Kapitel 8

Produktmaße und Summen unabhängiger Zufallsgrößen

Die Wahrscheinlichkeitsverteilungen voneinander unabhängiger Zufallsgrößen sind Produktmaße. In diesem Kapitel geht es zunächst um den Zusammenhang zwischen Integralen über Funktionen bez. Produktmaßen und sogenannten iterierten Integralen über diese Funktionen bez. der Einzelmaße. Der diesbezügliche Satz von Fubini mit samt seiner Folgerung, die nach Tonelli und Hobson benannt ist, bildet ein in der Wahrscheinlichkeitstheorie oft benutztes Werkzeug.
Im zweiten Abschnitt wird dann die Verteilung der Summen unabhängiger Zufallsgrößen untersucht. Sie ergibt sich als sogenannte Faltung der Verteilungen der einzelnen Zufallsgrößen.

8.1 Der Satz von Fubini

Es seien X und Y zwei Zufallsgrößen über $(\Omega, \mathfrak{A}, P)$ mit Werten in $(E, \mathfrak{E})$ bzw. $(F, \mathfrak{F})$. Der zufällige Vektor $(X, Y)^T$ ist eine Zufallsgröße mit Werten in $(E \times F, \mathfrak{E} \otimes \mathfrak{F})$, wobei $\mathfrak{E} \otimes \mathfrak{F}$ die Produkt-σ-Algebra von $\mathfrak{E}$ und $\mathfrak{F}$ in $E \times F$ ist, d. h. die kleinste σ-Algebra von Teilmengen von $E \times F$, die alle "Rechtecke" $B \times C$ mit $B \in \mathfrak{E}$ und $C \in \mathfrak{F}$ umfasst.

Sind X und Y voneinander unabhängig, so gilt für die Verteilung $P^{(X,Y)}$ von $(X, Y)^T$ die Beziehung

$$P^{(X,Y)}(B \times C) = P(X \in B, Y \in C) = P(X \in B)P(Y \in C) = P^X(B)P^Y(C),\ B \in \mathfrak{E}, C \in \mathfrak{F}. \tag{8.1}$$

Die gemeinsame Verteilung $P^{(X,Y)}$ von $(X,Y)^T$ ist also das *Produktmaß* der Randverteilungen P^X und P^Y. Ist h eine Borel-messbare Funktion von $(E \times F, \mathfrak{E} \otimes \mathfrak{F})$ in $(R_1, \mathscr{B}_1)$, so haben wir gemäß der Substitutionsformel (Aussage 7.14)

$$Eh(X,Y) = \int\limits_{E \times F} h(x,y) P^{(X,Y)}(dx, dy). \tag{8.2}$$

Der folgende Satz von Fubini liefert Bedingungen, unter denen man dieses Integral im Fall der Unabhängigkeit von X und Y auf Einzelintegrale bez. P^X bzw. P^Y zurückführen kann, die sich z. B. im Falle der Existenz von Dichten wiederum einfacher berechnen lassen. Für den Beweis dieses Satzes und den Aussagen dieses Abschnittes siehe z.B. Elstrodt (1999), Bauer (1990) bzw. die Vorlesung Maßtheorie.

Anstelle von P^X und P^Y verwenden wir im Folgenden irgend zwei Wahrscheinlichkeitsmaße Q_1 und Q_2 auf $(E, \mathfrak{E})$ bzw. $(F, \mathfrak{F})$.

Wir beginnen mit einer Aussage über Existenz und Eindeutigkeit des Produktmaßes $Q_1 \otimes Q_2$.

Aussage 8.1 *Es seien Q_1 und Q_2 zwei Wahrscheinlichkeitsmaße auf $(E, \mathfrak{E})$ bzw. $(F, \mathfrak{F})$. Dann besitzt die durch*

$$R(A \times B) := Q_1(A) Q_2(B) \tag{8.3}$$

auf der Menge $\mathscr{R}$ aller Rechtecke $A \times B$ mit $A \in \mathfrak{E}, B \in \mathfrak{F}$ definierte Mengenfunktion R eine eindeutige Fortsetzung zu einem Wahrscheinlichkeitsmaß auf $\mathfrak{E} \otimes \mathfrak{F}$. Diese Fortsetzung wird als Produktmaß $Q_1 \otimes Q_2$ von Q_1 und Q_2 bezeichnet.

