5** We consider a random experiment, in which a unit u is sampled and a value z of the covariate Z is assessed before the unit is assigned to one of the two treatment conditions. We also assume $P(Z{=}z) > 0$ for all values $z \in Z(\Omega)$. Then *Z-conditional randomized assignment* of a unit to one of the two treatment conditions refers to assigning the unit u to treatment condition 1 with probability

$$P^{Z=z}(X{=}1 \mid C_X) \underset{P^{Z=z}}{=} P^{Z=z}(X{=}1) = P(X{=}1 \mid Z{=}z). \tag{8.1}$$

This equation implies

$$\begin{aligned} P^{Z=z}(X{=}1 \mid U) &\underset{P^{Z=z}}{=} E^{Z=z}\big(P^{Z=z}(X{=}1 \mid C_X) \,\big|\, U\big) && \text{[SN-Box 10.2 (v)]} \\ &\underset{P^{Z=z}}{=} E^{Z=z}\big(P^{Z=z}(X{=}1) \,\big|\, U\big) && \text{[(8.1)]} \\ &\underset{P^{Z=z}}{=} P^{Z=z}(X{=}1) && \text{[SN-Box 10.2 (i)]} \\ &= P(X{=}1 \mid Z{=}z). && \text{[(3.28)]} \end{aligned}$$

Hence, all units with the same value z of the covariate Z have the same probability to be assigned to treatment 1. This assignment procedure allows for different treatment probabilities for units for which we observe different values z of the covariate Z. It creates $C_X \perp\!\!\!\perp X \,|\, Z$, which implies that $E(Y \,|\, X, Z)$ and all $E^{X=x}(Y \,|\, Z)$, $x \in X(\Omega)$, are unbiased.

▷ **Solution 8-6**

$$\begin{aligned} E\big(P(X{=}x \,|\, U)\big) &\underset{P}{=} E\big(E(1_{X=x} \,|\, U)\big) && [SN\text{-}(10.4)] \\ &\underset{P}{=} E(1_{X=x}) && [SN\text{-}Box\ 10.2\ (\text{iv})] \\ &= P(X{=}x). && [SN\text{-}(6.5)] \end{aligned}$$

▷ **Solution 8-7** If $P(X{=}x) > 0$ for all values x of X, then $X \perp\!\!\!\perp U$ is equivalent to

$$P(X{=}x \,|\, U) \underset{P}{=} P(X{=}x), \quad \forall\, x \in \{0, 1, \ldots, J\}.$$

If Z is measurable with respect to U, then $\sigma(U, Z) = \sigma(U)$ and this equation implies

$$P(X{=}x \,|\, U, Z) \underset{P}{=} P(X{=}x \,|\, U) \underset{P}{=} P(X{=}x), \quad \forall\, x \in \{0, 1, \ldots, J\}, \tag{8.2}$$

which is equivalent to

$$E(1_{X=x} \,|\, U, Z) \underset{P}{=} E(1_{X=x} \,|\, U) \underset{P}{=} E(1_{X=x}), \quad \forall\, x \in \{0, 1, \ldots, J\}. \tag{8.3}$$

Hence,

$$\begin{aligned} P(X{=}x \,|\, Z) &\underset{P}{=} E(1_{X=x} \,|\, Z) && [\text{SN-}(10.4)] \\ &\underset{P}{=} E\big(E(1_{X=x} \,|\, U, Z) \,\big|\, Z\big) && [\text{SN-Box 10.2 (v)}] \\ &\underset{P}{=} E\big(E(1_{X=x}) \,\big|\, Z\big) && [(8.3)] \\ &\underset{P}{=} E(1_{X=x}) && [\text{SN-Box 10.2, (i)}] \\ &= P(X{=}x) && [\text{SN-}(6.5)] \\ &\underset{P}{=} P(X{=}x \,|\, U, Z), && [(8.2)] \end{aligned}$$

for each $x = 0, 1, \ldots, J$. Hence,

$$P(X{=}x \,|\, U, Z) \underset{P}{=} P(X{=}x \,|\, Z), \quad \forall\, x \in \{0, 1, \ldots, J\}.$$

This equation is equivalent to $X \perp\!\!\!\perp U \,|\, Z$ if X is discrete with a finite number of values $x = 0, 1, \ldots, J$ (see SN-Rem. 16.27).

▷ **Solution 8-8** We check the implications listed in Table 8.1 row-wise.
$C_X \perp\!\!\!\perp X \;\Rightarrow\; \tau \perp\!\!\!\perp X$. This is the proposition of Theorem 8.14 (iii).
$C_X \perp\!\!\!\perp X \;\Rightarrow\; \tau \perp\!\!\!\perp X \,|\, Z$. This is the proposition of Theorem 8.14 (viii).
$C_X \perp\!\!\!\perp X \;\Rightarrow\; C_X \perp\!\!\!\perp X \,|\, Z$. This is the proposition of Theorem 8.14 (vi).
$C_X \perp\!\!\!\perp X \,|\, Z \;\Rightarrow\; \tau \perp\!\!\!\perp X \,|\, Z$. This is the proposition of Theorem 8.21 (iii).

Chapter 9
Reichenbach-Suppes Conditions

In chapter 8, we introduced the Fisher conditions, a first class of empirically testable causality conditions that focus on independence and Z-conditional independence of X and a global potential confounder C_X, where Z denotes a covariate of X. These causality conditions, $C_X \perp\!\!\!\perp X$ and $C_X \perp\!\!\!\perp X \,|\, Z$, also apply if X is continuous. We emphasized that $C_X \perp\!\!\!\perp X$ implies $C_X \perp\!\!\!\perp X \,|\, Z$. If X is discrete with values $0, 1, \ldots, J$, we showed that $C_X \perp\!\!\!\perp X \,|\, Z$ implies strong ignorability, that is, $\tau \perp\!\!\!\perp X \,|\, Z$, where $\tau = (\tau_0, \tau_1, \ldots, \tau_J)$ consists of the $J+1$ true outcome variables τ_x.

The causality conditions to be introduced in this chapter are called the Reichenbach-Suppes conditions. They are also empirically testable, also apply if X is continuous, and have a close relationship to strong ignorability, provided that X is discrete. We start introducing these causality conditions and treat their implication structure. Then we study their implications on other causality conditions and illustrate them by some examples. Finally, we discuss the methodological implications of this class of causality conditions. In particular, we discuss their role in covariate selection in the analysis of causal conditional and average total effects.

9.1 Reichenbach-Suppes Conditions

The Fisher conditions focus on (conditional) independence of the potential cause X and potential confounders of X. In contrast, the causality conditions introduced in this chapter, focus on conditional independence or conditional mean-independence of the outcome variable Y and potential confounders of X. These conditions will summarily be referred to as the *Reichenbach-Suppes conditions*, honoring two pioneers who made early contributions to the theory of causal effects (see, e. g., Reichenbach, 1956; Suppes, 1970).

9.1.1 X-Conditional Independence and Mean-Independence of Y and C_X

Remark 9.1 (X-Conditional Independence of Y and C_X) Under the Assumptions 8.1 we denote X-conditional independence of the outcome variable Y and a global potential confounder C_X of X by

$$Y \perp\!\!\!\perp C_X \,|\, X$$

[see Def. 7.1(ii)]. As mentioned before, an extensive treatment of conditional independence and its properties is found, for example, in SN-chapter 16. Note that this concept also applies if Y, C_X, and X are random variables that are not necessarily numerical. $\lhd$

Remark 9.2 (X-Conditional Mean-Independence) Under the Assumptions 8.1 and the additional assumption that Y is real-valued with $E(Y^2) < \infty$,

$$E(Y|X, C_X) \underset{P}{=} E(Y|X) \tag{9.1}$$

defines *X-conditional mean-independence of Y from C_X*. We use $Y \vdash C_X|X$ to abbreviate this equation, that is,

$$Y \vdash C_X|X \quad \Leftrightarrow \quad E(Y|X, C_X) \underset{P}{=} E(Y|X) \tag{9.2}$$

(see Rem. 3.71). ◁

Remark 9.3 (Intuitive Meaning of $Y \vdash C_X|X$) Referring to the filtration $(\mathscr{F}_t, t \in T)$ and the σ-algebra $\mathscr{F}_2$ introduced in Definition 4.4, note that Equation (9.1) is equivalent to

$$E(Y|\mathscr{F}_2) \underset{P}{=} E(Y|X), \tag{9.3}$$

because $\sigma(X, C_X) = \mathscr{F}_2$ (see Def. 4.4, SN-Def. 10.2, and SN-Rem. 10.3). Intuitively speaking, this equation means that X is the only cause of Y in the set of $\mathscr{F}_2$-measurable random variables, and that there is no potential confounder of X affecting the conditional expectation of Y, once we condition on X. By definition, a potential confounder of X is measurable with respect to $\mathscr{F}_1$ and we defined a global potential confounder C_X such that $\sigma(C_X) = \mathscr{F}_1$. Hence, Equations (9.1) and (9.3) do not exclude mediators and other intermediate variables (that are posterior to X and prior to Y) to determine the conditional expectation of Y over and above X. Nevertheless, $Y \vdash C_X|X$ is unrealistically restrictive in most applications. However, this limitation does not apply to its extension presented in section 9.1.2. ◁

Remark 9.4 (Falsifiability of $Y \vdash C_X|X$) The condition $Y \vdash C_X|X$ implies some propositions that lend themselves for falsification. If W is measurable with respect to C_X, then $Y \vdash C_X|X$ implies $Y \vdash W|X$, that is, it implies

$$E(Y|X, W) \underset{P}{=} E(Y|X) \tag{9.4}$$

(see Exercise 9-2). Hence, if $Y \vdash C_X|X$ holds, then there is no C_X-measurable random variable on $(\Omega, \mathscr{A}, P)$, that determines the conditional expectation of Y, once we condition on X. Therefore, $Y \vdash C_X|X$ can easily be tested empirically in the sense of falsifiability. We simply have to select a potential confounder of X, that is, a random variable W that is measurable with respect to C_X, and see if Equation (9.4) holds. If it does not hold, then the hypothesis $Y \vdash C_X|X$ is falsified. In an empirical study, this can be done using appropriate statistical procedures of regression analysis. ◁

Remark 9.5 (Implications of $Y \vdash C_X|X$ if U is C_X-Measurable) Assume that X represents a treatment variable and C_X is a global potential confounder such that the observational-unit variable U is measurable with respect to C_X Then $Y \vdash C_X|X$ implies

$$E(Y|X, U) \underset{P}{=} E(Y|X). \tag{9.5}$$

If $P(X{=}x, U{=}u) > 0$, for all pairs $(x, u) \in X(\Omega) \times U(\Omega)$, then this equation is equivalent to

$$\forall (x, u) \in X(\Omega) \times U(\Omega): \; E(Y|X{=}x, U{=}u) \; = \; E(Y|X{=}x).$$

Hence, in this case, the conditional expectation values $E(Y|X{=}x, U{=}u)$ are identical for all units u, and this holds for each value x of X. As mentioned before, this condition is unrealistically restrictive in most applications. However, this limitation does not apply to its extension presented in section 9.1.2. ◁

9.1.2 (X,Z)-Conditional Independence of Y and C_X

If X represents a treatment variable, often neither X-conditional independence of Y and C_X nor X-conditional mean-independence of Y from C_X will hold. Much more realistic causality conditions are (X,Z)-conditional independence of Y and C_X as well as (X,Z)-conditional mean-independence of Y from C_X.

Remark 9.6 ((X,Z)-Conditional Independence of Y and C_X) Under the Assumptions 8.4, (X,Z)-conditional independence of Y and a global potential confounder C_X of X is denoted by

$$Y \perp\!\!\!\perp C_X \,|\, (X,Z)$$

Again note that none of the random variables involved needs to be numerical [see again Def. 7.1(ii) and SN-chapter 16)]. ◁

As will be shown in Theorem 9.12, $Y \perp\!\!\!\perp C_X \,|\, (X,Z)$ follows from $Y \perp\!\!\!\perp C_X | X$. Furthermore, if Y is numerical, nonnegative, or with finite expectation, then it implies (X,Z)-conditional mean-independence of Y from the global potential confounder C_X (see Th. 9.13), which is also sufficient for unbiasedness of the conditional expectation $E(Y|X,Z)$, provided that we assume that all true outcome variables τ_x, $x = 0,1,\ldots,J$, are P-unique (see Th. 9.26).

Remark 9.7 ((X,Z)-Conditional Mean-Independence of Y from C_X) If, additionally to the Assumptions 8.4, Y is numerical, nonnegative or with finite expectation, then

$$E(Y|X,C_X) \underset{P}{=} E(Y|X,Z) \tag{9.6}$$

defines *(X,Z)-conditional mean-independence of Y from C_X*, which will be abbreviated by $Y \vdash C_X \,|\, (X,Z)$, that is,

$$Y \vdash C_X \,|\, (X,Z) \quad \Leftrightarrow \quad E(Y|X,C_X) \underset{P}{=} E(Y|X,Z) \tag{9.7}$$

(see again Rem. 3.71). ◁

Remark 9.8 (An Equivalent Formulation of $Y \vdash C_X \,|\, (X,Z)$) Because, according to Assumptions 8.4, Z is measurable with respect to C_X,

$$\sigma(X,Z,C_X) \;=\; \sigma(X,C_X)$$

(see SN-Defs. 2.26 and 2.43), which in turn implies

$$E(Y|X,Z,C_X) \underset{P}{=} E(Y|X,C_X) \tag{9.8}$$

[see SN-Eq. (10.1)]. Hence, if Z is measurable with respect to C_X, and this is part of the Assumptions 8.4, then

$$E(Y|X,Z,C_X) \underset{P}{=} E(Y|X,Z) \tag{9.9}$$

is equivalent to Equation (9.6). ◁

The conditions $Y \perp\!\!\!\perp C_X | X$, $Y \perp\!\!\!\perp C_X \,|\, (X,Z)$, $Y \vdash C_X | X$, and $Y \vdash C_X \,|\, (X,Z)$ will also be referred to as the *Reichenbach-Suppes conditions (for total effects)*. They share the focus on conditional independence or conditional mean-independence of Y and potential confounders of X. In contrast, the Fisher conditions treated in chapter 8 share the focus on (Z-conditional) independence of X and potential confounders of X.

