

1 Bedingte Erwartungswerte

Die folgenden Regeln sind das alltägliche Handwerkszeug für den Umgang mit bedingten Erwartungen und werden in diesem Abschnitt, allerdings ohne Beweise, zitiert. Es ist durchaus eine lohnenswerte Übung, die Aussagen zu beweisen oder die Beweise noch einmal nachzuarbeiten.

Sei (Ω, \mathcal{A}, P) ein fester Wahrscheinlichkeitsraum. Wir bezeichnen mit $L^1(\Omega, \mathcal{A}, P)$ die Menge aller \mathcal{A} -messbaren $\xi : \Omega \mapsto \mathbb{R}$, so dass $E|\xi| < \infty$. Wir schreiben $\mathcal{B} \subset \mathcal{A}$, falls \mathcal{B} eine σ -Algebra und Teilmenge von \mathcal{A} ist.

Definition 1. Sei $\xi \in L^1(\Omega, \mathcal{A}, P)$ und $\mathcal{B} \subset \mathcal{A}$. Eine Zufallsvariable $Z \in L^1(\Omega, \mathcal{A}, P)$ heisst *bedingte Erwartung* (von ξ gegeben \mathcal{B}) falls

- (i) Z ist \mathcal{B} -messbar
- (ii) für alle $B \in \mathcal{B}$ gilt

$$\int_B Z dP = \int_B \xi dP.$$

Für zwei Z_1 und Z_2 , welche beide (i) und (ii) dieser Definition erfüllen, gilt, dass $Z_1 = Z_2$ P -f.s.; und wir definieren durch alle solche Z eine Äquivalenzklasse, welche wir mit $E[\xi|\mathcal{B}]$ bezeichnen.

Die zugehörige bedingte Wahrscheinlichkeit definieren wir durch

$$P(A|\mathcal{B}) := E[1_A|\mathcal{B}].$$

Beispiel 2. Wir beginnen mit einer Reihe von Beispielen:

- (i) Ist $\mathcal{B} = \{\emptyset, \Omega\}$, so ist $E[\xi|\mathcal{B}] = \mathbb{E}[\xi]$.
- (ii) Wird \mathcal{B} durch paarweise disjunkte $A_i \in \mathcal{A}$ mit $P(A_i) > 0$ erzeugt, so gilt

$$E[\xi|\mathcal{B}] = \sum_{i=1}^{\infty} 1_{A_i} \frac{\int_{A_i} \xi dP}{P(A_i)}.$$

Der nun folgende erste Satz von Goldenen Regeln bildet die Grundausstattung.

Lemma 3. Sei $\xi, \xi_1, \xi_2 \in L^1(\Omega, \mathcal{A}, P)$ und $a, b \in \mathbb{R}$. Dann gilt

- (i) $\int_B E[\xi|\mathcal{B}] dP = \int_B \xi dP$ für alle $B \in \mathcal{B}$

1 Bedingte Erwartungswerte

- (ii) $E[E[\xi|\mathcal{B}]] = E[\xi]$
- (iii) $\xi = a$ *P-f.s.* $\Rightarrow E[\xi|\mathcal{B}] = a$ *P-f.s.*
- (iv) $0 \leq \xi$ *P-f.s.* $\Rightarrow 0 \leq E[\xi|\mathcal{B}]$ *P-f.s.*
- (v) $E[a_1\xi_1 + a_2\xi_2|\mathcal{B}] = a_1E[\xi_1|\mathcal{B}] + a_2E[\xi_2|\mathcal{B}]$ *P-f.s.*
- (vi) $\xi_1 \leq \xi_2$ *P-f.s.* $\Rightarrow E[\xi_1|\mathcal{B}] \leq E[\xi_2|\mathcal{B}]$ *P-f.s.*
- (vii) ξ ist \mathcal{B} -messbar $\Rightarrow E[\xi|\mathcal{B}] = \xi$.

Lemma 4. Sei $\xi \in L^1(\Omega, \mathcal{A}, \mathbb{P})$ und η \mathcal{B} -messbar (!), so dass $\eta \cdot \xi \in L^1(\Omega, \mathcal{A}, \mathbb{P})$. Dann gilt

$$E[\eta \cdot \xi|\mathcal{B}] = \eta E[\xi|\mathcal{B}] \quad \text{P-f.s..}$$

Diese Regel erlaubt es, messbare Faktoren aus dem bedingten Erwartungswert herauszuziehen.