Bemerkung: Die Aussage bleibt richtig, falls Q_1 und Q_2 keine Wahrscheinlichkeitsmaße, sondern sogenannte σ-finite Maße sind. Das Produktmaß $Q_1 \otimes Q_2$ ist dann ebenfalls ein σ-finites Maß. Das gilt insbesondere für die Lebesguemaße in R_1 bzw. R_2 und allgemeiner, in R_n.

Satz 8.2 (Satz von Fubini) *Es sei h eine reellwertige, $(\mathfrak{E}\otimes\mathfrak{F})-\mathscr{B}_1$-messbare Funktion, die nichtnegativ ist oder für die*

$$\int\limits_{E\times F} |h(x,y)| Q_1 \otimes Q_2(dx,dy) < \infty \tag{8.4}$$

richtig ist.

Dann gelten folgende zwei Aussagen:

1) *Die Funktion $x \to \int\limits_F h(x,y)Q_2(dy)$ ist $\mathfrak{E} - \mathscr{B}_1$-messbar. Sie ist nichtnegativ, falls h nichtnegativ ist, und sie ist Q_1-integrierbar, falls (8.4) gilt.*

   Die Funktion $y \to \int\limits_E h(x,y)Q_1(dx)$ ist $\mathfrak{F}$-$\mathscr{B}_1$-messbar. Sie ist nichtnegativ, falls h nichtnegativ ist, und sie ist Q_2-integrierbar, falls (8.4) gilt.

2) *Ist h nichtnegativ oder gilt (8.4), so haben wir*

$$\int\limits_{E\times F} h(x,y)Q_1 \otimes Q_2(dx,dy) = \int\limits_E \Big(\int\limits_F h(x,y)Q_2(dy)\Big) Q_1(dx)$$

$$= \int\limits_F \Big(\int\limits_E h(x,y)Q_1(dx)\Big) Q_2(dy) \tag{8.5}$$

Bemerkung: Der Satz von Fubini gilt auch für Maße Q_1, Q_2, die nicht notwendig Wahrscheinlichkeitsmaße sind. Insbesondere für Lebesguemaße.

Der Satz von Fubini besagt also, dass aus der Integrierbarkeit von h bezüglich des Produktmaßes $Q_1 \otimes Q_2$ folgt, dass die iterierten Integrale in (8.5) existieren und gleich dem Integral bez. dem Produktmaß sind.

Im allgemeinen folgt aus der Endlichkeit der iterierten Integrale, selbst wenn sie gleich sind, noch nicht die Eigenschaft (8.5). (Vgl. Elstrodt (1996), Kap. V §2).
Es gilt aber die

Folgerung 8.3 (Tonelli, Hobson) *Wenn gilt*

$$\int_E \Bigg(\int_F |h(x,y)| Q_2(dy) \Bigg) Q_1(dx) < \infty \qquad \textit{oder} \tag{8.6}$$

$$\int_F \Bigg(\int_E |h(x,y)| Q_1(dx) \Bigg) Q_2(dy) < \infty, \tag{8.7}$$

dann sind die Voraussetzungen des Satzes von Fubini erfüllt.