Remark 9.9 (Intuitive Meaning of $Y \vdash C_X \,|\, (X,Z)$) Referring to the filtration $(\mathscr{F}_t, t \in T)$ and the σ-algebra $\mathscr{F}_2$ introduced in Definition 4.4, note that Equation (9.6) is equivalent to

$$E(Y|\mathscr{F}_2) \underset{P}{=} E(Y|X,Z), \tag{9.10}$$

because $\sigma(X, C_X) = \mathscr{F}_2$ (see again Def. 4.4), provided, of course, that Z is a covariate of X. According to Equation (9.10), there are no $\mathscr{F}_2$-measurable random variables other than X and Z that determine the conditional expectation of Y, once we condition on X and Z. Intuitively speaking, conditioning on Z, the random variable X is the only cause of Y, in the set all potential confounders of X. Again, note that Equation (9.10) does not preclude mediators and intermediate variables to have an effect on Y over and above X and Z, because such intermediate variables are not $\mathscr{F}_2$-measurable. ◁

Remark 9.10 (Falsifiability of $Y \vdash C_X \,|\, (X,Z)$) The condition $Y \vdash C_X \,|\, (X,Z)$ implies a proposition that lends itself for falsification. If Z and W are measurable with respect to C_X, then $Y \vdash C_X \,|\, (X,Z)$ implies

$$E(Y|X,Z,W) \underset{P}{=} E(Y|X,Z) \tag{9.11}$$

(see Exercise 9-2). If $Y \vdash C_X \,|\, (X,Z)$, and with it Equation (9.6), holds, then there is no potential confounder W of X that determines the conditional expectation of Y, once we condition on X and Z.

Hence, the condition $Y \vdash C_X \,|\, (X,Z)$ can easily be tested empirically, again only in the sense of falsifiability. In order to test $Y \vdash C_X \,|\, (X,Z)$, we may select a variable W that is measurable with respect to C_X and see if Equation (9.11) holds. If it does not hold, then we can conclude that $Y \vdash C_X \,|\, (X,Z)$ is not satisfied. ◁

Remark 9.11 (Implications of $Y \vdash C_X \,|\, (X,Z)$ if U is C_X-Measurable) Assume that X represents a treatment variable and C_X is a potential global potential confounder such that the observational-unit variable U is measurable with respect to C_X. Then $Y \vdash C_X \,|\, (X,Z)$ implies

$$E(Y|X,Z,U) \underset{P}{=} E(Y|X,Z)\,. \tag{9.12}$$

If $P(X{=}x, Z{=}z, U{=}u) > 0$, then this equation in turn implies

$$E(Y|X{=}x, Z{=}z, U{=}u) \;=\; E(Y|X{=}x, Z{=}z)\,.$$

Hence, the conjunction of $Y \vdash C_X \,|\, (X,Z)$, $\sigma(U) \subset \sigma(C_X)$, and $P(X{=}x, Z{=}z, U{=}u) > 0$ implies that there are no differences between units u with respect to their conditional expectation values of the outcome variable Y given unit u, treatment condition x, and covariate value z. Table 9.1 (p. 211) displays an example. ◁

9.2 Implication Structure Among the Reichenbach-Suppes Conditions

Before studying the implications of the Reichenbach-Suppes conditions on the Rosenbaum-Rubin conditions and on unbiasedness, we consider the implication structure among the Reichenbach-Suppes conditions. Note that the assumptions neither include that X is discrete nor P-uniqueness of the true outcome variables.

Theorem 9.12 (An Implication of $Y \perp\!\!\!\perp C_X|X$)
Let the Assumptions 8.1 hold.

(i) If Y be real-valued with $E(Y^2) < \infty$, then

$$Y \perp\!\!\!\perp C_X|X \quad \Rightarrow \quad Y \vdash C_X|X. \tag{9.13}$$

(ii) If Z is a covariate of X, then

$$Y \perp\!\!\!\perp C_X|X \quad \Rightarrow \quad Y \perp\!\!\!\perp C_X|(X,Z). \tag{9.14}$$

(Proof p. 216)

According to Theorem 9.12 (i), $Y \perp\!\!\!\perp C_X|X$ implies X-conditional mean-independence of Y from C_X, if Y is real-valued with a finite second moment. Note that $Y \vdash C_X|X$ is another sufficient condition for unbiasedness of the conditional expectation $E(Y|X)$, provided that we assume that all true outcome variables τ_x, $x = 0,1,\ldots,J$, are P-unique (see Th. 9.21).

According to Theorem 9.12 (ii), $Y \perp\!\!\!\perp C_X|X$ also implies $Y \perp\!\!\!\perp C_X|(X,Z)$ if Z is a covariate of X. According to Theorem 9.13, $Y \perp\!\!\!\perp C_X|(X,Z)$ in turn implies $Y \vdash C_X|(X,Z)$.

Theorem 9.13 (An Implication of $Y \perp\!\!\!\perp C_X|(X,Z)$)
Let the Assumptions 8.1 hold and let Z be a covariate of X. Then

$$Y \perp\!\!\!\perp C_X|(X,Z) \quad \Rightarrow \quad Y \vdash C_X|(X,Z). \tag{9.15}$$

(Proof p. 216)

Hence, if Z is a covariate of X, then $Y \perp\!\!\!\perp C_X|(X,Z)$ implies (X,Z)-conditional mean-independence of Y from the global potential confounder C_X. Both conditions occuring in Proposition (9.15), the premise and the conclusion, are sufficient for unbiasedness of the conditional expectation $E(Y|X,Z)$, provided that we assume that all true outcome variables τ_x, $x = 0,1,\ldots,J$, are P-unique. This will be shown in Theorem 9.26.

In the following theorem we present two conditions that are equivalent to $Y \vdash C_X|X$ and $Y \vdash C_X|(X,Z)$, respectively, provided that, for all values x of X, $P(X{=}x) > 0$ and the true outcome variables τ_x are P-unique. Reading this lemma, remember that τ_x denotes a conditional expectation $E^{X=x}(Y|C_X)$ (see Def. 4.16), and that P-uniqueness of $E^{X=x}(Y|C_X)$ is equivalent to

$$P(X{=}x|C_X) \gneqq_P 0. \tag{9.16}$$

Theorem 9.14 (Equivalent Conditions of $Y \vdash C_X|X$ and $Y \vdash C_X|(X,Z)$)
Let the Assumptions 6.1 hold and assume that τ_x is P-unique for all $x \in X(\Omega)$.

(i) Then $Y \vdash C_X|X$ is equivalent to

$$\tau_x \;=\; E^{X=x}(Y|C_X) \;=_P\; E^{X=x}(Y), \quad \forall\, x \in X(\Omega). \tag{9.17}$$

(ii) If, additionally, Z is a covariate of X, then $Y \vdash C_X \,|\, (X,Z)$ is equivalent to

$$\tau_x \;=\; E^{X=x}(Y|C_X) \underset{P}{=} E^{X=x}(Y|Z), \quad \forall\, x \in X(\Omega)\,. \tag{9.18}$$

(Proof p. 216)

Remark 9.15 (Constant True Outcome Variables) According to Theorem 9.14 (i), $Y \vdash C_X|X$ is equivalent to

$$\tau \underset{P}{=} \big(E(Y|X=0), E(Y|X=1), \ldots, E(Y|X=J)\big), \tag{9.19}$$

where $\tau = (\tau_0, \tau_1, \ldots, \tau_J)$ is a vector of true outcome variables of the $J+1$ values $x \in X(\Omega)$. ◁

Remark 9.16 (True Outcome Variables as Functions of Z) According to Theorem 9.14 (ii), $Y \vdash C_X \,|\, (X,Z)$ is equivalent to

$$\tau \underset{P}{=} \big(E^{X=0}(Y|Z), E^{X=1}(Y|Z), \ldots, E^{X=J}(Y|Z)\big), \tag{9.20}$$

where τ is the vector specified in Remark 9.15. Hence, $Y \vdash C_X \,|\, (X,Z)$ means that Z contains all potential confounders of X that determine the C_X-conditional expectation of Y. It does not mean that there are no intermediate variables in between X and Y that determine Y over and above the (possibly multivariate) covariate Z. ◁

9.3 Further Implications of the Reichenbach-Suppes Conditions

Now we present the most important implications of the Reichenbach-Suppes conditions on the Rosenbaum-Rubin conditions and on unbiasedness.

9.3.1 Implications of X-Conditional Mean-Independence of Y from C_X

In Theorem 9.17, we consider a single value x of X with $P(X{=}x) > 0$ and the conditional expectation value $E(Y|X{=}x)$. Reading this theorem, remember that, according to Lemma 7.4, $\tau_x \perp\!\!\!\perp 1_{X=x}$ is equivalent to

$$P(X{=}x\,|\,\tau_x) \underset{P}{=} P(X{=}x).$$

The premise of Theorem 9.17 is $Y \vdash C_X \,|\, X{=}x$, which is defined by

$$Y \vdash C_X \,|\, X{=}x \quad :\Leftrightarrow \quad E^{X=x}(Y|C_X) \underset{P^{X=x}}{=} E^{X=x}(Y). \tag{9.21}$$

where

$$E^{X=x}(Y|C_X) \underset{P^{X=x}}{=} E^{X=x}(Y) \tag{9.22}$$

defines mean independence of Y from C_X with respect to the conditional probability measure $P^{X=x}$ (see Rem. 3.71 and section 3.3.6).

Theorem 9.17 (Implications of $Y \vdash C_X \,|\, X{=}x$)
Let the Assumptions 5.1 hold and assume that τ_x is P-unique. Then

$$Y \vdash C_X \,|\, X{=}x \quad \Rightarrow \quad \tau_x \perp\!\!\!\perp X. \tag{9.23}$$

(Proof p. 216)

Remark 9.18 (Further Implications of $Y \vdash C_X \,|\, X{=}x$) According to Proposition (9.23) and Table 7.5, $Y \vdash C_X \,|\, X{=}x$ also implies

(i) $\tau_x \vdash \mathbb{1}_{X=x}$.
(ii) $\tau_x \perp\!\!\!\perp \mathbb{1}_{X=x}$.
(iii) $\tau_x \vdash \mathbb{1}_{X=x}$.
(iv) $E(Y \,|\, X{=}x)$ is unbiased.

◁

Remark 9.19 (Methodological Implications) Remember, $\tau_x \perp\!\!\!\perp \mathbb{1}_{X=x}$ also follows from $C_X \perp\!\!\!\perp \mathbb{1}_{X=x}$ [see Th. 8.11 (ii)]. However, in contrast to $C_X \perp\!\!\!\perp \mathbb{1}_{X=x}$, assuming $Y \vdash C_X \,|\, X{=}x$ does not imply that τ_x is P-unique [cf. Th. 8.11 (i)]. This is why in Theorem 9.17 we need the additional assumption that τ_x is P-unique. ◁

Remark 9.20 (Falsifiability) It should also be emphasized that $Y \vdash C_X \,|\, X{=}x$ is empirically falsifiable, because $Y \vdash C_X \,|\, X{=}x$ implies

$$E^{X=x}(Y \,|\, W) \underset{P^{X=x}}{=} E^{X=x}(Y), \tag{9.24}$$

whenever W is a C_X-measurable random variable [see Prop. (9.21) and Exercise 9-1]. This mean-independence can be tested by standard regression techniques restricting the data analysis to the subsample in treatment x. Of course, Equation (9.24) will rarely hold in empirical applications. ◁

In Theorem 9.17, we conditioned on a single value x of X with $P(X{=}x) > 0$ and considered the conditional expectation value $E(Y \,|\, X{=}x)$. Now we condition on the random variable X and focus the conditional expectation $E(Y \,|\, X)$. Reading the following theorem, remember that $Y \vdash C_X \,|\, X$ is a shortcut for

$$E(Y \,|\, X, C_X) \underset{P}{=} E(Y \,|\, X)\,. \tag{9.25}$$

Furthermore, in Proposition (7.16) we already defined

$$\tau \vdash X \quad :\Leftrightarrow \quad \forall\, x \in X(\Omega)\colon\ E(\tau_x \,|\, X) \underset{P}{=} E(\tau_x). \tag{9.26}$$

Theorem 9.21 (Implications of $Y \vdash C_X \,|\, X$)
Let the Assumptions 6.1 hold, assume that τ_x is P-unique for all $x \in X(\Omega)$, and define $\tau = (\tau_0, \tau_1, \ldots, \tau_J)$.

(i) Then $Y \vdash C_X \,|\, X$ implies $\tau \perp\!\!\!\perp X$.

If, additionally, Z is a covariate of X, then $Y \vdash C_X \,|\, X$ implies:

(ii) $\tau \perp\!\!\!\perp X \,|\, Z$.
(iii) $Y \vdash C_X \,|\, (X, Z)$.

