

Exkurs A: Bedingte Erwartungswerte, bedingte Verteilungen

(Ω, \mathcal{A}, P) sei W-Raum, $X : \Omega \rightarrow \overline{\mathbb{R}}$ P-quasiintegrierbar, $\mathcal{F} \subset \mathcal{A}$ Unter- σ -Algebra.

$E(X|\mathcal{F}) = E_P(X|\mathcal{F})$ (Version des) **bedingter Erwartungswert von X unter \mathcal{F}** (bzgl.P):
P-f.s. eindeutig bestimmte P-quasiintegrierbare ZV $\Omega \rightarrow \overline{\mathbb{R}}$ mit

$$(1) \quad \mathcal{F} - \text{meßbar}$$

$$(2) \quad \int_F E(X|\mathcal{F}) dP = \int_F X dP \quad \forall F \in \mathcal{F} \text{ (Radon - Nikodym - Gleichung).}$$

Für $A \in \mathcal{A}$ heißt $P(A|\mathcal{F}) := E(1_A|\mathcal{F})$ **bedingte Wahrscheinlichkeit von A unter \mathcal{F}** .

Einige Rechenregeln:

Satz A.1 Es sei $X : \Omega \rightarrow \overline{\mathbb{R}}$ P-quasiintegrierbar, $\mathcal{F} \subset \mathcal{A}$.

$$(a) \quad EE(X|\mathcal{F}) = EX$$

$$(b) \quad X \mathcal{F} - \text{meßbar} \Rightarrow E(X|\mathcal{F}) = X \text{ P-f.s.}$$

$$(c) \quad (\text{Tower property}) \mathcal{F} \subset \mathcal{G} \subset \mathcal{A} \Rightarrow$$

$$E(E(X|\mathcal{G})|\mathcal{F}) = E(X|\mathcal{F}) = E(E(X|\mathcal{F})|\mathcal{G}) \text{ P-f.s.}$$

$$(d) \quad (\text{Taking out what is known}) Y : \Omega \rightarrow \overline{\mathbb{R}} \text{ sei } \mathcal{F}\text{-meßbar, } XY \text{ quasiintegrierbar}$$

$$\Rightarrow E(XY|\mathcal{F}) = YE(X|\mathcal{F}) \text{ P-f.s.}$$

$$(e) \quad X \text{ und } \mathcal{F} \text{ unabhängig}$$

$$\Rightarrow E(X|\mathcal{F}) = EX \text{ P-f.s.}$$

$$(f) \quad (\text{Jensensche Ungl.})$$

$I \subset \mathbb{R}$ sei offenes Intervall, $X \in \mathcal{L}^1(P)$, $P(X \in I) = 1$, $f : I \rightarrow \mathbb{R}$ konvex mit
 $f(X) \geq 0$ oder $f(X) \in \mathcal{L}^1(P)$

$$\Rightarrow E(X|\mathcal{F}) \in I \text{ P-f.s. ,}$$

$$f(E(X|\mathcal{F})) \leq E(f(X)|\mathcal{F}) \text{ P-f.s.}$$

Beweis. Bauer (4. Auflage, 1991), S. 120 - 124, Hoffmann-Jørgensen, Vol. I, S. 457. □

Satz A2 $\mathcal{F} \subset \mathcal{A}$.

$$(a) \quad (\text{Monotone Konvergenz}) X_n, X : \Omega \rightarrow \overline{\mathbb{R}} \text{ seien meßbar, } X_n \uparrow X \text{ f.s., } EX_1^- < \infty \\ \Rightarrow E(X_n|\mathcal{F}) \uparrow E(X|\mathcal{F}) \text{ f.s.}$$

$$\text{Falls } X_n \downarrow X, \text{ f.s., } EX_1^+ < \infty \Rightarrow E(X_n|\mathcal{F}) \downarrow E(X|\mathcal{F}) \text{ f.s.}$$

$$(b) \quad (\text{Beschränkte Konvergenz}) X_n \in \mathcal{L}^1(P), X : \Omega \rightarrow \mathbb{R} \text{ ZV, } Y \in \mathcal{L}^1(P), |X_n| \leq Y \\ \text{f.s. } \forall n \in \mathbb{N}, X_n \rightarrow X \text{ f.s.}$$

$$\Rightarrow X \in \mathcal{L}^1(P), E(X_n|\mathcal{F}) \rightarrow E(X|\mathcal{F}) \text{ f.s.}$$

(c) (Fatou) $X_n, Y : \Omega \rightarrow \overline{\mathbb{R}}$ seien meßbar, $n \in \mathbb{N}$.