Beweis: Für jede Indikatorfunktion $h = \mathbb{1}_C$ mit $C \in \mathfrak{E} \otimes \mathfrak{F}$ gilt der Satz von Fubini, d. h. $x \to Q_2(C(x))$ ist $\mathfrak{E}$-messbar und es besteht die Beziehung

$$\int_{E \times F} h dQ_1 \otimes Q_2 = \int_E Q_2(C(x)) Q_1(dx) < \infty$$

wobei $C(x)$ die sogenannte "Schnittmenge" von C bei x ist:

$$C(x) := \{y \in F | (x,y) \in C\}, \quad x \in E.$$

Damit ist auch für jede nichtnegative Elementarfunktion h

$$\int_{E \times F} h dQ_1 \otimes Q_2 = \int_E \Bigg(\int_F h dQ_2 \Bigg) dQ_1 < \infty.$$

Ist nun h nichtnegativ mit der Eigenschaft (8.6) und (h_n) eine h approximierende Folge von nichtnegativen Elementarfunktionen, so ergibt sich aus dem Satz über die monotone Konvergenz, dass

$$\int\limits_{E\times F} h dQ_1 \otimes Q_2 = \lim_{n\to\infty} \int\limits_{E\times F} h_n dQ_1 \otimes Q_2 =$$

$$= \lim_{n\to\infty} \int\limits_E \left(\int\limits_F h_n dQ_2 \right) dQ_1$$

richtig ist.

Wegen

$$\int\limits_F h_n dQ_2 \uparrow \int\limits_F h dQ_2$$

(Satz über monotone Konvergenz) erhalten wir wegen (8.6) die Eigenschaft (8.4).

Im allgemeinen Fall nutzen wir die Zerlegung $h = h^+ - h^-$.

Der Fall (7) wird analog behandelt. □

8.2 Faltungsformeln

Es seien X und Y zwei voneinander unabhängige reellwertige Zufallsgrößen über $(\Omega, \mathfrak{A}, P)$ und $Z := X + Y$.

Für die Wahrscheinlichkeitsverteilung P^Z der reellwertigen Zufallsgröße Z gilt wegen der Substitutionsformel (7.13), angewandt auf $h(X,Y) = \mathbb{1}_B(X+Y)$,

$$P^Z(B) = P(X + Y \in B) = E\mathbb{1}_B(X+Y) =$$

$$\int\limits_\Omega \mathbb{1}_B(X+Y)dP = \int\limits_{R_2} \mathbb{1}_B(x+y)P^X \otimes P^Y(dx,dy)$$

Auf Grund des Satzes von Fubini ist dieser Wert gleich

$$\int_{R_1} P^Y(\{y : x + y \in B\})P^X(dx).$$

Für $B = (-\infty, z]$ ergibt sich

$$F_Z(z) = P^Z((-\infty, z]) = P(Z \le z) =$$

$$\int_{R_1} P^Y((-\infty, z - x])P^X(dx) =$$

$$\int_{R_1} F_Y(z - x)P^X(dx). \tag{8.8}$$

Bemerkung: Da die Verteilungsfunktion F_X die Wahrscheinlichkeitsverteilung P^X eindeutig bestimmt, schreibt man häufig statt

$$\int_{R_1} f(x)P^X(dx) \text{ auch } \int_{R_1} f(x)F_X(dx).$$

Wir haben dann also anstelle (8.8) die Gleichung

$$F_Z(z) = \int_{R_1} F_Y(z-x)F_X(dx) \tag{8.8$'$}$$

Definition 8.4 *Die durch die rechte Seite von (8.8') aus F_X und F_Y gebildete Verteilungsfunktion F_Z bezeichnet man als Faltung der beiden Verteilungsfunktionen F_X und F_Y und schreibt $F_Z = F_X * F_Y$.*

Offensichtlich gilt $F_X * F_Y = F_Y * F_X$.

Spezialfälle 8.5

a) Sind X und Y unabhängig und diskret verteilt mit Werten aus den ganzen Zahlen, gilt also

$$P(X = k) = p_k \ , \ P(Y = k) = q_k \ , \ k \in \Gamma := \{0, \pm 1, \pm 2, \ldots\}$$

mit $p_k, q_k \geq 0$ und $\sum p_k = \sum q_k = 1$, so erhalten wir

$$P(Z = k) = \sum_{l \in \Gamma} q_{k-l} p_l = \sum_{l \in \Gamma} p_{k-l} q_l, \ k \in \Gamma. \tag{8.9}$$