(Proof p. 216)

Remark 9.22 (Further Implications of $Y \vdash C_X \,|\, X$) If we assume that all τ_x, $x \in X(\Omega)$, are P-unique, then, according to Theorem 9.21 (i) and Table 7.5, $Y \vdash C_X \,|\, X$ also implies that $E(Y \,|\, X)$ is unbiased and all other causality conditions listed in Table 7.5. Furthermore, according to Theorem 9.21 (ii) and Table 7.6, $Y \vdash C_X \,|\, X$ also implies all other causality conditions listed in Table 7.6, and that $E(Y \,|\, X, Z)$ is unbiased. ◁

9.3.2 Implications of (X, Z)-Conditional Mean-Independence of Y from C_X

Now we extend the results of section 9.3.1 to the case in which we condition not only on X but additionally on a covariate Z of X. We start considering a single conditional expectation $E^{X=x}(Y \,|\, Z)$ and a single conditional expectation value $E(Y \,|\, X=x, Z=z)$. Remember, the term $E^{X=x}(Y \,|\, Z)$ denotes the Z-conditional expectation of Y with respect to the $(X=x)$-conditional probability measure $P^{X=x}$. Also remember that *unbiasedness* of the conditional expectation $E^{X=x}(Y \,|\, Z)$ is defined by assuming P-uniqueness of τ_x and

$$E^{X=x}(Y \,|\, Z) \underset{P}{=} E(\tau_x \,|\, Z) \tag{9.27}$$

[see Def. 6.11 (ii)]. Analogously to (9.21) we define

$$Y \vdash C_X \,|\, X=x, Z \quad :\Leftrightarrow \quad E^{X=x}(Y \,|\, C_X) \underset{P^{X=x}}{=} E^{X=x}(Y \,|\, Z), \tag{9.28}$$

where $E^{X=x}(Y \,|\, C_X)$ and $E^{X=x}(Y \,|\, Z)$ denote conditional expectations with respect to the measure $P^{X=x}$ (see Def. 3.79).

Theorem 9.23 (Implications of $Y \vdash C_X \,|\, X=x, Z$)
Let the Assumptions 5.1 hold, let Z be a covariate of X, and assume that τ_x is P-unique. Then

$$Y \vdash C_X \,|\, X=x, Z \quad \Rightarrow \quad \tau_x \perp\!\!\!\perp X \,|\, Z. \tag{9.29}$$

(Proof p. 217)

Remark 9.24 (Further Implications of $Y \vdash C_X \,|\, X=x, Z$) If Z is a covariate of X and we assume that τ_x is P-unique, then, according to Proposition (9.29) and Table 7.6, $Y \vdash C_X \,|\, X=x, Z$ also implies

(i) $\tau_x \vdash 1_{X=x} \,|\, Z$.
(ii) $\tau_x \perp\!\!\!\perp 1_{X=x} \,|\, Z$.
(iii) $\tau_x \vdash X \,|\, Z$.
(iv) $E^{X=x}(Y \,|\, Z)$ is unbiased.

◁

Remember, if Z is a covariate of X, $E^{X=x}(Y|C_X)$ is $P^{Z=z}$-unique, $P(X=x, Z=z) > 0$, and

$$E(Y|X=x, Z=z) \;=\; E(\tau_x|Z=z), \tag{9.30}$$

then $E(Y|X=x, Z=z)$ is unbiased [see Def. 6.11 (i)]. Also remember that P-uniqueness of $E^{X=x}(Y|C_X)$ implies that it is also $P^{Z=z}$-unique [see SN-Box 14.1 (v)]. In this sense, $P^{Z=z}$-uniqueness of $E^{X=x}(Y|C_X)$ is less restrictive than assuming that $E^{X=x}(Y|C_X)$ is P-unique.

Theorem 9.25 (Unbiasedness of $E(Y|X=x, Z=z)$)
Let the Assumptions 5.1 hold, let Z be a covariate of X with value space $(\Omega'_Z, \mathscr{A}'_Z)$, assume that $z \in \Omega'_Z$ such that $P(X=x, Z=z) > 0$, and that τ_x is $P^{Z=z}$-unique. Then $Y \vdash C_X|X=x, Z$ implies that $E(Y|X=x, Z=z)$ is unbiased.

(Proof p. 217)

Reading the following theorem, remember that $\sigma(Z) \subset \sigma(C_X)$ if we assume that Z is a covariate of X. Under this assumption,

$$Y \vdash C_X|(X, Z) \quad \Leftrightarrow \quad E(Y|X, C_X) \underset{P}{=} E(Y|X, Z)\,. \tag{9.31}$$

Theorem 9.26 (Implications of $Y \vdash C_X|(X, Z)$)
Let the Assumptions 5.1 hold, let Z be a covariate of X, assume that, for all $x \in X(\Omega)$, the true outcome variables τ_x are P-unique, and define $\tau = (\tau_0, \tau_1, \ldots, \tau_J)$. Then

$$Y \vdash C_X|(X, Z) \quad \Rightarrow \quad \tau \perp\!\!\!\perp X|Z. \tag{9.32}$$

(Proof p. 218)

Remark 9.27 (Further Implications of $Y \vdash C_X|(X, Z)$) If we assume that Z is a covariate of X and τ_x is P-unique for all $x \in X(\Omega)$, then, according to Proposition (9.32) and Table 7.6, $Y \vdash C_X|(X, Z)$ also implies that $E(Y|X, Z)$ is unbiased and that all eight other causality conditions listed in that table hold.

◁

If we presume that z is a value of a covariate Z of X such that $P(X=x, Z=z) > 0$, then, according to the following theorem $Y \vdash C_X|(X, Z)$ also implies that $E(Y|X=x, Z=z)$ is unbiased, provided that we assume that τ_x is at least $P^{Z=z}$-unique. [Remember, according to SN-Box 14.1 (v), P-uniqueness of τ_x implies that τ_x is also $P^{Z=z}$-unique.]

Theorem 9.28 ($Y \vdash C_X|(X, Z)$ Implies Unbiasedness of $E(Y|X=x, Z=z)$)
Let the Assumptions 5.1 hold, let Z be a covariate of X with value space $(\Omega'_Z, \mathscr{A}'_Z)$, assume that $z \in \Omega'_Z$ such that $P(X=x, Z=z) > 0$, and that τ_x is $P^{Z=z}$-unique. Then

$$Y \vdash C_X|(X, Z) \quad \Rightarrow \quad E(Y|X=x, Z=z) \text{ is unbiased.} \tag{9.33}$$

(Proof p. 218)

Remark 9.29 (Covariate Selection) In contrast to $C_X \perp\!\!\!\perp X|Z$ (see Th. 8.21), the condition $Y \vdash C_X|(X, Z)$ cannot be created by a design technique such as randomization. Instead

we can only try to find a (possibly multivariate) covariate Z such that $Y \vdash C_X \,|\, (X,Z)$ holds. In other words, $Y \vdash C_X \,|\, (X,Z)$ can only be used for covariate selection (see section 9.5 for more details). ◁

Prima Facie Effect Functions

Remember, a Z-conditional prima facie effect function is defined by

$$PFE_{Z;xx'} \;=\; E^{X=x}(Y|Z) \;-\; E^{X=x'}(Y|Z).$$

Hence, unbiasedness of the conditional expectations $E^{X=x}(Y|Z)$ and $E^{X=x'}(Y|Z)$ has implications on unbiasedness of $PFE_{Z;xx'}$ and on the $(Z{=}z)$-conditional prima facie effect

$$PFE_{Z;xx'}(z) \;=\; E(Y|X{=}x, Z{=}z) - E(Y|X{=}x', Z{=}z),$$

which is uniquely defined if $P(X{=}x, Z{=}z)$, $P(X{=}x', Z{=}z) > 0$.

Corollary 9.30 (Unbiasedness of Conditional Prima Facie Effects)
Let the Assumptions 5.1 hold and let Z be a covariate of X with value space $(\Omega'_Z, \mathcal{A}'_Z)$.

(i) If τ_x and $\tau_{x'}$ are P-unique, then $Y \vdash C_X \,|\, (X,Z)$ implies

$$PFE_{Z;xx'}(Z) \;\underset{P}{=}\; CTE_{Z;xx'}(Z). \tag{9.34}$$

(ii) If $z \in \Omega'_Z$, $P(X{=}x, Z{=}z)$, $P(X{=}x', Z{=}z) > 0$, and τ_x, $\tau_{x'}$ are $P^{Z=z}$-unique, then $Y \vdash C_X \,|\, (X,Z)$ implies

$$PFE_{Z;xx'}(z) \;=\; CTE_{Z;xx'}(z). \tag{9.35}$$

(Proof p. 218)

Hence, under the assumptions of Corollary 9.30, the Z-conditional prima facie effect function $PFE_{Z;xx'}(Z)$ is unbiased if (a) τ_x and $\tau_{x'}$ are P-unique and (b) $Y \vdash C_X \,|\, (X,Z)$ holds. Correspondingly, the prima facie effect $PFE_{Z;xx'}(z)$ is unbiased, provided that (a) τ_x and $\tau_{x'}$ are $P^{Z=z}$-unique, (b) $P(X{=}x, Z{=}z) > 0$ and $P(X{=}x', Z{=}z) > 0$, and (c) $Y \vdash C_X \,|\, (X,Z)$.

9.4 Example

Now we illustrate (X,Z)-conditional mean-independence of Y from C_X by a numerical example.

Example 9.31 Table 9.1 (p. 211) shows an example in which $Y \vdash C_X \,|\, (X,Z)$ holds. In this example, $C_X = U$ and Z is U-measurable, which implies

$$E(Y|X,Z,C_X) \;\underset{P}{=}\; E(Y|X,C_X) \;\underset{P}{=}\; E(Y|X,U).$$

Therefore, the values of the true outcome variables are the individual conditional expectation values $E(Y|X{=}x, U{=}u)$. As is easily seen in the table, the true outcomes under control

Table 9.1. (X,Z)-conditional mean-independence of Y on C_X

		Fundamental parameters				Derived parameters		
Person variable U	Sex Z	$P(U=u)$	$P(X=1\,\vert\,U)$	$\tau_0 = E^{X=0}(Y\,\vert\,U)$	$\tau_1 = E^{X=1}(Y\,\vert\,U)$	$\delta_{10} = \tau_1 - \tau_0$	$P(U=u\,\vert\,X=0)$	$P(U=u\,\vert\,X=1)$
u_1	male	1/6	6/8	90	95	5	2/19	6/19
u_2	male	1/6	7/8	90	95	5	1/19	7/19
u_3	male	1/6	6/8	90	95	5	2/19	6/19
u_4	male	1/6	7/8	90	95	5	1/19	7/19
u_5	female	1/6	1/8	100	110	10	7/19	1/19
u_6	female	1/6	2/8	100	110	10	6/19	2/19

	$x=0$	$x=1$	
$E(\tau_x)$:	93.333	100	$ATE_{10} = 6.667$
$E(Y\,\vert\,X=x)$:	96.842	96.552	$PFE_{10} = -.29$
$E(\tau_x\,\vert\,Z=m)$:	90	95	$CTE_{Z;10}(m) = 5$
$E(Y\,\vert\,X=x,Z=m)$:	90	95	$PFE_{Z;10}(m) = 5$
$E(\tau_x\,\vert\,Z=f)$:	100	110	$CTE_{Z;10}(f) = 10$
$E(Y\,\vert\,X=x,Z=f)$:	100	110	$PFE_{Z;10}(f) = 10$

and the true total effects of treatment 1 vs. 0 are the same for all male individuals and they are the same for all female individuals. For all males, the true outcomes *under control* are

$$E(Y\,\vert\,X=0,Z=m,U=u_i) \;=\; E(Y\,\vert\,X=0,Z=m) \;=\; 90, \quad i=1,\ldots,4,$$

and for all females they are

$$E(Y\,\vert\,X=0,Z=f,U=u_i) \;=\; E(Y\,\vert\,X=0,Z=f) \;=\; 100, \quad i=5,6.$$

Under treatment, the true outcomes for all males are

$$E(Y\,\vert\,X=1,Z=m,U=u_i) \;=\; E(Y\,\vert\,X=1,Z=m) \;=\; 95, \quad i=1,\ldots,4,$$

and for all females they are

$$E(Y\,\vert\,X=1,Z=f,U=u_i) \;=\; E(Y\,\vert\,X=1,Z=f) \;=\; 110, \quad i=5,6.$$

Looking at the columns for τ_0 and τ_1 in Table 9.1 shows that, in both treatment conditions, the expectations of the outcomes only depend on Z, that is,

$$\tau_x \underset{P}{=} E^{X=x}(Y\,\vert\,U) \underset{P}{=} E^{X=x}(Y\,\vert\,Z), \quad x=0,1.$$

Furthermore, the true outcome variables τ_0 and τ_1 are uniquely defined (see SN-Th. 10.17), because, in this example, $C_X = U$ and $0 < P(X=1\,|\,U) < 1$. This implies that τ_0 and τ_1 are also P-unique. According to SN-Remark 14.34,

$$E(Y\,|\,X,U) \underset{P}{=} \sum_x E^{X=x}(Y\,|\,U)\cdot 1_{X=x}.$$

Hence, we conclude that in this example

$$E(Y\,|\,X,C_X) \underset{P}{=} E(Y\,|\,X,U) \underset{P}{=} E(Y\,|\,X,Z).$$

Because $C_X = U$ and $\sigma(Z) \subset \sigma(C_X)$, this is $Y \vdash C_X\,|\,(X,Z)$. Furthermore, because τ_0 and τ_1 are P-unique, according to Remark 9.27, $Y \vdash C_X\,|\,(X,Z)$ implies that $E(Y\,|\,X,Z)$ is unbiased.