$X_n \geq Y$ f.s. $\forall n \in \mathbb{N}, EY^- < \infty$

$\Rightarrow E(\liminf_{n \rightarrow \infty} X_n | \mathcal{F}) \leq \liminf_{n \rightarrow \infty} E(X_n | \mathcal{F})$ f.s.

$X_n \leq Y$ f.s. $\forall n \in \mathbb{N}, EY^+ < \infty$

$\Rightarrow E(\limsup_{n \rightarrow \infty} X_n | \mathcal{F}) \geq \limsup_{n \rightarrow \infty} E(X_n | \mathcal{F})$ f.s. □

Beispiel A.3 $A_i \in \mathcal{A}, i \in I$ seien paarweise disjunkt, I abzählbar, $\bigcup_{i \in I} A_i = \Omega$, $\mathcal{F} = \sigma(A_i, i \in I)$. Dann gilt

$$E(X | \mathcal{F}) = \sum_{i \in I} E(X | A_i) 1_{A_i}$$

mit

$$E(X | A_i) = \begin{cases} \int_{A_i} X dP / P(A_i) & , P(A_i) > 0 \\ 0 \text{ (z.B.)} & , \text{sonst.} \end{cases}$$

$(\Omega_t, \mathcal{A}_t)$ seien meßb. Räume, $Y_t : \Omega \rightarrow \Omega_t$ meßbar, $t \in T$.

$E(X | Y_t, t \in T) := E(X | \sigma(Y_t, t \in T))$ heißt **bedingter Erwartungswert von X unter $(Y_t)_{t \in T}$** .

$(\Omega_2, \mathcal{A}_2)$ sei meßb. Raum, $Y : \Omega \rightarrow \Omega_2$ meßbar.

$E(X | Y = y)$ **bedingter Erwartungswert von X unter $Y = y$** , $y \in \Omega_2$:

P^Y -f.s. eindeutig bestimmte P^Y -quasiintegrierbare $ZV : \Omega_2 \rightarrow \overline{\mathbb{R}}$ mit

$$(3) \quad \int_{A_2} E(X | Y = y) dP^Y(y) = \int_{Y^{-1}(A_2)} X dP \quad \forall A_2 \in \mathcal{A}_2.$$

Es gilt

$$E(X | Y) = E(X | Y = \cdot) \circ Y \quad \text{P-f.s.}$$

Für $X \in \mathcal{L}^1(P)$ heißt

$$\text{Var}(X | \mathcal{F}) := E((X - E(X | \mathcal{F}))^2 | \mathcal{F})$$

bedingte Varianz von X unter \mathcal{F} . Es gilt

$$\text{Var}(X | \mathcal{F}) = E(X^2 | \mathcal{F}) - (E(X | \mathcal{F}))^2$$

falls $X \in \mathcal{L}^2(P)$.

Zu bedingten Verteilungen:

Definition A.4 $(\Omega_i, \mathcal{A}_i)$ seien meßbare Räume, $X : \Omega \rightarrow \Omega_1$ meßbar, $Y : \Omega \rightarrow \Omega_2$ meßbar, $\mathcal{F} \subset \mathcal{A}$.

(a) Ein Markov-Kern $K = K_P$ von (Ω, \mathcal{F}) nach $(\Omega_1, \mathcal{A}_1)$ mit

$$K(\cdot, A_1) = P(X \in A_1 | \mathcal{F}) \text{ P-f.s. } \forall A_1 \in \mathcal{A}_1$$

heißt (eine Version der regulären) **bedingte Verteilung von X unter \mathcal{F}** (bzgl. P). Bezeichnung $P^{X|\mathcal{F}} := K$. RN- Gleichung (2) hat die Form

$$(4) \quad \int_F P^{X|\mathcal{F}}(\omega, A_1) dP(\omega) = P(X^{-1}(A_1) \cap F) \quad \forall A_1 \in \mathcal{A}_1, F \in \mathcal{F}.$$

$P^{X|Y} := P^{X|\sigma(Y)}$ heißt **bedingte Verteilung von X unter Y**.