Beweis: Man verwende

$$P(Z = k) = F_Z(k) - F_Z(k - \frac{1}{2}) \text{und} \int\limits_{R_1} h(x) F_X(dx) = \sum_k h(k) p_k$$

bzw. $\int\limits_{R_1} h(x) F_Y(dx) = \sum_k h(k) q_k$ für alle nichtnegativen Funktionen h auf der Menge der ganzen Zahlen. □

b) Sind X und Y unabhängig und haben sie Dichten f_X bzw. f_Y, so hat auch $Z = X + Y$ eine Dichte f_Z, und es gilt

$$f_Z(z) = \int\limits_{R_1} f_Y(z - x) f_X(x) dx = \int\limits_{R_1} f_X(z - y) f_Y(y) dy. \tag{8.10}$$

Beweis: Mittels der Substitutionsformel (7.13) folgt aus (8.8) die Beziehung

$$F_Z(z) = \int\limits_{R_1} \Bigg(\int\limits_{(-\infty, z-x]} f_Y(y) dy \Bigg) f_X(x) dx =$$

$$\int_{R_1} \left(\int_{(-\infty,z]} f_Y(y-x)dy \right) f_X(x)dx$$

und der Satz von Fubini mit seiner Folgerung von Tonelli-Hobson ergibt damit

$$F_Z(z) = \int_{(-\infty,z]} \left(\int_R f_Y(y-x) f_X(x)dx \right) dy.$$

Daraus folgt nach Definition der Dichten die Behauptung. □

Bemerkung 8.6 Bereits wenn nur eine der beiden unabhängigen Zufallsgrößen X und Y eine Dichte besitzt, hat auch $Z = X+Y$ eine Dichte. (Übung)

Beispiele 8.7 Es seien X und Y unabhängige Zufallsgrößen.

1. Sind X und Y binomialverteilt mit den Parametern (n, p) bzw. (m, p), so ist $X + Y$ binomialverteilt mit den Parametern $(n + m, p)$.

2. Sind X und Y Poissonverteilt mit den Parametern λ bzw. μ, so ist $X+Y$ Poissonverteilt mit dem Parameter $\lambda + \mu$.

3. Sind X und Y negativ binomialverteilt mit den Parametern (p, v_X) bzw. (p, v_Y), so ist $X+Y$ negativ binomialverteilt mit den Parametern (p, v_X+v_Y).

4. Sind X und Y normalverteilt mit den Parametern (μ_X, σ_X^2) bzw. (μ_Y, σ_Y^2), so ist $X + Y$ normalverteilt mit den Parametern $(\mu_X + \mu_Y, \sigma_X^2 + \sigma_Y^2)$.

5. Sind X und Y Gammaverteilt mit den Parametern (α_X, λ) bzw. (α_Y, λ), so ist $X + Y$ Gammaverteilt mit den Parametern $(\alpha_X + \alpha_Y, \lambda)$.

Der Beweis von 1. - 3. ergibt sich sofort aus der Formel

$$g_{X+Y}(s) = Es^{X+Y} =$$

$$Es^X \cdot Es^Y = g_X(s) \cdot g_Y(s), |s| < 1$$

für die erzeugenden Funktionen.

Für den Beweis von 4. und 5. verwendet man die sogenannten charakteristischen Funktionen, die wir später definieren werden.

Die Faltungsformeln (8.9) und (8.10) führen auch zum Ziel, sind aber häufig mit längeren Rechnungen verbunden.

Definition 8.8 *Eine Familie* $(F_\vartheta, \vartheta \in \Theta \subseteq R_k)$ *von Verteilungsfunktionen auf* R_1 *heißt faltungsstabil, falls für alle* $\vartheta, \eta \in \Theta$ *ein* $\xi \in \Theta$ *existiert mit* $F_\vartheta * F_\eta = F_\xi$.

Die Familie aller eindimensionalen Normalverteilungen ist zum Beispiel faltungsstabil.

Wir schließen dieses Kapitel mit einigen Eigenschaften von Erwartungswerten und Varianzen unabhängiger Zufallsgrößen.