Because, $P(X=x, Z=z) > 0$ for all values of X and Z, according to Corollary 9.30, $Y \vdash C_X\,|\,(X,Z)$ also implies that the conditional prima facie effects

$$PFE_{Z;10}(m) \;=\; E(Y\,|\,X=1, Z=m) - E(Y\,|\,X=0, Z=m) \;=\; 95-90 \;=\; 5$$

and

$$PFE_{Z;10}(f) \;=\; E(Y\,|\,X=1, Z=f) - E(Y\,|\,X=0, Z=f) \;=\; 110-100 \;=\; 10$$

are unbiased as well, that is, $PFE_{Z;10}(m) = CTE_{Z;10}(m)$ and $PFE_{Z;10}(f) = CTE_{Z;10}(f)$. In other words, the $(Z=z)$-conditional prima facie effects are equal to the causal $(Z=z)$-conditional total effects.

Another conclusion is that the expectation of the conditional total effects is equal to the causal average total effect [see Eq. (6.38)], that is,

$$\begin{aligned} ATE_{10} \;=\; E[PFE_{Z;10}(Z)] \;&=\; PFE_{Z;10}(m)\cdot 4/6 + PFE_{Z;10}(f)\cdot 2/6 \\ &=\; 5\cdot 4/6 + 10\cdot 2/6 \;\approx\; 6.667. \end{aligned}$$

Finally, note that, if $Y \vdash C_X\,|\,(X,Z)$ holds, then the individual treatment probabilities $P(X=1\,|\,U=u)$ that played a crucial role in our examples for independence and Z-conditional independence of X and C_X (see ch. 8) do not matter any more in the following sense: Even though, in this example, the individual treatment probabilities $P(X=1\,|\,U=u)$ are different within the male and within the female subpopulations, the $(Z=z)$-conditional prima facie effects are unbiased. ◁

9.5 Methodological Conclusions

Now we discuss the implications of the causility conditions $Y \vdash C_X|X$ and $Y \vdash C_X\,|\,(X,Z)$ for the design and analysis of experiments and quasi-experiments.

9.5.1 Methodological Conclusions From $Y \vdash C_X|X$

First of all, we have to realize that $Y \vdash C_X|X$ has no implications on any design technique. Either the empirical phenomenon we study (i. e., the random experiment that we consider in a specific empirical application) is such that $Y \vdash C_X|X$ holds, or it is such that it does not hold. In contrast to $C_X \perp\!\!\!\perp X$, which can be created by randomized assignment of the unit to one of the treatment conditions, there is nothing we can do that implies $Y \vdash C_X|X$.

In data analysis, however, we can test the hypothesis

$$E(Y \mid X, W) \underset{P}{=} E(Y \mid X). \tag{9.36}$$

And, if W is C_X-measurable, then this equation follows from $Y \vdash C_X \mid X$. Hence, if this equation is rejected, then we also reject the hypothesis that $Y \vdash C_X \mid X$ holds. For a test of the hypothesis that Equation (9.36) holds, standard procedures of regression analysis can be used.

9.5.2 Methodological Conclusions From $Y \vdash C_X \mid (X, Z)$

In contrast to $Y \vdash C_X \mid X$, we can use $Y \vdash C_X \mid (X, Z)$ for *covariate selection* in quasi-experiments aiming at the analysis of causal conditional and average total effects. By definition, there is no randomization in a quasi-experiment. However, even under initial randomization, systematic attrition of subjects may invalidate the Fisher condition $X \perp\!\!\!\perp C_X$ (see, e. g., Abraham & Russell, 2004; Fichman & Cummings, 2003; Graham & Donaldson, 1993; Shadish et al., 2002). In this case, we will say that randomization failed and that the initially randomized experiment turned into a quasi-experiment.

In quasi-experiments, selecting the random variables Z_i in the m-variate covariate $Z := (Z_1, \ldots, Z_m)$ for which we can hope that $Y \vdash C_X \mid (X, Z)$ holds, is a useful strategy in the analysis of causal conditional and average total treatment effects. Again, compared to the causality condition $X \perp\!\!\!\perp C_X \mid Z$, it might be easier to find covariates such that $Y \vdash C_X \mid (X, Z)$ holds. If, for example, X is a treatment variable, in many cases a pre-test of the outcome variable Y will already suffice. In contrast, there might be many covariates determining the treatment probabilities. For instance, the *severity of the disorder, knowing about the treatment*, and *availability of the treatment* are candidates for such covariates. Often only $Z :=$ *pre-test severity of the disorder* may be important as a covariate to satisfy $Y \vdash C_X \mid (X, Z)$ if the outcome variable Y is *post-test severity of the disorder*. However, there is no guarantee that the pre-test is sufficient as a covariate for $Y \vdash C_X \mid (X, Z)$ to hold. Whether or not it holds depends on the application considered.

Remark 9.32 (Falsifiability) As already mentioned before, in contrast to some of the other causality conditions treated in chapters 6 and 7, the Reichenbach-Suppes conditions can be tested in empirical applications, at least in the sense that some consequences of these assumptions can be checked. Falsifiability is important, because otherwise we would not have any criterion for covariate selection, that is, for deciding whether or not a specific covariate should be included in the m-variate covariate $Z := (Z_1, \ldots, Z_m)$ for which we hope that either $C_X \perp\!\!\!\perp X \mid Z$ or $Y \vdash C_X \mid (X, Z)$ holds.

It is easy to test $Y \vdash C_X \mid (X, Z)$. For example, if W is measurable with respect to C_X, then we can test the hypothesis

$$E(Y \mid X, Z, W) \underset{P}{=} E(Y \mid X, Z), \tag{9.1}$$

which follows from $Y \vdash C_X \mid (X, Z)$. If we reject this hypothesis, then we also reject the causality condition $Y \vdash C_X \mid (X, Z)$. Again, we can use the well-known techniques of regression analysis for such a test (see, e. g., Aiken & West, 1996; Allen, 1997; Cohen et al., 2003; Draper & Smith, 1998; Gelman & Hill, 2007; Györfi, Kohler, Krzyzak, & Walk, 2002; von Eye & Schuster, 1998; West & Aiken, 2005). ◁

Box 9.1 Reichenbach-Suppes conditions

Let the Assumptions 5.1 hold, let Z be a covariate of X, and assume that τ_x is P-unique for all values $x \in X(\Omega)$.

Conditions implying unbiasedness of $E(Y|X)$ and $E(Y|X,Z)$

$Y \perp\!\!\!\perp C_X | X$ — *X-conditional independence of Y from C_X.*

$Y \vdash C_X | X$ — *X-conditional mean-independence of Y from C_X.* It is defined by

$$E(Y|X,C_X) \underset{P}{=} E(Y|X).$$

It may or may not hold depending on the empirical phenomenon considered. If, for all values x of X, τ_x is P-unique, then $Y \vdash C_X | X$ implies that $E(Y|X)$ and all its values $E(Y|X{=}x)$ are unbiased and that all $E^{X=x}(Y|Z)$ as well as $E(Y|X,Z)$ are unbiased.

Conditions implying unbiasedness of $E(Y|X,Z)$

$Y \perp\!\!\!\perp C_X | (X,Z)$ — *(X,Z)-conditional independence of Y from C_X.*

$Y \vdash C_X | (X,Z)$ — *(X,Z)-conditional mean-independence of Y from C_X.* It is defined by

$$E(Y|X,C_X) \underset{P}{=} E(Y|X,Z).$$

We can try to select the m-variate covariate $Z = (Z_1,\ldots,Z_m)$ such that $Y \vdash C_X | (X,Z)$ holds. If, for all values x of X, the true outcome variable τ_x is P-unique, then $Y \vdash C_X | (X,Z)$ implies that $E(Y|X,Z)$ and all $E^{X=x}(Y|Z)$, $x = 1,\ldots,J$, are unbiased. If τ_x is $P^{Z=z}$-unique and z is a value of Z satisfying $P(X{=}x, Z{=}z) > 0$, then $E(Y|X{=}x, Z{=}z)$ is unbiased.

9.6 Summary and Conclusions

In chapter 6 we introduced unbiasedness of the conditional expectation values $E(Y|X{=}x)$ and $E(Y|X{=}x,Z{=}z)$, and unbiasedness of the conditional expectations $E(Y|X)$, $E^{X=x}(Y|Z)$ and $E(Y|X,Z)$. There we showed that unbiasedness of these terms is crucial for computing causal conditional and average total effects. In chapter 7, we treated the Rosenbaum-Rubin conditions that imply unbiasedness of those conditional expectation values and conditional expectations. However, the Rosenbaum-Rubin conditions as well as the unbiasedness conditions themselves cannot be tested empirically.

In chapter 8, we introduced the Fisher conditions, a first class of empirically testable causality conditions that focus on independence and Z-conditional independence of X and a global potential confounder C_X. There we showed, for example, that $X \perp\!\!\!\perp C_X | Z$ implies Rosenbaum and Rubin's *strong ignorability*, that is, it implies $\tau \perp\!\!\!\perp X | Z$, where $\tau = (\tau_0, \tau_1, \ldots, \tau_J)$ consists of the $J+1$ true outcome variables τ_x and Z is a covariate of X. The Fisher conditions share the focus on (conditional) independence of X and potential confounders of X.

In contrast, the causality conditions introduced in this chapter, share the focus on conditional independence or conditional mean-independence of the outcome variable Y and

Table 9.2. Implications among Reichenbach-Suppes conditions

	$\tau \perp\!\!\!\perp X$	$Y \vdash C_X \mid X$	$Y \perp\!\!\!\perp C_X \mid X$	$\tau \perp\!\!\!\perp X \mid Z$	$Y \vdash C_X \mid (X,Z)$	$Y \perp\!\!\!\perp C_X \mid (X,Z)$
$\tau \perp\!\!\!\perp X$	$\Leftrightarrow$					
$Y \vdash C_X \mid X$	$\Rightarrow$	$\Leftrightarrow$		$\Rightarrow$	$\Rightarrow$	
$Y \perp\!\!\!\perp C_X \mid X$	$\Rightarrow$	$\Rightarrow$	$\Leftrightarrow$	$\Rightarrow$	$\Rightarrow$	$\Rightarrow$
$\tau \perp\!\!\!\perp X \mid Z$				$\Leftrightarrow$		
$Y \vdash C_X \mid (X,Z)$				$\Rightarrow$	$\Leftrightarrow$	
$Y \perp\!\!\!\perp C_X \mid (X,Z)$				$\Rightarrow$	$\Rightarrow$	$\Leftrightarrow$

Note: An entry $\Rightarrow$ ($\Leftrightarrow$) means that the condition in the row implies (is equivalent to) the condition in the column, provided that the Assumptions 5.1 hold. Z is a covariate of X, and we assume that, for all values x of X, the true outcome variables τ_x are P-unique. The symbols $\vdash$ and $\perp\!\!\!\perp$ are explained in Box 7.1. The implications of $\tau \perp\!\!\!\perp X$ and of $\tau \perp\!\!\!\perp X \mid Z$ on other causality conditions are summarized in Table 7.5.

potential confounders of X. These conditions have been referred to as the *Reichenbach-Suppes conditions*. Box 9.1 displays their symbols and their definitions. We also studied the implication relations among the Reichenbach-Suppes conditions and showed that they imply the Rosenbaum-Rubin conditions. Table 9.2 displays the details. Combining these results with those already displayed in Tables 7.5 and 7.6 reveals the far reaching implications of the Reichenbach-Suppes conditions. Unlike unbiasedness and the Rosenbaum-Rubin conditions, and in line with the Fisher conditions, the Reichenbach-Suppes conditions are empirically testable. Therefore, their most important methodological implication is that they provide a viable alternative for covariate selection aiming at creating unbiasedness. In contrast to $C_X \perp\!\!\!\perp X \mid Z$, which can be used to select the covariate $Z = (Z_1, \ldots, Z_m)$ such that X and the global potential confounder C_X are Z-conditionally independent, we can use $Y \vdash C_X \mid (X,Z)$ to select Z such that Y is (X,Z)-conditionally mean-independent from C_X.

Last but not least it should be emphasized that the Reichenbach-Suppes conditions may also hold if X is a continuous random variable. In this case the theory of true outcome variables does not apply any more, because it rests on the assumption that the values x of X have a nonzero probability. This assumption does not hold if X is continuous. Nevertheless, we can meaningfully talk about causal dependencies if one of the four causality conditions listed in Box 9.1 holds. Hence, the Reichenbach-Suppes conditions are relevant beyond the true outcome theory of causal effects.

9.7 Proofs

Proof of Theorem 9.12

(i) This proposition immediately follows from SN-Remark 16.35.
(ii) This proposition immediately follows from SN-Box 16.3 (x).

Proof of Theorem 9.13

This proposition immediately follows from SN-Remark 16.35.