(b) Ein Markov-Kern $K = K_P$ von $(\Omega_2, \mathcal{A}_2)$ nach $(\Omega_1, \mathcal{A}_1)$ mit

$$K(\cdot, A_1) = P(X \in A_1 | Y = \cdot) \text{ P}^Y\text{-f.s. } \forall A_1 \in \mathcal{A}_1$$

heißt **bedingte Verteilung von X unter Y = y**, $y \in \Omega_2$. Bezeichnung $P^{X|Y=y}(A_1) := K(y, A_1)$. RN- Gleichung (3) hat also die Form

$$(5) \quad \begin{aligned} \int_{A_2} P^{X|Y=y}(A_1) dP^Y(y) &= P(X^{-1}(A_1) \cap Y^{-1}(A_2)) \\ &= P^{(X,Y)}(A_1 \times A_2) \quad \forall A_i \in \mathcal{A}_i. \end{aligned}$$

In Situation wie in A.4 (b) :

Lemma A.5 $Y : \Omega \rightarrow \Omega_2$ sei meßbar, K Markov-Kern von $(\Omega_2, \mathcal{A}_2)$ nach $(\Omega_1, \mathcal{A}_1)$

\Rightarrow

(i) $\Omega_2 \rightarrow \mathbb{R}_+, y \mapsto K(y, D_y)$ ist \mathcal{A}_2 - meßbar $\forall D \in \mathcal{A}_1 \otimes \mathcal{A}_2$.

(ii) $\tilde{K}(y, D) := K(y, D_y)$ def. Markovkern von $(\Omega_2, \mathcal{A}_2)$ nach $(\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \mathcal{A}_2)$.

(iii) $K \otimes P^Y(D) := \int K(y, D_y) dP^Y(y)$, $D \in \mathcal{A}_1 \otimes \mathcal{A}_2$, ist W-Maß auf $\mathcal{A}_1 \otimes \mathcal{A}_2$.

(iv) (Fubini für Markov-Kerne)

$f : \Omega_1 \times \Omega_2 \rightarrow \overline{\mathbb{R}}$ sei $K \otimes P^Y$ - quasiintegrierbar

\Rightarrow

$$\int f dK \otimes P^Y = \int \int f(x, y) K(y, dx) dP^Y(y).$$

Beweis. (i) $\mathcal{D} := \{D \in \mathcal{A}_1 \otimes \mathcal{A}_2 : y \mapsto K(y, D_y) \text{ } \mathcal{A}_2\text{- meßb.}\}$ ist ein Dynkin-System. Für $\mathcal{E} = \{A_1 \times A_2 : A_i \in \mathcal{A}_i\}$ gilt $\mathcal{E} \subset \mathcal{D}$, denn

$$K(y, (A_1 \times A_2)_y) = K(y, A_1) 1_{A_2}(y).$$

Ferner ist \mathcal{E} ein \cap - stabiler Erzeuger von $\mathcal{A}_1 \otimes \mathcal{A}_2$. Es folgt

$$\mathcal{A}_1 \otimes \mathcal{A}_2 = \sigma(\mathcal{E}) = \delta(\mathcal{E}) \subset \mathcal{D} \subset \mathcal{A}_1 \otimes \mathcal{A}_2,$$

wobei $\delta(\mathcal{E})$ das von \mathcal{E} erzeugte Dynkin - System bezeichnet.

(ii) und (iii) sind klar.

(iv) Für $f = 1_D$, $D \in \mathcal{A}_1 \otimes \mathcal{A}_2$, ist das die Definition von $K \otimes P^Y$. Rest folgt über Aufbau der $\mathcal{A}_1 \otimes \mathcal{A}_2$ -meßbaren Funktionen. \square

Beispiel A.6 $X : \Omega \rightarrow \Omega_1$ und $Y : \Omega \rightarrow \Omega_2$ seien meßbar.

- (a) X, Y unabhängig $\Rightarrow P^{X|Y=y} := P^X$ ist eine bedingte Verteilung von X unter $Y = y$, denn

$$\begin{aligned} \int_{A_2} P^X(A_1) dP^Y(y) &= P^X(A_1) P^Y(A_2) \\ &= P^X \otimes P^Y(A_1 \times A_2) \\ &= P^{(X,Y)}(A_1 \times A_2), A_i \in \mathcal{A}_i. \end{aligned}$$

Also gilt die definierende RN-Gleichung (5). (Die Umkehrung gilt auch.)

