Bedingter Erwartungswert: Definition

Definition 6.1

Seien X und Y Zufallsvariablen auf dem Wahrscheinlichkeitsraum $\langle \Omega, \mathfrak{A}, P \rangle$. Ist Y reellwertig mit endlich vielen Werten $y_1, ..., y_n$ und ist $P(X = x) > 0$, dann ist der *bedingte Erwartungswert* von Y gegeben $X = x$ die mit den bedingten Wahrscheinlichkeiten $P(Y = y_i \mid X = x)$ gewichtete Summe ihrer Werte:

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

Bedingter Erwartungswert: Abbildung

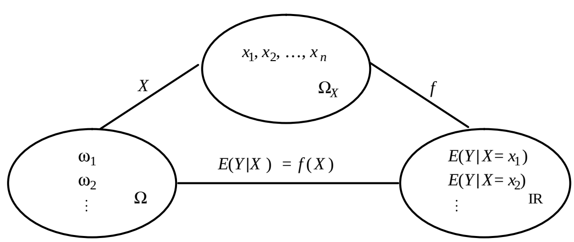

Abbildung 6.1. Die Beziehung zwischen der Menge Ω der möglichen Ergebnisse, dem Regressor X und den bedingten Erwartungswerten $E(Y \mid X = x)$.

Bedingter Erwartungswert: Spezialfall

3

Nimmt Y nur die beiden Werte 1 und 0 an, dann folgt aus der Gleichung:

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

$$= P(Y = 1 \mid X = x).$$

Conditional Expected Value: Special Case

4

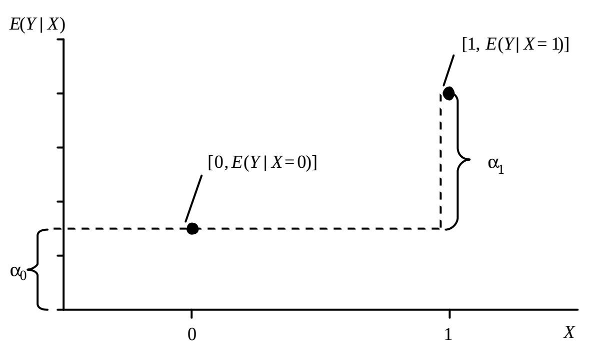

Abbildung 6.2. Darstellung einer linearen Regression im Fall, dass X nur die Werte 0 und 1 annimmt.

Bedingter Erwartungswert: Rechenregeln

Sind X und Y sowie Y_1 und Y_2 numerische Zufallsvariablen auf $\langle \Omega, \mathfrak{A}, P \rangle$ mit endlichen Erwartungswerten sowie α und $\beta \in \mathbb{R}$, dann gelten:

(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

Unter den gleichen Voraussetzungen wie in Definition 6.1 kann die Regression $E(Y|X)$ als diejenige Funktion von X definiert werden, deren Werte die bedingten Erwartungswerte $E(Y|X = x)$ von Y gegeben $X = x$ sind.

Regression: Rechenregeln

(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) $E[f(X) \mid X] = f(X)$, falls $f(X)$ numerisch ist

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

(vii) $E[f(X) \cdot Y \mid X] = f(X) \cdot E(Y \mid X)$, falls $f(X)$ numerisch ist

Residuum: Definition

Definition 6.3

Das *Residuum* e bezüglich einer Regression $E(Y \mid X)$ ist definiert als Abweichung der Zufallsvariablen Y von $E(Y \mid X)$, d. h.:

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

Residuum: Beispiel

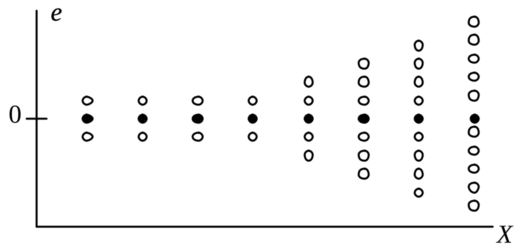

Abbildung 6.3

Die Regressionen von e auf einen numerischen Regressor X. In diesem Beispiel sind die bedingten Varianzen des Residuums von X abhängig. Das Zeichen • markiert die Werte der Regression $E(e \mid X)$ und ∘ die Werte des Residuums e.

Residuum: Eigenschaften

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

(ii) $E(e) = 0$

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

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

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

(vi) $Cov(e, X) = 0$, falls X numerisch ist

(vii) $Cov(e, X_i) = 0$, $i = 1, ..., m$, falls $X = (X_1, ..., X_m)$ numerisch ist

(viii) $Cov[e, f(X)] = 0$, falls $f(X)$ numerisch ist

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

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

Determinationskoeffizient

Ein Begriff, der unmittelbar auf den oben behandelten Eigenschaften des Residuums e [insb. auf Regel (x)] basiert, ist der des *Determinationskoeffizienten*, der für jede numerische Zufallsvariable Y mit endlichem Erwartungswert $E(Y)$ sowie endlicher und positiver Varianz $Var(Y)$ definiert ist, und zwar durch:

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

Multiple Korrelation

Die positive Wurzel aus dem Determinationskoeffizienten $R^2_{Y|X}$ heißt die *multiple Korrelation* von Y bezüglich X und wird mit $R_{Y|X}$ notiert.

Der Determinationskoeffizient lässt sich als *der durch X determinierte Varianzanteil von Y* interpretieren. Wie man aus der vorangegangenen Gleichung sehen kann, addiert er sich mit dem Residualvarianzanteil von Y zu 1 auf, falls $Var(Y) > 0$:

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

Regression: Allgemeine Definition I

Definition 6.4

Seien Y eine numerische Zufallsvariable auf dem Wahrscheinlichkeitsraum $\langle \Omega, \mathfrak{A}, P \rangle$ mit endlichem Erwartungswert [d. h. $-\infty < E(Y) < +\infty$] und X eine diskrete Zufallsvariable auf $\langle \Omega, \mathfrak{A}, P \rangle$ mit den Werten $x_1, ..., x_n$, für die jeweils $P(X = x_i) > 0$ gilt. Dann heißt die Zufallsvariable

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

Regression von Y auf X oder bedingte Erwartung von Y unter X.

Regression: Allgemeine Definition II

Definition 6.5

Seien $\langle \Omega, \mathfrak{A}, P \rangle$ ein Wahrscheinlichkeitsraum, Y eine numerische Zufallsvariable auf $\langle \Omega, \mathfrak{A}, P \rangle$ mit endlichem Erwartungswert, $X: \Omega \to \Omega'$ eine Zufallsvariable auf $\langle \Omega, \mathfrak{A}, P \rangle$, $\mathfrak{A}$ eine σ-Algebra auf Ω, $\overline{\mathfrak{B}}$ die Borelsche σ-Algebra auf $\overline{\mathbb{R}}$ und $X^{-1}(\mathfrak{A}') := \{X^{-1}(A'): A' \in \mathfrak{A}'\}$ die Urbild-σ-Algebra von X. Dann heißt jede numerische Zufallsvariable $Z: \Omega \to \overline{\mathbb{R}}$ auf $\langle \Omega, \mathfrak{A}, P \rangle$ *bedingte Erwartung* oder *Regression* von Y auf X, falls sie die folgenden Bedingungen erfüllt:

(a) $Z^{-1}(B) \in X^{-1}(\mathfrak{A}')$, für alle $B \in \overline{\mathfrak{B}}$;

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

Anstelle von Z schreibt man für eine Regression von Y auf X meist $E(Y \mid X)$.