Proof of Theorem 9.14

We only have to prove (ii) because (i) is a special case of (ii) for Z being a constant. Note that our assumptions include that Z is C_X-measurable, implying that Equations (9.6) and (9.9) are equivalent to each other. Because we assume $P(X{=}x) > 0$ for all $x \in X(\Omega)$, according to SN-Remark 14.70, Equation (9.9) implies

$$E^{X=x}(Y|C_X) \underset{P^{X=x}}{=} E^{X=x}(Y|Z), \quad \forall\, x \in X(\Omega)\,.$$

Assuming P-uniqueness of $E^{X=x}(Y|C_X)$ immediately yields (9.18). Vice versa, if Equation (9.18) holds, then

$$\begin{aligned}
E(Y|X,C_X) &\underset{P}{=} \sum_{x\in X(\Omega)} E^{X=x}(Y|C_X)\cdot 1_{X=x} && \text{[SN-Rem. 14.34]}\\
&\underset{P}{=} \sum_{x\in X(\Omega)} E^{X=x}(Y|Z)\cdot 1_{X=x} && \text{[(9.18)]}\\
&\underset{P}{=} E(Y|X,Z). && \text{[SN-Rem. 14.34]}
\end{aligned}$$

Proof of Theorem 9.17

$$\begin{aligned}
Y \vdash C_X \,|\, X{=}x \;&\Leftrightarrow\; E^{X=x}(Y|C_X) \underset{P^{X=x}}{=} E^{X=x}(Y) && \text{[(9.21)]}\\
&\Rightarrow\; E^{X=x}(Y|C_X) \underset{P}{=} E^{X=x}(Y) && \text{[}\tau_x \text{ is } P\text{-unique]}\\
&\Rightarrow\; \tau_x \underset{P}{=} E^{X=x}(Y) && \text{[}\tau_x \underset{P}{=} E^{X=x}(Y|C_X)\text{]}\\
&\Rightarrow\; \tau_x \perp\!\!\!\perp X. && \text{[SN-Lem. 5.51]}
\end{aligned}$$

Proof of Theorem 9.21

(i).

$$\begin{aligned}
&Y \vdash C_X \,|\, X\\
\Leftrightarrow\;& E(Y|X,C_X) \underset{P}{=} E(Y|X) && \text{[(9.2)]}\\
\Leftrightarrow\;& \forall\, x \in X(\Omega)\colon\; E^{X=x}(Y|C_X) \underset{P}{=} E^{X=x}(Y) && \text{[Lem. 9.14 (i)]}
\end{aligned}$$

$$\begin{aligned}
&\Rightarrow \forall x \in X(\Omega)\colon\ \tau_x \underset{P}{=} E^{X=x}(Y) && [\tau_x = E^{X=x}(Y|C_X),\ \tau_x \text{ is } P\text{-unique}]\\
&\Rightarrow \tau \underset{P}{=} \big(E^{X=0}(Y), E^{X=1}(Y), \ldots, E^{X=J}(Y)\big) && [\tau = (\tau_0, \tau_1, \ldots, \tau_J)]\\
&\Rightarrow \tau \perp\!\!\!\perp X. && [\text{SN-Lem. 5.51}]
\end{aligned}$$

(ii). This proposition follows from $\tau \underset{P}{=} \big(E^{X=0}(Y), E^{X=1}(Y), \ldots, E^{X=J}(Y)\big)$ and SN-Box 16.3 (xii). (iii).

$$\begin{aligned}
E(Y|X,Z) &\underset{P}{=} E\big(E(Y|X,C_X)\,\big|\,X,Z\big) && [\sigma(Z) \subset \sigma(C_X), \text{SN-Box 10.2 (v)}]\\
&\underset{P}{=} E\big(E(Y|X)\,\big|\,X,Z\big) && [(9.25)]\\
&\underset{P}{=} E(Y|X). && [\text{SN-Box 10.2 (vii)}]\\
&\underset{P}{=} E(Y|X,C_X). && [(9.25)]
\end{aligned}$$

According to (9.7) this proves the proposition.

Proof of Theorem 9.23

$$\begin{aligned}
Y \vdash C_X | X{=}x, Z \ &\Leftrightarrow\ E^{X=x}(Y|C_X) \underset{P^{X=x}}{=} E^{X=x}(Y|Z) && [(9.21)]\\
&\Rightarrow\ \tau_x \underset{P}{=} E^{X=x}(Y|Z). && [\tau_x \underset{P}{=} E^{X=x}(Y|C_X),\ \tau_x \text{ is } P\text{-unique}]
\end{aligned}$$

Furthermore,

$$\begin{aligned}
&E^{X=x}(Y|Z) \perp\!\!\!\perp X \,|\, Z && [\sigma\big(E^{X=x}(Y|Z)\big) \subset \sigma(Z), \text{SN-Box 16.3 (iv)}]\\
\Rightarrow\ &\tau_x \perp\!\!\!\perp X \,|\, Z. && [\tau_x \underset{P}{=} E^{X=x}(Y|Z), \text{SN-Box 16.3 (xi)}]
\end{aligned}$$

Proof of Theorem 9.25

$Y \vdash C_X | X{=}x, Z$ is defined by Proposition (9.28) and, because we assume that Z is C_X-measurable, it implies $E^{X=x}(Y|Z) \in \mathscr{E}^{X=x}(Y|C_X)$ (see SN-Rem. 10.14). Because $P(X{=}x, Z{=}z) > 0$ implies $P(Z{=}z) > 0$, and because we assume that τ_x is $P^{Z=z}$-unique, according SN-Remark 10.13,

$$\tau_x \underset{P^{Z=z}}{=} E^{X=x}(Y|Z), \quad \forall\, \tau_x \in \mathscr{E}^{X=x}(Y|C_X). \tag{9.2}$$

Furthermore,

$$\begin{aligned}
E(\tau_x | Z{=}z) &= E\big(E^{X=x}(Y|Z)\,\big|\,Z{=}z\big) && [P(Z{=}z) > 0, (9.2)]\\
&= E^{X=x}(Y|Z{=}z) && [\text{SN-(14.22)}]\\
&= E(Y|X{=}x, Z{=}z). && [\text{SN-(14.37)}]
\end{aligned}$$

Proof of Theorem 9.26

$$
\begin{aligned}
& Y \vdash C_X \,|\, (X,Z) \\
\Leftrightarrow\ & E(Y|X,C_X) \underset{P}{=} E(Y|X,Z) && [(9.7)] \\
\Leftrightarrow\ & \forall\, x \in X(\Omega)\colon\ E^{X=x}(Y|C_X) \underset{P}{=} E^{X=x}(Y\,|\,Z) && [\text{Th. 9.14 (ii)}] \\
\Rightarrow\ & \forall\, x \in X(\Omega)\colon\ \tau_x \underset{P}{=} E^{X=x}(Y\,|\,Z) && [\tau_x \underset{P}{=} E^{X=x}(Y|C_X),\ \tau_x \text{ is } P\text{-unique}] \\
\Rightarrow\ & \tau \underset{P}{=} \left(E^{X=0}(Y\,|\,Z), E^{X=1}(Y\,|\,Z), \ldots, E^{X=J}(Y\,|\,Z)\right) && [\tau = (\tau_0, \tau_1, \ldots, \tau_J)] \\
\Rightarrow\ & \tau \perp\!\!\!\perp X \,|\, Z. && [\sigma(\tau) \subset \sigma(Z),\ \text{SN-Box 16.3 (iv), (xi)}]
\end{aligned}
$$

Note that SN-Box 16.3 (xi) implies $\tau \perp\!\!\!\perp X \,|\, Z$ for all versions of τ, even those that are not Z-measurable.

Proof of Theorem 9.28

$Y \vdash C_X \,|\, (X,Z)$ is defined by Proposition (9.7) and, because $\sigma(Z) \subset \sigma(C_X)$, according to SN-Proposition (14.81), it implies

$$\forall\, x \in X(\Omega)\colon\ E^{X=x}(Y|C_X) \underset{P^{X=x}}{=} E^{X=x}(Y|Z)\,. \tag{9.3}$$

Using the shortcut introduced in Proposition (9.28), this can equivalently be written

$$\forall\, x \in X(\Omega)\colon\ Y \vdash C_X \,|\, X{=}x, Z. \tag{9.4}$$

Now Theorem 9.25 implies the proposition.

Proof of Corollary 9.30

(i). This proposition follows from Theorem 9.26, Theorem 9.23, Equations (5.2), (5.16), and (6.29).

(ii) If $P(X{=}x, Z{=}z) > 0$, $P(X{=}x', Z{=}z) > 0$ and τ_x, $\tau_{x'}$ are $P^{Z=z}$-unique, then, according to Theorem 9.28, the conditional expectation values $E(Y\,|\,X{=}x, Z{=}z)$ and $E(Y\,|\,X{=}x', Z{=}z)$ are unbiased. Equations (5.2), (5.14), and (6.30) then yield the proposition.

9.8 Exercises

▷ **Exercise 9-1** Show that $Y \vdash C_X \,|\, X{=}x$ implies $E^{X=x}(Y\,|\,W) \underset{P^{X=x}}{=} E^{X=x}(Y)$, if we assume that W is a C_X-measurable random variable.

▷ **Exercise 9-2** Show that $Y \vdash C_X \,|\, (X,Z)$ implies $E(Y\,|\,X,Z,W) \underset{P}{=} E(Y\,|\,X,Z)$ if both Z and W are measurable with respect to the global potential confounder C_X.

▷ **Exercise 9-3** Let the Assumptions 6.1 hold, let Z be a covariate of X, and assume that τ_x is P-unique for all $x = 0, 1, \ldots, J$. Which unbiasedness conditions follow from $Y \vdash C_X \,|\, X$ and which ones from $Y \vdash C_X \,|\, (X,Z)$?

▷ **Exercise 9-4** Check that all implications listed in Table 9.2 have been proven in this chapter. Use the Assumptions 6.1, that Z is a covariate of X, and the additional assumption that all true outcome variables τ_x are P-unique.

Solutions

▷ **Solution 9-1** By definition, $Y \vdash C_X \,|\, X{=}x$ is equivalent to $E^{X=x}(Y|C_X) \underset{P^{X=x}}{=} E^{X=x}(Y)$ [see (9.21)]. Hence,

$$\begin{aligned} E^{X=x}(Y|W) &\underset{P^{X=x}}{=} E^{X=x}\big(E^{X=x}(Y|C_X)\,|\,W\big) && [\sigma(W)\subset\sigma(C_X),\ \text{SN-Box 10.2 (v)}] \\ &\underset{P^{X=x}}{=} E^{X=x}\big(E^{X=x}(Y)\,|\,W\big) && [(9.22)] \\ &\underset{P^{X=x}}{=} E^{X=x}(Y). && [\text{SN-Box 10.2 (i)}] \end{aligned}$$

▷ **Solution 9-2**

$$\begin{aligned} E(Y|X,Z,W) &\underset{P}{=} E\big(E(Y|X,C_X)\,\big|\,X,Z,W\big) && [\sigma(Z)\subset\sigma(C_X),\ \text{SN-Box 10.2 (v)}] \\ &\underset{P}{=} E\big(E(Y|X,Z)\,\big|\,X,Z,W\big) && [Y\vdash C_X\,|\,(X,Z)] \\ &\underset{P}{=} E(Y|X,Z)\,. && [\text{Box SN-10.2 (vii)}] \end{aligned}$$

▷ **Solution 9-3** Under the assumptions mentioned in the exercise, $Y \vdash C_X|X$ implies that $E(Y|X)$, all its values $E(Y|X{=}x)$, and the prima facie effects $PFE_{xx'}$ are unbiased. Under the same assumptions $Y \vdash C_X\,|\,(X,Z)$ implies that $E(Y|X,Z)$ and all conditional expectations $E^{X=x}(Y|Z)$, $x = 0,1,\ldots,J$, are unbiased. Furthermore, if τ_x is P-unique and z is a value of Z satisfying $P(X{=}x, Z{=}z) > 0$, then $Y \vdash C_X\,|\,(X,Z)$ also implies that $E(Y|X{=}x, Z{=}z)$ is unbiased.

▷ **Solution 9-4** We check the implications listed in Table 9.2 row-wise.

(a) $Y \vdash C_X\,|\,X \;\Rightarrow\; \tau \perp\!\!\!\perp X$. This is Proposition (i) of Theorem 9.21.

(b) $Y \vdash C_X\,|\,X \;\Rightarrow\; \tau \perp\!\!\!\perp X\,|\,Z$. This is Proposition (ii) of Theorem 9.21.

(c) $Y \vdash C_X\,|\,X \;\Rightarrow\; Y \vdash C_X\,|\,(X,Z)$. This is Proposition (iii) of Theorem 9.21.

(d) $Y \perp\!\!\!\perp C_X\,|\,X \;\Rightarrow\; \tau \perp\!\!\!\perp X$. This proposition immediately follows from (e) and (a).

(e) $Y \perp\!\!\!\perp C_X\,|\,X \;\Rightarrow\; Y \vdash C_X\,|\,X$. This proposition follows from Theorem 9.12 (i) and Proposition (9.2).

(f) $Y \perp\!\!\!\perp C_X\,|\,X \;\Rightarrow\; \tau \perp\!\!\!\perp X\,|\,Z$. This proposition immediately follows from (e) and (b).

(g) $Y \perp\!\!\!\perp C_X\,|\,X \;\Rightarrow\; Y \vdash C_X\,|\,(X,Z)$. This proposition immediately follows from (e) and (c).

(h) $Y \perp\!\!\!\perp C_X\,|\,X \;\Rightarrow\; Y \perp\!\!\!\perp C_X\,|\,(X,Z)$. This is the proposition of Theorem 9.12 (ii).

(i) $Y \vdash C_X\,|\,(X,Z) \;\Rightarrow\; \tau \perp\!\!\!\perp X\,|\,Z$. This is the proposition of Theorem 9.26.

(j) $Y \perp\!\!\!\perp C_X\,|\,(X,Z) \;\Rightarrow\; \tau \perp\!\!\!\perp X\,|\,Z$. This proposition follows from (k) and (i).

(k) $Y \perp\!\!\!\perp C_X\,|\,(X,Z) \;\Rightarrow\; Y \vdash C_X\,|\,(X,Z)$. This proposition follows from Theorem 9.13 and Proposition (9.7)

References

Abraham, W. T., & Russell, D. W. (2004). Missing data: A review of current methods and applications in epidemiological research. *Current Opinion in Psychiatry*, *17*, 315–321.

Agresti, A. (2007). *An introduction to categorical data analysis* (2nd ed.). Hoboken, NJ: Wiley.

