Conditional Expected Value: Definition

Definition 6.1

Let X and Y be random variables on the probability space $\langle \Omega, \mathfrak{A}, P \rangle$. If Y is real-valued with finite values $y_1, \ldots, y_n$ and if $P(X = x) > 0$, then the *conditional expected value* of Y given $X = x$ is defined as the sum of the values y_i weighted with their conditional probabilities $P(Y = y_i \mid X = x)$:

$$E(Y \mid X = x) := \sum_{i=1}^{n} y_i \cdot P(Y = y_i \mid X = x)\,.$$

Conditional Expected Value: Special Case

If Y has only two values 1 and 0, then the equation implicits:

$E(Y \mid X = x) = 1 \cdot P(Y = 1 \mid X = x) + 0 \cdot P(Y = 0 \mid X = x) \; = P(Y = 1 \mid X = x)$.

Conditional Expected Value: Special Case

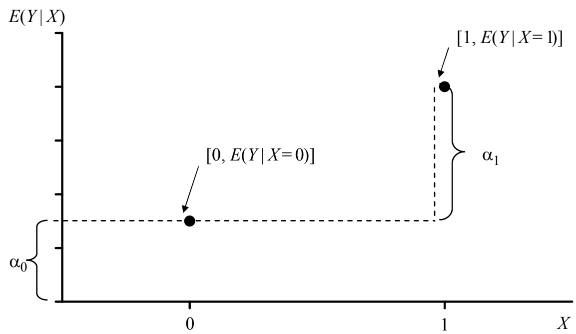

Figure 6.2. A regression in which X has only the two values 0 and 1.

Conditional Expected Value: Rules of Computation

If X and Y as well as Y_1 and Y_2 are numerical random variables on the probability space $\langle \Omega, \mathfrak{A}, P \rangle$ with finite expected values and if α and

$\beta \in \mathbb{R}$, then:

(i) $E(\alpha \mid X = x) = \alpha$

(ii) $E(\alpha \cdot Y \mid X = x) = \alpha \cdot E(Y \mid X = x)$

(iii) $E(\alpha \cdot Y_1 + \beta \cdot Y_2 \mid X = x) = \alpha \cdot E(Y_1 \mid X = x) + \beta \cdot E(Y_2 \mid X = x)$

(iv) $E(Y \mid X = x) = \sum_z E(Y \mid X = x, Z = z) \cdot P(Z = z \mid X = x)$

Regression: Definition

Definition 6.2

Given the assumptions of Definition 6.1, the regression $E(Y|X)$ can be defined as the function of X, the values of which are the conditional expected values $E(Y|X=x)$ of Y given $X=x$.

Conditional Expected Value: Figure

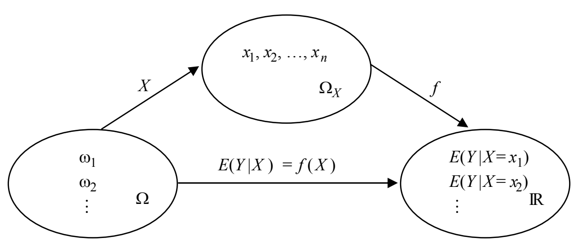

Figure 6.1 The relationship between the set Ω of possible outcomes, the regressor X and the conditional expected values $E(Y|X=x)$.

Regression: Rules of Computation

(i) $E(\alpha \mid X) = \alpha, \quad \alpha \in \mathbb{R}$

(ii) $E(\alpha Y \mid X) = \alpha\ E(Y \mid X), \quad \alpha \in \mathbb{R}$

(iii) $E(\alpha Y_1 + \beta Y_2 \mid X) = \alpha\ E(Y_1 \mid X) + \beta\ E(Y_2 \mid X), \quad \alpha, \beta \in \mathbb{R}$

(iv) $E[E(Y \mid X)] = E(Y)$

(v) $E[f(X) \mid X] = f(X)$, if $f(X)$ is numerical

(vi) $E[E(Y \mid X) \mid f(X)] = E[Y \mid f(X)]$

(vii) $E[f(X) \cdot Y \mid X] = f(X) \cdot E(Y \mid X)$, if $f(X)$ is numerical.