Aussage 8.9 *Sind* X *und* Y *zwei voneinander unabhängige reellwertige Zufallsgrößen über einem Wahrscheinlichkeitsraum* $(\Omega, \mathfrak{A}, P)$ *und gilt* $E|X| < \infty$, $E|Y| < \infty$, *so folgt*

$$E|XY| < \infty \textit{ und } E(XY) = (EX)(EY). \tag{8.11}$$

Beweis: Wegen der Unabhängigkeit ist die gemeinsame Verteilung von X und Y gleich dem Produktmaß $P^X \otimes P^Y$. Der Satz von Fubini (Folgerung 8.3) impliziert

$$E(XY) = \int\limits_{R_2} xyP^X \otimes P^Y(dx, dy) =$$

$$\int\limits_{R_1}\Big(\int\limits_{R_1} xyP^X(dx)\Big)P^Y(dy) = \int\limits_{R_1} y\Big(\int\limits_{R_1} xP^X(dx)\Big)P^Y(dy) = (EX)(EY).$$

Folgerung 8.10 Sind $X_1, X_2, \ldots, X_n$ voneinander unabhängige reellwertige Zufallsgrößen über einem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, P)$ und gilt $E|X_k| < \infty$, $k = 1, \ldots, n$, so folgt

$$E|X_1 \cdot X_2 \cdot \ldots \cdot X_n| < \infty \text{ und}$$

$$E\Big(\prod_{k=1}^n X_k\Big) = \prod_{k=1}^n EX_k.$$

Die Gleichung (8.11) hat zur Konsequenz, dass für zwei unabhängige Zufallsgrößen X und Y, deren Erwartungswerte endlich sind, die Kovarianz auch endlich und gleich Null ist.

$$Kov(X, Y) = E(XY) - EXEY = 0.$$

Unabhängige Zufallsgrößen mit endlichem Erwartungswert sind also unkorreliert.
Daraus ergibt sich die

Aussage 8.11 *Sind $X_1, X_2, \ldots, X_n$ voneinander unabhängige Zufallsgrößen mit $D^2X_k < \infty, k = 1, \ldots, n$, so gilt*

$$D^2\Big(\sum_{k=1}^n X_k\Big) = \sum_{k=1}^n D^2X_k \tag{8.12}$$

Beweis:

$$D^2\Big(\sum_{k=1}^n X_k\Big) = E\Big(\sum_{k=1}^n (X_k - EX_k)\Big)^2 =$$

$$\sum_{k,\ell=1}^{n} E(X_k - EX_k)(X_\ell - EX_\ell) =$$

$$\sum_{k=1}^{n} D^2 X_k + 2\sum_{k<\ell} Kov(X_k, X_\ell) = \sum_{k=1}^{n} D^2 X_k.$$

Folgerung 8.12 Unter den gleichen Voraussetzungen wie in der eben bewiesenen Aussage gilt für die Varianz des arithmetischen Mittels:

$$D^2\Big(\frac{1}{n}\sum_{k=1}^{n} X_k\Big) = \frac{1}{n^2}\sum_{k=1}^{n} D^2 X_k \leq \frac{1}{n} \max_{k=1,\ldots,n} D^2 X_k.$$

Beispiel: Wirft man einen regelmäßigen Spielwürfel unabhängig voneinander 100mal und bildet das arithmetische Mittel M_{100} der auftretenden Augenzahlen, so gilt

$$EM_{100} = 3,5 \text{ und } D^2 M_{100} = \frac{2,9167}{100} = 0,0292.$$

Das arithmetische Mittel wird bei häufiger Wiederholung von 100 Würfen also meist in der Nähe von $3,5$ liegen. Die Tschebyschev'sche Ungleichung liefert nämlich

$$P(|M_{100} - 3,5| \geq 0,5) \leq \frac{0,0292}{0,25} = 0,12.$$

Das arithmetische Mittel M_{100} ist also "weniger zufällig" als jeder einzelne Wurf des Würfels.

Index