Aiken, L. S., & West, S. G. (1996). *Multiple regression: Testing and interpreting interactions.* Thousand Oaks, CA: Sage.

Aitkin, M. (1978). The analysis of unbalanced cross-classifications. *Journal of the Royal Statistical Society, Series A: Statistics in Society*, *141*, 195–223.

Allen, M. P. (1997). *Understanding regression analysis.* New York, NY: Plenum Press.

Angrist, J. D., Imbens, G. W., & Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*, *91*, 444–455.

Appelbaum, M. I., & Cramer, E. M. (1974). Some problems in the nonorthogonal analysis of variance. *Psychological Bulletin*, *81*, 335–343.

Arbuckle, J. L. (2006). AMOS 7.0 user's guide [Computer software manual]. Chicago, IL: SPSS.

Barnard, J., Du, J., Hill, J. L., & Rubin, D. B. (1998). A broader template for analyzing broken randomized experiments. *Sociological Methods & Research*, *27*, 285–317.

Barnard, J., Frangakis, C. E., Hill, J. L., & Rubin, D. B. (2003, June). Principal stratification approach to broken randomized experiments: A case study of school choice vouchers in new york city. *Journal of the American Statistical Association*, *98*(462), 299–323.

Bentler, P. M. (1995). EQS Structural Equations program manual [Computer software manual]. Encino, CA: Multivariate Software.

Berger, M. P. F., & Wong, W. K. (Eds.). (2005). *Applied optimal designs.* Chichester, England: Wiley.

Biesanz, J. C., Deeb-Sossa, N., Aubrecht, A. M., Bollen, K. A., & Curran, P. J. (2004). The role of coding time in estimating and interpreting growth curve models. *Psychological Methods*, *9*, 30–52.

Bollen, K. A. (1995). Structural equation models that are nonlinear in latent variables: A least-squares estimator. *Sociological Methodology*, *25*, 223–251.

Bollen, K. A., & Curran, P. J. (2006). *Latent curve models: A structural equation approach.* Hoboken, NJ: Wiley.

Bonney, G. E. (1987). Logistic regression for dependent binary observations. *Biometrics*, *43*, 951–973.

Borooah, V. K. (2001). *Logit and probit: Ordered and multinomial models.* Thousand Oaks, CA: Sage.

Browne, M. W., & Mels, G. (1998). Path analysis: RAMONA. In *SYSTAT for Windows: Advanced Applications (Version 8) [Computer software manual]*. Evanston, IL: SYSTAT.

Campbell, D. T., & Stanley, J. C. (1963). Experimental and quasi-experimental designs for research on teaching. In N. L. Gage (Ed.), *Handbook of research on teaching*. Chicago, IL: Rand McNally.

Campbell, D. T., & Stanley, J. C. (1966). *Experimental and quasi-experimental designs for research*. Boston, MA: Houghton Mifflin.

Carlson, J. E., & Timm, N. H. (1974). Analysis of nonorthogonal fixed-effects designs. *Psychological Bulletin*, *81*, 563–570.

Cartwright, N. (1979). Causal laws and effective strategies. *Noûs*, *13*, 419–437.

Cheng, J., & Small, D. S. (2006). Bounds on causal effects in three-arm trials with noncompliance. *Journal of the Royal Statistical Society, Series B: Statistical Methodology*, *68*, 815–836.

Cochran, W. G. (1957). Analysis of covariance: Its nature and uses. *Biometrics*, *13*, 261–281.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression / correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum.

Cook, E. F., & Goldman, L. (1989). Performance of tests of significance based on stratification by a multivariate confounder score or by a propensity score. *Journal of Clinical Epidemiology*, *42*, 317–324.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Chicago, IL: Rand McNally.

Cox, D. R., & Wermuth, N. (2004). Causality: A statistical view. *International Statistical Review*, *72*, 285–305.

Dawid, A. P. (1979). Conditional independence in statistical theory. *Journal of the Royal Statistical Society, Series B: Statistical Methodology*, *41*(1), 1–31.

Draper, N. R., & Smith, H. (1998). *Applied regression analysis*. New York, NY: Wiley.

Dunn, G., Maracy, M., Dowrick, C., Ayuso-Mateos, J. L., Dalgard, O. S., Page, H., ... Wilkinson, G. (2003). Estimating psychological treatment effects from a randomised controlled trial with both non-compliance and loss to follow-up. *British Journal of Psychiatry*, *183*, 323–331.

Fahrmeir, L., Hamerle, A., & Tutz, G. (1996). *Multivariate statistische Verfahren. [Multivariate statistical methods]* (2nd ed.). Berlin, Germany: de Gruyter.

Fahrmeir, L., & Tutz, G. (2001). *Multivariate statistical modelling based on Generalized Linear Models* (2nd ed.). New York, NY: Springer.

Fichman, M., & Cummings, J. N. (2003). Multiple imputation for missing data: Making the most of what you know. *Organizational Research Methods*, *6*, 282–308.

Fisher, R. A. (1925/1946). *Statistical methods for research workers* (10th ed.). Edinburgh, England: Oliver and Boyd.

Gelman, A., & Hill, J. L. (2007). *Data analysis using regression and multilevel / hierarchical models*. New York, NY: Cambridge University.

Gosslee, D. G., & Lucas, H. L. (1965). Analysis of variance of disproportionate data when interaction is present. *Biometrics*, *21*, 115–133.

Graham, J. W., & Donaldson, S. I. (1993). Evaluating interventions with differential attrition: The importance of nonresponse mechanisms and use of follow-up data. *Journal of Applied Psychology*, *78*, 119–128.

Green, W. H. (2003). *Econometric analysis* (5th ed.). Upper Saddle River, NJ: Prentice-Hall.

Greenland, S. (1996). Basic methods for sensitivity analysis of biases. *International Journal of Epidemiology*, *25*, 1107–1116.

Greenland, S. (2000). Causal analysis in the health sciences. *Journal of the American Statistical Association*, *95*, 286–289.

Greenland, S. (2001). Sensitivity analysis, Monte Carlo risk analysis and Bayesian uncertainty assessment. *Risk Analysis*, *21*, 579–583.

Greenland, S. (2004). An overview of methods for causal inference from observational studies. In A. Gelman & X.-L. Meng (Eds.), *Applied bayesian modeling and causal inference from incomplete-data perspectives* (pp. 3–14). Chichester, England: Wiley.

Greenland, S., & Robins, J. M. (1986). Identifiability, exchangeability, and epidemiological confounding. *International Journal of Epidemiology*, *15*, 413–419.

Györfi, L., Kohler, M., Krzyzak, A., & Walk, H. (2002). *A distribution-free theory of nonparametric regression*. New York, NY: Springer.

Hancock, G. R. (2004). Experimental, quasi-experimental, and nonexperimental design and analysis with latent variables. In D. Kaplan (Ed.), *The Sage handbook of quantitative methodology for the social sciences*. Thousand Oaks, CA: Sage.

Hernán, M. A., Clayton, D., & Keiding, N. (2011). The Simpson's paradox unraveled. *International Journal of Epidemiology*, 1–6. (Advance Access published March 30, 2011) doi: 10.1093/ije/dyr041

Höfler, M. (2005). Causal inference based on counterfactuals. *BMC Medical Research Methodology*, *5*, 1–12.

Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association*, *81*, 945–960.

Horvitz, D. G., & Thompson, D. J. (1952). A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association*, *47*, 663–685.

Hosmer, D. W., & Lemeshow, S. (2000). *Applied logistic regression* (2nd ed.). New York, NY: Wiley.

Hoyer, J., & Klein, A. (2000). Self-reflection and well-being: Is there a healthy amount of introspection? *Psychological Reports*, *86*, 135–141.

Huet, S., Bouvier, A., Poursat, M.-A., & Jolivet, E. (2004). *Statistical tools for nonlinear regression: A practical guide with S-PLUS and R examples* (2nd ed.). New York, NY: Springer.

Jamieson, J. (2004). Analysis of covariance ANCOVA with difference scores. *International Journal of Psychophysiology*, *52*, 277–283.

Jennings, E., & Green, J. L. (1984). Resolving nonorthogonal ANOVA disputes using cell means. *Journal of Experimental Education*, *52*, 159–162.

Jo, B. (2002a). Estimation of intervention effects with noncompliance: Alternative model specifications. *Journal of Educational and Behavioral Statistics*, *27*, 385–409.

Jo, B. (2002b). Model misspecification sensitivity analysis in estimating causal effects of interventions with non-compliance. *Statistics in Medicine*, *21*, 3161–3181.

Jo, B. (2002c). Statistical power in randomized intervention studies with noncompliance. *Psychological Methods*, *7*, 178–193.

Jo, B., Asparouhov, T., Muthén, B. O., Ialongo, N. S., & Brown, C. H. (2008). Cluster randomized trials with treatment noncomplicance. *Psychological Methods*, *13*, 1–18.

Jöreskog, K. G., & Sörbom, D. (1996/2001). *LISREL 8: User's Reference Guide. [Computer software manual]* (2nd ed.). Lincolnwood, IL: Scientific Software International.

Keele, L. (2008). *Semiparametric regression for the social sciences*. Chichester, England: Wiley.

Kenny, D. A. (1975). Cross-lagged panel correlation: A test for spuriousness. *Psychology Bulletin*, *82*, 887–903.

Kenny, D. A., & Judd, C. M. (1984). Estimating the nonlinear and interactive effects of latent variables. *Psychological Bulletin*, *96*, 201–210.

Keren, G., & Lewis, C. (1976). Nonorthogonal designs: Sample versus population. *Psychological Bulletin*, *83*, 817–826.

King, G., & Zeng, L. (2001). Improving forecasts of state failure. *World Politics*, *53*, 623–658.

Klauer, K. J., Willmes, K., & Phye, G. D. (2002). Inducing inductive reasoning: Does it transfer to fluid intelligence? *Contemporary Educational Psychology*, *27*, 1–25.

Klein, A. G., & Moosbrugger, H. (2000). Maximum likelihood estimation of latent interaction effects with the LMS method. *Psychometrika*, *65*, 457–474.

Klein, A. G., & Muthén, B. O. (2007). Quasi-maximum likelihood estimation of structural equation models with multiple interaction and quadratic effects. *Multivariate Behavioral Research*, *42*, 647–673.

Klenke, A. (2013). *Probability theory – A comprehensive course* (2nd ed.). London, England: Springer. doi: 10.1007/978-1-4471-5361-0

Kolmogorov, A. N. (1933/1977). *Grundbegriffe der Wahrscheinlichkeitsrechnung [Foundations of the Theory of Probability]* (Reprinted ed.). Berlin, Germany: Springer.

Kolmogorov, A. N. (1956). *Foundations of the theory of probability* (2nd ed.; N. Morrison, Trans.). New York, NY: Chelsea.

Kramer, C. Y. (1955). On the analysis of variance of a two-way classification with unequal sub-class numbers. *Biometrics*, *11*, 441–452.

Langsrud, Ø. (2003). ANOVA for unbalanced data: Use Type II instead of Type III sums of squares. *Statistics and Computing*, *13*, 163–167.

Lee, B. K., Lessler, J., & Stuart, E. A. (2010). Improving propensity score weighting using machine learning. *Statistics in Medicine*, *29*, 337–346. doi: 10.1002/sim.3782

Levy, P. S., & Lemeshow, S. (2003). *Sampling of populations: Methods and applications* (3rd ed.). New York, NY: Wiley.

Liao, T. F. (1994). *Interpreting probability models: Logit, probit, and other generalized linear models*. London, England: Sage.

Lord, F. M. (1967). A paradox in the interpretation of group comparisons. *Psychological Bulletin*, *68*, 304–305.

Marsh, H. W., Wen, Z., & Hau, K. T. (2004). Structural equation models of latent interactions: Evaluation of alternative estimation strategies and indicator construction. *Psychological Methods*, *9*, 275–300.

Maxwell, S. E., & Delaney, H. D. (2004). *Designing experiments and analyzing data: A model comparison perspective* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum.

Mayer, A., Dietzfelbinger, L., Rosseel, Y., & Steyer, R. (2016). The effectliter approach for analyzing average and conditional effects. *Multivariate Behavioral Research*, *51*, 374–391. doi: 10.1080/00273171.2016.1151334

Mayer, A., Thoemmes, F., Rose, N., Steyer, R., & West, S. G. (2014). Theory and analysis of total, direct, and indirect causal effects. *Multivariate Behavioral Research*, *49*(5), 425–442. doi: 10.1080/00273171.2014.931797

McArdle, J. J. (2001). A latent difference score approach to longitudinal dynamic structural analysis. In R. Cudeck, S. du Toit, & D. Sörbom (Eds.), *Structural equation modeling:*

Present and future (pp. 341–380). Lincolnwood, IL: Scientific Software International.

McCaffrey, D. F., Ridgeway, G., & Morral, A. R. (2004). Propensity score estimation with boosted regression for evaluating causal effects in observational studies. *Psychological Methods, 9*(4), 403–425.

McCullagh, P., & Nelder, J. A. (1989). *Monographs on statistics and applied probability: Vol. 37. Generalized linear models* (2nd ed.; D. R. Cox, D. V. Hinkley, N. Reid, D. B. Rubin, & D. V. Silverman, Eds.). Chapman & Hall.

Meredith, M., & Tisak, J. (1990). Latent curve analysis. *Psychometrika, 55*, 107–122.

Mill, J. S. (1843/1865). Of the four methods of experimental inquiry. In *A system of logic, ratiocinative and inductive: Volume 1. Being a connected view of the principles of evidence, and the methods of scientific investigation.* London, England: Longmans, Green, and Co.

Morgan, S. L., & Winship, C. (2007). *Counterfactuals and causal inference. methods and principles for social research.* New York, NY: Cambridge University.