Lemma 5. Für $\xi \in L^1(\Omega, \mathcal{A}, \mathbb{P})$ und $\mathcal{B}_1 \subset \mathcal{B}_2 \subset \mathcal{A}$ gilt

$$E[E[\xi|\mathcal{B}_1]|\mathcal{B}_2] = E[\xi|\mathcal{B}_1] = E[E[\xi|\mathcal{B}_2]|\mathcal{B}_1] \quad \text{P-f.s..}$$

Was hat man in diesem neuen Kontext unter Unabhängigkeit zu verstehen? Zunächst definieren wir Unabhängigkeit für σ -Algebren.

Definition 6. \mathcal{B}_1 und \mathcal{B}_2 heißen *unabhängig*, falls

$$P(B_1 \cap B_2) = P(B_1)P(B_2) \quad \text{für alle } B_i \in \mathcal{B}_i, \quad i = 1, 2.$$

Für Zufallsvariablen ξ_1 und ξ_2 welche jeweils B_i -messbar sind, erhalten wir $E[\xi_1\xi_2] = E[\xi_1]E[\xi_2]$. Die von einer Zufallsvariablen *erzeugte* σ -Algebra definieren wir wie folgt:

$$\sigma(\xi) := \left\{ \xi^{-1}(D) : D \subset \mathcal{B}(\mathbb{R}) \right\},$$

wobei $\mathcal{B}(\mathbb{R})$ die Borel σ -Algebra auf \mathbb{R} bezeichnet. Zwei Zufallsvariablen ξ_1 und ξ_2 heißen *unabhängig*, falls $\sigma(\xi_1)$ und $\sigma(\xi_2)$ unabhängig sind.

Lemma 7. Sei $\xi \in L^1(\Omega, \mathcal{A}, \mathbb{P})$ und $\sigma(\xi)$ unabhängig von \mathcal{B} . Dann gilt

$$E[\xi|\mathcal{B}] = E[\xi] \quad \text{P-f.s..}$$

Den bedingten Erwartungswert kann man als Erwartungswert gegeben der in \mathcal{B} enthaltenen Information interpretieren. Das vorige Lemma bedeutet intuitiv, dass unter Unabhängigkeit die zusätzliche Information keinen weiteren Nutzen bringt: der bedingte Erwartungswert ist gleich dem unbedingten Erwartungswert.

Im Folgenden schreiben wir für die *bedingte Erwartung bezüglich* η

$$E[\xi|\eta] := E[\xi|\sigma(\eta)].$$

Lemma 8. *Es existiert eine messbare Funktion $f : \mathbb{R} \mapsto \mathbb{R}$, so dass*

$$E[\xi|\eta] = f(\eta).$$

Diese Funktion f nennen wir *Faktorisierung* von ξ bezgl. η und setzen

$$E[\xi|\eta = x] := f(x).$$

In der Statistik nennt man die Funktion f auch *Regressionsfunktion*.

Bedingte Dichte. Haben die Zufallsvariablen eine gemeinsame Dichte, so kann man die bedingte Dichte ausrechnen: Bezeichne $f(x, y)$ die gemeinsame Dichte von (ξ, η) und f die Dichte von η . Dann gilt auf $\{x : f(x) > 0\}$

$$E[\xi|\eta = x] = \frac{\int y f(x, y) dy}{f(x)} =: \int y f(y|x) dy.$$

Die bedingte Dichte von ξ gegeben $\eta = x$ ist also gerade

$$f(y|x) = \frac{f(x, y)}{f(x)}.$$

Die folgende Regel ermöglicht es, Strukturen mit Unabhängigkeiten effektiv auszunutzen.

Lemma 9. ξ und η seien unabhängig und die Funktion $T : \mathbb{R}^2 \mapsto \mathbb{R}$ erfülle $T(\xi, \eta) \in L^1(\Omega, \mathcal{A}, \mathbb{P})$. Mit P_η bezeichnen wir die Verteilung von η . Dann gilt

$$E[T(\xi, \eta)|\eta = x] = E[T(\xi, x)] \quad P_\eta - \text{f.s.}$$

Beispiel 10. Ist $T(x, y) = xy$, so erhält man, dass

$$E[\xi\eta|\eta = x] = xE[\xi]$$

und, durch Einsetzen von η , dass $E[\xi\eta|\eta] = \eta E[\xi]$.

Bedingte Unabhängigkeit. Im Rahmen der bedingten Wahrscheinlichkeiten erhält man auch ein neues Konzept der bedingten Unabhängigkeit.