- (b) Es existiere die bedingte Verteilung $P^{X|Y=y}, h : (\Omega_1, \mathcal{A}_1) \rightarrow (\Omega_3, \mathcal{A}_3)$ sei meßbar \Rightarrow

$$P^{h(X)|Y=y} := (P^{X|Y=y})^h \text{ (Bildmaß)}$$

ist eine bedingte Verteilung von $h(X)$ unter $Y = y$. Dazu sei $K(y, A_3) := P^{X|Y=y}(h^{-1}(A_3)), y \in \Omega_2$. Dann gilt: K ist Markov-Kern von $(\Omega_2, \mathcal{A}_2)$ nach $(\Omega_3, \mathcal{A}_3)$ und

$$\begin{aligned} \int_{A_2} K(y, A_3) dP^Y(y) &= \int_{A_2} P^{X|Y=y}(h^{-1}(A_3)) dP^Y(y) \\ &= \\ (5) \quad &P^{(X,Y)}(h^{-1}(A_3) \times A_2) \\ &= P^{(h(X),Y)}(A_3 \times A_2) \end{aligned}$$

\Rightarrow Beh.

Falls bedingte Verteilungen existieren, lassen sich bedingte Erwartungswerte als Erwartungswerte bzgl. bedingter Verteilungen berechnen.

Satz A.7 $X : \Omega \rightarrow \Omega_1$ und $Y : \Omega \rightarrow \Omega_2$ seien meßbar, $P^{X|Y=y}$ existiere, $f : \Omega_1 \times \Omega_2 \rightarrow \overline{\mathbb{R}}$ sei $P^{(X,Y)}$ -quasiintegrierbar \Rightarrow

- (i) $P^{(X,Y)} = P^{X|Y=\cdot} \otimes P^Y$ (s.A.5)
- (ii) $E(f(X, Y)|Y = y) = \int f(x, y) P^{X|Y=y}(dx)$ P^Y -f.s
- (iii) $E f(X, Y) = \int \int f(x, y) P^{X|Y=y}(dx) dP^Y(y)$

Beweis. (i) Nach A.5 (iii) ist

$$Q := P^{X|Y=\cdot} \otimes P^Y$$

ein W-Maß auf $\mathcal{A}_1 \otimes \mathcal{A}_2$. Wegen (5) gilt

$$\begin{aligned} Q(A_1 \times A_2) &= \int P^{X|Y=y}(A_1) 1_{A_2}(y) dP^Y(y) \\ &= P^{(X,Y)}(A_1 \times A_2), A_i \in \mathcal{A}_i, \end{aligned}$$

also $Q|_{\mathcal{E}} = P^{(X,Y)}|_{\mathcal{E}}$ für den \cap -stabilen Erzeuger $\mathcal{E} = \{A_1 \times A_2 : A_i \in \mathcal{A}_i\}$ von $\mathcal{A}_1 \otimes \mathcal{A}_2$. Es folgt $Q = P^{(X,Y)}$.

- (ii) Die rechte Seite ist \mathcal{A}_2 -meßbar in y . Ferner gilt mit Teil (i) und A.5 (iv)

$$\begin{aligned} &\int_{A_2} \int f(x, y) P^{X|Y=y}(dx) dP^Y(y) \\ &= \int \int f 1_{\Omega_1 \times A_2} dP^{X|Y=\cdot} \otimes P^Y \\ &= \int \int f 1_{\Omega_1 \times A_2} dP^{(X,Y)} \\ &= \int_{Y^{-1}(A_2)} f(X, Y) dP \quad \forall A_2 \in \mathcal{A}_2. \end{aligned}$$

Dies ist definierende RN-Gleichung (3).

(iii) Wähle $A_2 = \Omega_2$ im Beweis von (ii). □

Korollar A.8 $X : \Omega \rightarrow \Omega_1$ und $Y : \Omega \rightarrow \Omega_2$ seien meßbar.

(a) $P^{X|Y=y}$ existiere, $h : \Omega_1 \rightarrow \overline{\mathbb{R}}$ sei P^X -quasiintegrierbar

$$\Rightarrow E(h(X)|Y=y) = \int h(x)P^{X|Y=y}(dx) \text{ P}^Y\text{-f.s.}$$

($\mathcal{F} \subset \mathcal{A}$, $P^{X|\mathcal{F}}$ existiere

$$\Rightarrow E(h(X)|\mathcal{F})(\omega) = \int h(x)P^{X|\mathcal{F}}(\omega, dx) \text{ P-f.s.})$$

(b) (Substitutionsregel bei Unabhängigkeit)

X, Y seien unabhängig, $f : \Omega_1 \times \Omega_2 \rightarrow \overline{\mathbb{R}}$ sei $P^{(X,Y)}$ -quasiintegrierbar

$$\begin{aligned} \Rightarrow E(f(X, Y)|Y=y) &= \int f(x, y)dP^X(x) \\ &= E f(X, y) \text{ P}^Y\text{-f.s.} \end{aligned}$$

Beweis. (a) Setze $f(x, y) = h(x)$ in A.7 (ii).