Muthén, L. K., & Muthén, B. O. (1998-2007). Mplus User's Guide (5th ed.) [Computer software manual]. Los Angeles, CA: Muthén & Muthén.

Muthén, L. K., & Muthén, B. O. (2002). How to use a Monte Carlo study to decide on sample size and determine power. *Structural Equation Modeling, 9*, 599–620.

Nagengast, B. (2009). *Causal inference in multilevel models* (Unpublished doctoral dissertation). Friedrich-Schiller-Universität Jena, Thüringen, Germany.

Nagengast, B., Kröhne, U., Bauer, M., & Steyer, R. (2007). Causal Effects Explorer: A didactic tool for teaching the theory of individual and average causal effects [Computer software manual]. University of Jena, Thüringen, Germany.

Nelder, J. A., & Lane, P. W. (1995). The computer analysis of factorial experiments: In memoriam – Frank Yates. *The American Statistician, 49*, 382–385.

OpenMx. (2009). *OpenMx - Advanced Structural Equation Modeling [Computer Software].* Retrieved from http://openmx.psyc.virginia.edu/.

Overall, J. E., & Spiegel, D. K. (1969). Concerning least squares analysis of experimental data. *Psychological Bulletin, 72*(5), 311–322.

Overall, J. E., & Spiegel, D. K. (1973a). Comment on "Regression analysis of proportional cell data". *Psychological Bulletin, 80*, 28–30.

Overall, J. E., & Spiegel, D. K. (1973b). Comments on Rawlings' nonorthogonal analysis of variance. *Psychological Bulletin, 79*, 164–167.

Overall, J. E., Spiegel, D. K., & Cohen, J. (1975). Equivalence of orthogonal and nonorthogonal analysis of variance. *Psychological Bulletin, 82*, 182–186.

Pearl, J. (2009). *Causality: Models, reasoning, and inference* (2nd ed.). Cambridge, UK.

Pearson, K., Lee, A., & Bramley-Moore, L. (1899). Genetic (reproductive) selection: Inheritance of fertility in man, and of fecundity in thoroughbred racehorses. *Series A, Containing Papers of a Mathematical or Physical Character: Philosophical Transactions of the Royal Society of London, 192*, 257–330.

Porta, M. (2014). *A dictionary of epidemiology.* Oxford University Press. doi: 10.1093/acref/9780199976720.001.0001

Pukelsheim, F. (2006). *Optimal design of experiments.* Philadelphia, PA: SIAM (Society for Industrial and Applied Mathematics).

Rao, C. R. (1973). *Linear statistical inference and its applications* (2nd ed.). New York, NY: Wiley.

Reichardt, C. S. (1979). The statistical analysis of data from nonequivalent group designs. In T. D. Cook & D. C. Campbell (Eds.), *Quasi-experimentation: Design and analysis issues for field settings.* Orlando, FL: Houghton Mifflin.

Reichardt, C. S. (2005). Nonequivalent group design. In B. Everitt & D. Howell (Eds.), *Encyclopedia of behavioral statistics.* New York, NY: Wiley.

Reichenbach, H. (1956). *The direction of time* (M. Reichenbach, Ed.). University of California Press.

Robins, J., & Rotnitzky, A. (2004). Estimation of treatment effects in randomised trials with non-compliance and a dichotomous outcome using structural mean models. *Biometrika*, *91*, 763–783.

Robins, J. M. (1998). Correction for non-compliance in equivalence trials. *Statistics in Medicine*, *17*, 269–302.

Robins, J. M., Hérnan, M. A., & Brumback, B. (2000). Marginal structural models and causal inference in epidemiology. *Epidemiology*, *11*, 550–560.

Rogosa, D. (1980a). Comparing nonparallel regression lines. *Psychological Bulletin*, *88*, 307–321.

Rogosa, D. (1980b). A critique of cross-lagged correlation. *Psychological Bulletin*, *88*, 245–258.

Rosenbaum, P. R. (2002a). Attributing effects to treatment in matched observational studies. *Journal of the American Statistical Association*, *97*(457), 183–192. doi: 10.1198/016214502753479329

Rosenbaum, P. R. (2002b). *Observational studies* (2nd ed.). New York, NY: Springer.

Rosenbaum, P. R., & Rubin, D. B. (1983a). Assessing sensitivity to an unobserved binary covariate in an observational study with binary outcome. *Journal of the Royal Statistical Society, Series B: Statistical Methodology*, *45*, 212–218.

Rosenbaum, P. R., & Rubin, D. B. (1983b). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*, 41–55. doi: 10.1093/biomet/70.1.41

Rosenbaum, P. R., & Rubin, D. B. (1984). Reducing bias in observational studies using subclassification on the propensity score. *Journal of the American Statistical Association*, *79*, 516–524.

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1–36.

Rothman, K. J., Greenland, S., & Lash, T. L. (2008). *Modern epidemiology* (3rd ed.). Philadelphia, PA: Lippincott Williams & Wilkins.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, *66*, 688–701.

Rubin, D. B. (2004). On principles for modeling propensity scores in medical research. *Pharmacoepidemiology and Drug Safety*, *13*, 855–857. Retrieved from `ISI:000225802100003`

Rubin, D. B. (2005). Causal inference using potential outcomes: Design, modeling, decisions. *Journal of the American Statistical Association*, *100*, 322–331. doi: 10.1198/016214504000001880

Rubin, D. B. (2006). *Matched sampling for causal effects.* New York, NY: Cambridge University.

Rubin, D. B., & Thomas, N. (1996). Matching using estimated propensity scores: Relating theory to practice. *Biometrics*, *52*, 249–264.

Särndal, C.-E., Swensson, B., & Wretman, J. H. (2003). *Model assisted survey sampling* (2nd ed.). Berlin, Germany: Springer.

Searle, S. R. (1971). *Linear models.* New York, NY: Wiley.

Searle, S. R., Casella, G., & McCulloch, C. E. (1992). *Variance components.* New York, NY: Wiley.

Senn, S. (2006). Change from baseline and analysis of covariance revisited. *Statistics in Medicine*, *25*, 4334–4344.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference.* Boston, MA: Houghton Mifflin.

Simpson, E. H. (1951). The interpretation of interaction in contingency tables. *Journal of the Royal Statistical Society, Series B: Statistical Methodology*, *13*, 238–241.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence.* New York, NY: Oxford University.

Song, X. Y., & Lee, S. Y. (2006). Bayesian analysis of structural equation models with nonlinear covariates and latent variables. *Multivariate Behavioral Research*, *41*, 337–365.

Sörbom, D. (1978). An alternative to the methodology for analysis of covariance. *Psychometrika*, *43*, 381–396.

Spirtes, P., Glymour, C., & Scheines, R. (2000). *Causation, prediction, and search* (2nd ed.). Cambridge, MA: MIT.

Splawa-Neyman, J. (1923/1990). On the application of probability theory to agricultural experiments. Essay on principles. Section 9 (Reprinted from roczniki nauk rolniczych tom 10, pp. 1–51, 1923). *Statistical Science*, *5*, 465–480.

Spohn, W. (1980). Stochastic independence, causal independence, and shieldability. *Journal of Philosophical Logic*, *9*, 73–99.

Stegmüller, W. (1983). *Erklärung, Begründung, Kausalität: Probleme und Resultate der Wissenschaftstheorie und analytischen Philosophie. [Explanation, justification, causality: Problems and findings of philosophy of science and analytic philosophy].* Berlin, Germany: Springer.

Steyer, R. (1984). Causal linear stochastic dependencies: An introduction. In J. R. Nesselroade & A. von Eye (Eds.), *Individual development and social change: Explanatory analysis.* New York, NY: Academic Press.

Steyer, R. (1992). *Theorie kausaler Regressionsmodelle [Theory of causal regression models].* Stuttgart, Germany: Fischer.

Steyer, R. (2001). Classical test theory. In C. Ragin & T. D. Cook (Eds.), *International encyclopedia of the social and behavioural sciences: Logic of inquiry and research design* (pp. 481–520). Oxford, England: Pergamon.

Steyer, R. (2005). Analyzing individual and average causal effects via structural equation models. *Methodology*, *1*, 39–54.

Steyer, R., & Eid, M. (2001). *Messen und Testen: Ein Lehrbuch [Measurement and testing].* Berlin, Germany: Springer.

Steyer, R., Eid, M., & Schwenkmezger, P. (1997). Modeling true intraindividual change: True change as a latent variable. *Methods of Psychological Research Online*, *2*, 21–33.

Steyer, R., Gabler, S., von Davier, A. A., & Nachtigall, C. (2000). Causal regression models II: Unconfoundedness and causal unbiasedness. *Methods of Psychological Research Online*, *5*, 55–87.

Steyer, R., Mayer, A., Geiser, C., & Cole, D. (2015). A theory of states and traits - revised. *Annual Review of Clinical Psychology*, *11*, 71–98.

Steyer, R., & Nagel, W. (2017). *Probability and conditional expectation. Fundamentals for the Empirical Sciences.* Chichester, England: Wiley.

Steyer, R., & Partchev, I. (2007). EffectLite: User's manual – A program for the uni- and multivariate analysis of unconditional, conditional and average mean differences between groups [Computer software manual]. University of Jena, Thüringen, Germany.

Suppes, P. (1970). *A probabilistic theory of causality.* Amsterdam, Netherlands: North-Holland.

Takezawa, K. (2005). *Introduction to nonparametric regression.* New York, NY: Wiley.

Tisak, J., & Tisak, M. S. (2000). Permanency and ephemerality of psychological measures with application to organizational commitment. *Psychological Methods*, *5*, 175–198.

van Breukelen, G. J. P. (2006). ANCOVA versus change from baseline had more power in randomized studies and more bias in nonrandomized studies. *Journal of Clinical Epidemiology*, *59*, 920–925. doi: 10.1016/j.jclinepi.2006.02.007

von Eye, A., & Schuster, C. (1998). *Regression analysis for social sciences.* San Diego, CA: Academic Press.

Wainer, H. (1991). Adjusting for differential base rates: Lord's paradox again. *Psychological Bulletin*, *109*, 147–151.

Wall, M. M., & Amemiya, Y. (2003). A method of moments technique for fitting interaction effects in structural equation models. *British Journal of Mathematical and Statistical Psychology*, *56*, 47–63.

Watkins, M. W., Lei, P., & Canivez, G. L. (2007). Psychometric intelligence and achievement: A cross-lagged panel analysis. *Intelligence*, *35*, 59–68.

West, S. G., & Aiken, L. S. (2005). Multiple linear regression. In B. S. Everitt & D. C. Howell (Eds.), *Encyclopedia of statistics in behavioral science.* Chichester, England: Wiley.

West, S. G., Biesanz, J. C., & Pitts, S. C. (2000). Causal inference and generalization in field settings experimental and quasi-experimental designs. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (pp. 40–84). New York, NY: Cambridge University.

West, S. G., & Sagarin, B. J. (2000). Participant selection and loss in randomized experiments. In L. Bickman (Ed.), *Research design: Donald Campbell's legacy* (pp. 117–151). Thousand Oaks, CA: Sage.

Williams, J. D. (1972). Two way fixed effects analysis of variance with disproportionate cell frequencies. *Multivariate Behavioral Research*, *7*, 57–83.

Winship, C., & Morgan, S. L. (1999). The estimation of causal effects from observational data. *Annual Review of Sociology*, *25*, 659–707.

Wolf, F. M., Chandler, T. A., & Spies, C. J. (1981). A crossed-lagged panel analysis of quality of school life and achievement responsibility. *Journal of Educational Research*, *74*, 363–368.

Woo, M.-J., Reiter, J. P., & Karr, A. F. (2008). Estimation of propensity scores using generalized additive models. *Statistics in Medicine*, *27*, 3805–3816.

Woodward, J. A., & Bonett, G. B. (1991). Simple main effects in factorial designs. *Journal of Applied Statistics*, *18*(2), 255–264. doi: 10.1080/02664769100000019

Wright, S. (1918). On the nature of size factors. *Genetics*, *3*, 367–374.

Wright, S. (1921). Correlation and causation. *Journal of Agricultural Research*, *20*(7), 557–585.

Wright, S. (1923). The theory of path coefficients: A reply to Niles's criticism. *Genetics*, *8*, 239–255.

Wright, S. (1934). The method of path coefficients. *Annals of Mathematical Statistics*, *5*, 161–215.

Wright, S. (1960a). Path coefficients and path regressions: Alternative or complementary concepts? *Biometrics*, *16*(2), 189–202.

Wright, S. (1960b). The treatment of reciprocal interaction, with or without lag, in path analysis. *Biometrics*, *16*(3), 423–445.

Yule, G. U. (1903). Notes on the theory of association of attributes of statistics. *Biometrika*, *2*, 121–134.