Definition 11. Sei $\mathcal{B} \subset \mathcal{A}$ eine σ -Algebra. A_1, \dots, A_n heißen *bedingt unabhängig* bezgl. \mathcal{B} , falls

$$P\left(\bigcap_{i=1}^n A_i | \mathcal{F}\right) = \prod_{i=1}^n P(A_i | \mathcal{F}) \quad P\text{-f.s.}$$

Sind weiterhin $\mathcal{B}_1, \mathcal{B}_2, \mathcal{B} \subset \mathcal{A}$ σ -Algebren, so nennen wir \mathcal{B}_1 und \mathcal{B}_2 *bedingt unabhängig* bezgl. \mathcal{B} , falls für alle $B_i \in \mathcal{B}_i$, $i = 1, 2$,

$$P(B_1 \cap B_2 | \mathcal{B}) = P(B_1 | \mathcal{B}) P(B_2 | \mathcal{B}).$$

1 Bedingte Erwartungswerte

Bedingte Unabhängigkeit kann man mit Hilfe von Zufallsvariablen charakterisieren, wie folgendes Resultat zeigt.

Lemma 12. \mathcal{B}_1 und \mathcal{B}_2 sind bedingt unabhängig bezgl. \mathcal{B} genau dann, wenn für jedes $\xi \in L^1(\Omega, \mathcal{B}_1, P)$

$$E[\xi | \sigma(\mathcal{B}_2 \cup \mathcal{B})] = E[\xi | \mathcal{B}] \quad P\text{-f.s. .}$$

Martingale in diskreter Zeit

Die wichtigste Anwendung bedingter Erwartungswerte sind Martingale. Das sind Prozesse, die im Mittel einen Aufwärtstrend, einen Abwärtstrend oder eben gar keinen Trend vorweisen.

Wir betrachten die Zeitpunkte $0, 1, 2, \dots$. Ein *stochastischer Prozess* (in diskreter Zeit) ist eine Familie von Zufallsvariablen $M = (M_s)_{s=0,1,2,\dots}$. Man kann natürlich auch andere Zeiträume betrachten, wie etwa $\{0, \dots, T\}$, oder $\{0, t_1, t_2, \dots\}$ bzw. $[0, t]$ oder $\mathbb{R}_{\geq 0}$. In den letzten beiden Fällen spricht man von kontinuierlicher Zeit.

Information wird mit der Hilfe von σ -Algebren modelliert. Das keine Information verloren geht, führt dazu, dass die betrachtete Folge von σ -Algebren wachsend ist.

Definition 13. Eine *Filtration* ist eine wachsende Folge von Sub- σ -Algebren

$$\mathcal{F}_0 \subset \mathcal{F}_1 \subset \mathcal{F}_2 \subset \dots \subset \mathcal{F}.$$

Ein Prozess M heißt *adaptiert*, falls M_t \mathcal{F}_t -messbar ist für alle $t = 0, 1, 2, \dots$. Er heißt *vorhersehbar*, falls M_t bereits \mathcal{F}_{t-1} -messbar ist.

Definition 14. Ein Prozess M heißt *integrierbar*, falls $E[|M_t|] < \infty$. Ein integrierbarer, adaptierter Prozess heißt

- (i) *Martingal*, falls $E[M_t | \mathcal{F}_s] = M_s$,
- (ii) *Submartingal*, falls $E[M_t | \mathcal{F}_s] \geq M_s$,
- (iii) *Supermartingal*, falls $E[M_t | \mathcal{F}_s] \leq M_s$,

jeweils P -fast sicher für $0 \leq s \leq t$,

Für Submartingal M und N erhalten wir

- (i) $aM + bN$ ist wieder ein Submartingal, falls $a, b \geq 0$,
- (ii) $M \vee N$ ist ein Submartingal
- (iii) ist ϕ konvex, so ist auch $\phi(M)$ ein Submartingal, so lange der Prozess integrierbar ist.

Der folgende, Doobsche Zerlegungssatz, besagt, dass ein Submartingal zerlegbar ist in ein Martingal und einen wachsenden, vorhersehbaren Prozess.

Satz 15. *Sei X ein Submartingal. Dann existiert ein Martingal M und ein wachsender Prozess A , so dass*

$$X = A + M.$$

A kann hierbei vorhersehbar gewählt werden. Ist $A_0 = 0$, so ist A eindeutig.

Literatur

- Bauer: Wahrscheinlichkeitstheorie
- Georgii: Stochastik
- Billingsley: Probability Theory
- Kallenberg: Foundations of Modern Probability