Residual: Definition

Definition 6.3

The *residual* ε with respect to the regression $E(Y \mid X)$ is defined to be the deviation of the random variable Y from $E(Y \mid X)$, i.e.

$$\varepsilon := Y - E(Y \mid X).$$

Residual: Figure

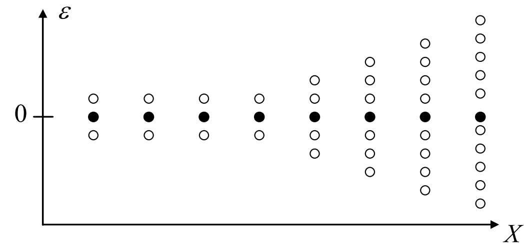

Figure 6.3

The regression of ε on a numerical regressor X. The conditional variances of the residual depend on X. The symbol • marks the values of the regression $E(\varepsilon \mid X)$ and ○ the values of the residual ε.

Residual: Properties

(i) $Y = E(Y \mid X) + \varepsilon$

(ii) $E(\varepsilon) = 0$

(iii) $E(\varepsilon \mid X) = 0$

(iv) $E[\varepsilon \mid f(X)] = 0$

(v) $E[\varepsilon \mid E(Y \mid X)] = 0$

(vi) $Cov(\varepsilon, X) = 0$, if X is numerical

(vii) $Cov(\varepsilon, X_i) = 0$, $i = 1, ..., m$, if $X = (X_1, ..., X_m)$ is numerical

(viii) $Cov[\varepsilon, f(X)] = 0$, if $f(X)$ is numerical

(ix) $Cov[\varepsilon, E(Y \mid X)] = 0$

(x) $Var(Y) = Var[E(Y \mid X)] + Var(\varepsilon)$

Coefficient of Determination

A term that is direcly related to the properties of the residual ε [particularly to property (x)] is the *coefficient of determination*. It is defined for every numerical random variable Y with finite expected value $E(Y)$ as well as finite and positive variance $Var(Y)$ by:

$$R^2_{Y|X} := \frac{Var[E(Y|X)]}{Var(Y)}, \quad if\ Var(Y) > 0$$

Multiple Correlation

The positve square root of the coefficient of determination $R^2_{Y|X}$ is called the *multiple correlation* of Y with respect to X and is denoted $R_{Y|X}$.

The coefficient of determination can be interpreted as the *proportion of the variance* of Y that is *determined* by X. Together with the variance of the residual it sums up to 1, i.e., if $Var(Y) > 0$:

$$1 = \frac{Var(Y)}{Var(Y)} = \frac{Var[E(Y|X)]}{Var(Y)} + \frac{Var(\varepsilon)}{Var(Y)}.$$

Regression: General Definition I

Definition 6.4

Let Y be a random variable on the probability space $\langle \Omega, \mathfrak{A}, P \rangle$ with finite expected value [i.e. $-\infty < E(Y) < +\infty$] and X a discrete random variable on $\langle \Omega, \mathfrak{A}, P \rangle$ with values $x_1, \ldots, x_n$, and $P(X = x_i) > 0$. The random variable

$$E(Y \mid X) := \sum_{i=1}^{n} E(Y \mid X = x_i) \cdot I_{X = x_i}$$

is called the regression of Y on X or conditional expectation of Y with respect to X.

Regression: General Definition II

Definition 6.5

Let $\langle \Omega, \mathfrak{A}, P \rangle$ be a probability space, Y a numerical random variable $\langle \Omega, \mathfrak{A}, P \rangle$ with finite expected value, $X: \Omega \to \Omega'$ a random variable on $\langle \Omega, \mathfrak{A}, P \rangle$, $\mathfrak{A}$ a σ-algebra on Ω, $\bar{\mathfrak{B}}$ the Borel σ-algebra on $\overline{\mathbb{R}}$ and $X^{-1}(\mathfrak{A}') := \{X^{-1}(A') \colon A' \in \mathfrak{A}'\}$. Then every numerical random variable $Z: \Omega \to \overline{\mathbb{R}}$ on $\langle \Omega, \mathfrak{A}, P \rangle$ is called *conditional expectation* or *regression* of Y on X, if the following conditions hold:

(a) $Z^{-1}(B) \in X^{-1}(\mathfrak{A}')$, for all $B \in \bar{\mathfrak{B}}$;

(b) $E(I_C \cdot Z) = E(I_C \cdot Y)$, for all $C \in X^{-1}(\mathfrak{A}')$.

Instead of Z we usually write a regression of Y on X as $E(Y \mid X)$.