Author Index

Abraham, W. T., 53, 336
Agresti, A., 7, 53
Aiken, L. S., 19, 53, 54, 62
Aitkin, M., 8, 18, 53
Allen, M. P., 53
Amemiya, Y., 62, 321, 325
Angrist, J. D., 53
Appelbaum, M. I., 8, 18, 53
Arbuckle, J. L., XIV, 53
Asparouhov, T., 19, 56
Aubrecht, A. M., 42, 53
Austin, J. T., 31, 58
Ayuso-Mateos, J. L., 54

Barnard, J., 53
Baron, R. M., 13, 15, 53
Bauer, H., 53
Bauer, M., XIV, 59
Bentler, P. M., X, XIV, 31, 53, 312
Berger, M. P. F., 53, 267
Biesanz, J. C., 18, 42, 53, 62
Bollen, K. A., 31, 42, 53, 54
Bonett, G. B., 63, 118
Bonney, G. E., 54
Borooah, V. K., 54
Borsboom, D., 31, 54
Bouvier, A., 56, 316
Bovaird, J. A., 31, 57
Bramley-Moore, L., 3, 59
Brown, C. H., 19, 56
Browne, M. W., XIV, 54
Brumback, B., 59

Campbell, D. T., V, VIII, IX, XIII, 18, 54, 60, 341
Canivez, G. L., 42, 62
Card, N. A., 31, 57
Carlson, J. E., 8, 54
Cartwright, N., XIII, 54, 190
Casella, G., 60
Chandler, T. A., 42, 62
Cheng, J., 19, 54
Clayton, D., 4, 56
Cochran, W. G., 54, 307
Cohen, J., 9, 19, 54, 59, 307
Cohen, P., 19, 54
Cole, D., 40, 62
Cook, E. F., 54
Cook, T. D., V, VIII, XIII, 18, 54, 60, 341
Cox, D. R., XIV, 31, 54
Cramer, E. M., 8, 18, 53
Crandall, C. S., 31, 57
Cummings, J. N., 55, 336
Curran, P. J., 42, 53, 54

Dalgard, O. S., 54
Dawid, A. P., 54
Deeb-Sossa, N., 42, 53
Delaney, H. D., 18, 19, 58, 318
Dietzfelbinger, L., XIV, 58
Donaldson, S. I., 55, 336
Dowrick, C., 54
Draper, N. R., 54
Du, J., 53
Dunn, G., 19, 54

Eid, M., 42, 61, 319

Fahrmeir, L., 55, 316
Feller, W., 55
Fichman, M., 55, 336
Fisher, R. A., 55

Gabler, S., 61, 190
Geiser, C., 40, 62
Gelman, A., 55
Georgii, H.-O., 55
Glymour, C., X, 61
Goldman, L., 54
Gosslee, D. G., 8, 18, 55
Graham, J. W., 55, 336
Green, J. L., 8, 56
Green, W. H., 55
Greenland, S., 5, 18, 31, 55, 60, 150, 336, 342
Györfi, L., 55, 316

Hamerle, A., 55
Hancock, G. R., 55, 321, 325
Harding, D. J., 58, 224
Hau, K. T., 58, 321
Hernán, M. A., 4, 56
Hérnan, M. A., 59
Hill, J. L., 53, 55
Höfler, M., 18, 56
Holland, P. W., VIII, 3, 5, 31, 39, 56, 114, 125
Horvitz, D. G., 56
Hosmer, D. W., 56
Hoyer, J., 56, 321, 325
Huet, S., 56, 316

Ialongo, N. S., 19, 56
Imbens, G. W., 53

Jamieson, J., 56, 304
Jennings, E., 8, 56
Jo, B., 19, 29, 56, 342
Jolivet, E., 56, 316
Jöreskog, K. G., X, XIV, 56, 312, 318, 326
Judd, C. M., 57, 321, 325

Kaplan, D., 31, 56
Karr, A. F., 63, 274
Keele, L., 56, 274, 283
Keiding, N., 4, 56
Kenny, D. A., 13, 15, 42, 53, 56, 57, 321, 325
Keren, G., 8, 57
King, G., 57, 274, 283
Klauer, K. J., 57, 321, 325, 338
Klein, A., 56, 321, 325
Klein, A. G., 57, 321, 325
Klenke, A., 57
Kohler, M., 55
Kolmogorov, A. N., XI, 57
Kramer, C. Y., 8, 57
Kröhne, U., XIV, 59
Krzyzak, A., 55

Lane, P. W., 9, 59
Langsrud, Ø., 9, 11, 57
Lash, T. L., 5, 60
Lechner, M., 57
Lee, A., 3, 59
Lee, B. K., 57, 274, 283
Lee, S.-Y., 31, 57
Lee, S. Y., 61, 321, 325
Lehtinen, V., 54
Lei, P., 42, 62
Lemeshow, S., 56, 57
Lessler, J., 57, 274
Levy, P. S., 57
Lewis, C., 8, 57
Liao, T. F., 57
Little, T. D., 31, 57
Loève, M., 57
Lord, F. M., 19, 57
Lucas, H. L., 8, 18, 55

MacCallum, R. C., 31, 58
MacKinnon, D. P., 13, 15, 58
Maracy, M., 54
Marsh, H. W., 58, 321, 325
Maxwell, S. E., 18, 19, 58, 318
Mayer, A., XII, XIV, 13, 40, 58, 62
McArdle, J. J., 42, 58
McCaffrey, D. F., 58, 274, 283
McCullagh, P., 58
McCulloch, C. E., 60
Mellenbergh, G. J., 31, 54
Mels, G., XIV, 54
Meredith, M., 42, 58
Mill, J. S., VI, 58, 119
Moosbrugger, H., 57, 321, 325
Morgan, S. L., 18, 58, 62, 139, 224
Morral, A. R., 58, 274
Muthén, B. O., X, XIV, 19, 31, 56–58, 312, 318, 321, 325, 326

Muthén, L. K., X, XIV, 31, 58, 312, 318, 326

Nachtigall, C., 61, 190
Nagel, W., XIV, 5, 7, 15, 21–23, 29, 30, 42, 62, 216
Nagengast, B., XIV, 37, 58, 59, 137
Nelder, J. A., 9, 58, 59

OpenMx, XIV, 59
Overall, J. E., 8, 9, 18, 59

Page, H., 54
Partchev, I., 62, 312, 318, 326, 335, 338
Pearl, J., X, XIV, 14, 18, 31, 55, 59
Pearson, K., 3, 59
Phye, G. D., 57, 321
Pitts, S. C., 18, 62
Poursat, M.-A., 56, 316
Preacher, K. J., 31, 57
Pukelsheim, F., 59, 267

Rao, C. R., 59, 307
Reichardt, C. S., 18, 59, 310
Reiter, J. P., 63, 274
Ridgeway, G., 58, 274
Robins, J., 19, 59
Robins, J. M., 19, 31, 55, 59, 150, 342
Rogosa, D., 42, 59, 60, 313
Rose, N., XII, 58
Rosenbaum, P. R., 18, 19, 29, 60, 215, 220, 270
Rosseel, Y., XIV, 58, 60
Rothman, K. J., 5, 60
Rotnitzky, A., 19, 59
Rubin, D. B., 18, 19, 39, 53, 60, 215, 220, 270
Russell, D. W., 53, 336

Sagarin, B. J., 62
Särndal, C.-E., 60
Scheines, R., X, 61
Schuster, C., 62
Schwenkmezger, P., 42, 61
Searle, S. R., 60
Senn, S., 19, 60
Shadish, W. R., V, IX, XIII, 18, 60, 334, 336, 341, 343
Simpson, E. H., 3, 4, 61
Singer, J. D., 42, 61
Small, D. S., 19, 54
Smith, H., 54
Sobel, M. E., 39, 61
Song, X. Y., 61, 321, 325
Sörbom, D., X, XIV, 56, 61, 312, 318, 321, 325, 326, 338
Spiegel, D. K., 8, 9, 59
Spies, C. J., 42, 62
Spirtes, P., X, XIV, 15, 18, 31, 61
Splawa-Neyman, J., 18, 61
Spohn, W., XIII, 61
Stanley, J. C., V, VIII, XIII, 18, 54, 341
Stegmüller, W., XIII, 61
Steyer, R., XII, XIV, 5, 7, 15, 21–23, 29, 30, 40, 42, 58, 59, 61, 62, 190, 216, 312, 318, 319, 326, 335, 338
Stuart, E. A., 57, 274
Suppes, P., XIII, 62
Swensson, B., 60

Takezawa, K., 62
Thoemmes, F., XII, 58
Thomas, N., 60
Thompson, D. J., 56
Timm, N. H., 8, 54
Tisak, J., 42, 58, 62
Tisak, M. S., 42, 62
Tutz, G., 55, 316

van Breukelen, G. J. P., 19, 62
van Heerden, J., 31, 54
von Davier, A. A., 61, 190
von Eye, A., 62

Wainer, H., 19, 62
Walk, H., 55
Wall, M. M., 62, 321, 325
Watkins, M. W., 42, 62
Wen, Z., 58, 321
Wermuth, N., XIV, 31, 54
West, S. G., XII, 18, 19, 53, 54, 58, 62
Willett, J. B., 42, 61
Williams, J. D., 9, 62
Willmes, K., 57, 321
Winship, C., 18, 58, 62, 139
Wolf, F. M., 42, 62
Wong, W. K., 53, 267

Woo, M.-J., 63, 274, 283
Woodward, J. A., 63, 118
Wretman, J. H., 60
Wright, S., VIII, 15, 63
Wu, E. J. C., X, 53

Yule, G. U., 3, 4, 63

Zeng, L., 57, 274, 283

Subject Index

analysis of covariance, 19
ANOVA
    nonorthogonal, 8
atomic total effect variable, 77
average causal effect, VIII, IX
average direct effect, 114
average indirect effect
    given a value of a covariate, 117
average of the conditional prima facie effects, 131
average of the conditional total effects, 131
average prima facie effect, 6
average total effect, 103, 117, 118, 120, 121, 137
    definition, 104, 117
    in a subpopulation, 138
    of the treated, 110
    vs. prima facie effect, 153

baseline bias, 139–141, 153, 333
    of the conditional *PFE*-function, 139, 140
between-group designs, IX
between-group experiment, IX
between-group quasi-experiment, IX, X
bias of the conditional prima facie effect, *139*
bias of the prima facie effect, *139*
Bias Theorem
    second, 139
biased covariate-treatment regression
    example, 127
biased treatment regression
    example, 127, 130

causal inference
    vs. statistical inference, 9, 19, 21
completeness
    example, 175
condition independence
    of true outcomes and treatments given a covariate, 231
conditional causal effect, IX
conditional direct effect
    given a value of V, 114
    given a value of a covariate, 115
conditional direct total effect
    given a value of V, 117
conditional expectations of outcomes
    and of the individual expected outcomes, 137
conditional independence
    of true outcomes and treatments
      definition, 219
    of units and treatments given a covariate, 231
conditional prima facie effect, 138
    given the value of a covariate, 136
conditional random assignment, 187, 237
conditional randomization, *178–179*, 336
conditional regressive independence
    of true outcomes and treatments, 224
    of true outcomes from X, 226
    of true outcomes from treatments, 231
conditional total effect
    given a treatment, *109*, 120, 121
      interpretation, 111
      substantive meaning, 109
    given a value of a covariate, 117, 121

given the value v of a variable V, 103
given the value of a covariate, 120
given the value of a random variable V, 106
given treatment and covariate, *112*
of the untreated, 110, 121
conditional total effects
and average total effects, 131
conditional unit-treatment homogeneity, 231
conditionally randomized experiment, 28
conditioning on a covariate, 132
counterfactual perspective, 118
covariate
multi-dimensional, 29
qualitative, 29
quantitative, 29
uni-dimensional, 29
covariate selection, 336

derived parameter, 127
differential item functioning, X

effect bias, 139–141, 154, 333
of the conditional *PFE*-function, 139, 140
effect function, 311, 316
EffectLiteR, XIV
experiment
vs. quasi-experiment, 28
exposition, 18

fallible covariate, 31
filtration, 48
fundamental parameter, 127

independence
of true outcomes and treatments, *218–223*, 231
definition, 219
example, 221
testability, 220
of units and treatments, 231
independence of units and treatments, 336
example, 171
testability, 180
individual causal effect, VII, IX
individual total effect, 108
intercept function, 311, 316
interpretation
of the outcome variable, 342
of the random experiment, 343
of the treatment variable, 342
intervention, 18

main effect, 311, 317
in analysis of variance, 104, 105, 118
main effects, 9
mean adjustment, 335, 337
with propensities, 339
mean differences
and their meaning, 148
measurement invariance, X

non-compliance, 19
non-equivalent control-group design, 18
nonorthogonal ANOVA, 8, 18

observational unit, IX
observational-unit variable, IX
outcome variable, 26, 32, 34, 36, 37, 39, 40
outcome weighting, 335, 340

pre-facto vs. counterfactual perspective, 118
prima facie effect, 3, 17, 137
average, 6
in a subpopulation, 5, 10
in the total population, 9
three kinds of, 12, 17, 20
prima facie effects, 118, 120, 331, 333
average, 311, 317
propensity, 161
propensity score methods, 19

quasi-experiment, 17, 334
in between-group designs, X
vs. true experiment, X

random experiment, IX, 125
randomization, 177, 237, 336
within subsets of units, 178
randomized experiment, VI, 28, 51, 130

randomized trial, 18
regressive independenc
    of true outcomes from treatments, 231
regressive independence
    of true outcomes and treatments, *224–230*
repeated-measures design, VIII
reversal of effect, 17

self-selection of subjects to treatments, 129
Simpson's paradox, *3–121*
    population parameters, 4
single-unit trial, 30
    vs. sample, 27
statistical inference
    vs. causal inference, 9, 19, 21
stochastic causality, X
strong ignorability, 220, 233

treatment, 18
treatment probability, 127
treatment variable, 26, 31, 36, 37, 39, 40, 52
true experiment, 334
    vs. quasi-experiment, X
true-outcome variable, 74, 76

unbiased covariate-treatment regression
    example, 128, 130
unbiased treatment regression
    example, 128
unbiasedness, 121, 144
    of the covariate-treatment regression, 334
    of the prima facie effects, 124
    of the*PFE*-function, 124
unconfounded covariate-treatment regression
    example, 194
unconfoundedness, 204
    example, 194
unit-treatment homogeneity, 231

weighting theorem, *284–289*
within-group designs, IX