(b) Nach A.6 (a) ist P^X eine bedingte Verteilung von X unter $Y = y$. Die Beh. folgt damit aus A.7 (ii). □

A.8 (b) kann man verschärfen: Y sei \mathcal{F} -meßbar, $\mathcal{F} \subset \mathcal{A}$, X und \mathcal{F} seien unabhängig

$$\begin{aligned} \Rightarrow E(f(X, Y) | \mathcal{F}) &= E(f(X, Y) | Y) \\ &= \int f(x, Y)dP^X(x) \text{ P-f.s.} \end{aligned}$$

(s. Hoffmann-Jørgensen, Vol. I, S. 452). A 8 (b) entspricht $\mathcal{F} = \sigma(Y)$.

Satz A.9 (Existenz) $X : \Omega \rightarrow \Omega_1$ sei meßbar, Ω_1 polnisch, $\mathcal{A}_1 = \mathcal{B}(\Omega_1)$.

(a) $Y : \Omega \rightarrow \Omega_2$ sei meßbar

\Rightarrow es ex. bedingte Verteilung $P^{X|Y=y}$.

(b) $\mathcal{F} \subset \mathcal{A}$

\Rightarrow es ex. bedingte Verteilung $P^{X|\mathcal{F}}$.

Beweis. Bauer, S. 397, Gänsler/Stute, S. 196. □

Jede offene und jede abgeschlossene Teilmenge von \mathbb{R}^n ist polnisch.

Bemerkung A.10 (Eindeutigkeit) $X : \Omega \rightarrow \Omega_1$ und $Y : \Omega \rightarrow \Omega_2$ seien meßbar, K_1, K_2 seien bedingte Verteilungen von X unter $Y = y$, \mathcal{A}_1 sei abzählbar erzeugt $\Rightarrow \exists N \in \mathcal{A}_2, P^Y(N) = 0$ mit $K_1(y, A_1) = K_2(y, A_1) \forall y \notin N \forall A_1 \in \mathcal{A}_1$.

Beweis. Bauer, S. 397, Gänsler/ Stute, S.197. □

$X : \Omega \rightarrow \Omega_1$ und $Y : \Omega \rightarrow \Omega_2$ seien meßbar. Interessant für die Berechnung der bedingten Verteilung ist der Fall, dass $P^{(X,Y)}$ durch ein Produktmaß dominiert wird,

$$P^{(X,Y)} \ll \mu_1 \otimes \mu_2, f^{(X,Y)} = dP^{(X,Y)} / d\mu_1 \otimes \mu_2.$$

Dabei ist μ_i ein σ -endliches Maß auf \mathcal{A}_i . Die μ_2 - Dichte f^Y von P^Y lässt sich als Randdichte berechnen,

$$f^Y(y) = \int f^{(X,Y)}(x,y) d\mu_1(x).$$

Sei

$$f^{X|Y=y}(x) := \begin{cases} \frac{f^{(X,Y)}(x,y)}{f^Y(y)} & , f^Y(y) > 0 \\ g(x) & , \text{sonst} \end{cases}$$

die **bedingte Dichte von X unter Y = y**. (g ist eine beliebige W-Dichte bzgl. μ_1 .)

Satz A.11 In obiger Situation sei

$$K(y, A_1) := \int_{A_1} f^{X|Y=y}(x) d\mu_1(x), y \in \Omega_2, A_1 \in \mathcal{A}_1$$

$\Rightarrow K$ ist eine bedingte Verteilung von X unter $Y = y$.

Beweis. K ist offenbar ein Markov-Kern von $(\Omega_2, \mathcal{A}_2)$ nach $(\Omega_1, \mathcal{A}_1)$. Es gilt

$$P^Y(f^Y = 0) = \int_{\{f^Y=0\}} f^Y d\mu_2 = 0.$$

Ferner

$$\begin{aligned} \int_{A_2} K(y, A_1) dP^Y(y) &= \int_{A_2 \cap \{f^Y > 0\}} K(y, A_1) dP^Y(y) \\ &= \int_{A_2 \cap \{f^Y > 0\}} \int_{A_1} f^{X|Y=y}(x) d\mu_1(x) f^Y(y) d\mu_2(y) \\ &= \int_{A_2 \cap \{f^Y > 0\}} \int_{A_1} f^{(X,Y)}(x,y) d\mu_1(x) d\mu_2(y) \\ &= P^{(X,Y)}(A_1 \times A_2) \forall A_i \in \mathcal{A}_i. \end{aligned}$$

Also gilt die RN- Gleichung (5). □