Anhang zu Kapitel 5: n-dimensionale Zufallsvariablen

Häufig reicht die Betrachtung 2-dimensionaler Zufallsvariablen nicht aus. Zum Beispiel bildet eine *Stichprobe der Größe n* in der (mathematischen) Statistik eine n-dimensionale Zufallsvariable, wobei $n \in \mathbb{N}$ ist. Um im nächsten Kapitel Grenzwertsätze der Wahrscheinlichkeitstheorie betrachten zu können, müssen wir ebenfalls $n \in \mathbb{N}$ zulassen und dann n gegen ∞ gehen lassen. In diesem Anhang sollen kurz die wesentlichen Definitionen, Konzepte und Aussagen über mehrdimensionale Zufallsvariablen zusammengestellt werden. Da er vollkommen analog zu der 2-dimensionalen Situation ist, wird der Stoff dieses Anhangs nicht in der Vorlesung vorgetragen, dient aber als Grundlage für Kapitel 6 und 7 der Vorlesung. Sei $(\Omega, \mathcal{F}, P)$ stets ein Wahrscheinlichkeitsraum.

Definition 0.1. Sei $n \in \mathbb{N}$ und seien $X_1, \ldots, X_n : \Omega \to \mathbb{R}$ reelle (diskrete, stetige oder weder noch) Zufallsvariablen. Dann heißt der Vektor $X := (X_1, \ldots, X_n) : \Omega \to \mathbb{R}^n$ *n-dimensionale Zufallsvariable* oder auch *n-dimensionaler Zufallsvektor*.

Definition 0.2. Eine n-dimensionale Zufallsvariable $X := (X_1, \ldots, X_n) : \Omega \to \mathbb{R}^n$ heißt *diskret*, wenn $X_1, \ldots, X_n$ diskrete reelle Zufallsvariablen sind, d.h. wenn ihre Wertebereiche $S_1 := X_1(\Omega), \ldots, S_n := X_n(\Omega)$ abzählbar sind. In diesem Fall heißt die Abbildung $P_X := P_{(X_1,\ldots,X_n)} : \mathcal{P}(S_1 \times \ldots \times S_n) \to [0,1]$ mit

$$P_{(X_1,\ldots,X_n)}(B) := P\big((X_1, \ldots, X_n) \in B\big) := P\big((X_1, \ldots, X_n)^{-1}(B)\big)$$

die *gemeinsame Verteilung* von $X_1, \ldots, X_n$. Weiter heißt die Funktion
$\varrho_X := \varrho_{(X_1,\ldots,X_n)} : S_1 \times \ldots \times S_n \to \mathbb{R}$ mit

$$\varrho_{(X_1,\ldots,X_n)}(x_1, \ldots, x_n) := P(X_1 = x_1, \ldots, X_n = x_n)$$

die *gemeinsame Zähldichte* von $X_1, \ldots, X_n$.

Bemerkung 0.3. Wie für 2-dimensionale diskrete Zufallsvariablen in Bemerkung 5.3 erklärt, lässt sich auch im n-dimensionalen Fall die Zähldichte $\varrho_{(X_1,\ldots,X_n)}$ zu einer Funktion auf $\mathbb{R}^n$ fortsetzen, indem man $\varrho_{(X_1,\ldots,X_n)}(x_1, \ldots, x_n) := 0$ setzt für $(x_1, \ldots, x_n) \in \mathbb{R}^n \setminus S_1 \times \ldots \times S_n$ und die gemeinsame Verteilung $P_{(X_1,\ldots,X_n)}$ lässt sich auf $\mathcal{P}(\mathbb{R}^n)$ fortsetzen. Dann gilt für jede Teilmenge $B \subseteq \mathbb{R}^n$:

$$P_{(X_1,\ldots,X_n)}(B) := \sum_{(x_1,\ldots,x_n)\in B} \varrho_{(X_1,\ldots,X_n)}(x_1, \ldots, x_n)$$

Insbesondere ist die gemeinsame Verteilung also durch die gemeinsame Zähldichte vollständig beschrieben.

Definition 0.4. Eine n-dimensionale Zufallsvariable $X = (X_1, \ldots, X_n) : \Omega \to \mathbb{R}^n$ heißt *stetig*, wenn es eine integrierbare Funktion $f : \mathbb{R}^n \to \mathbb{R}$ mit folgenden Eigenschaften gibt:

(a) f ist nichtnegativ, d.h. $f(x_1, \ldots, x_n) \geq 0$ für alle $(x_1, \ldots, x_n) \in \mathbb{R}^n$

(b) Für alle Intervalle $I_1, \ldots, I_n \subseteq \mathbb{R}$ gilt:

$$\begin{aligned} P\big((X_1, \ldots, X_n) \in I_1 \times \ldots \times I_n\big) &= P\big(X \in I_1, \ldots, X_n \in I_n\big) \\ &= \int_{I_1} \ldots \int_{I_n} f(x_1, \ldots, x_n) dx_n \ldots dx_1 \end{aligned}$$

(Nach dem Satz von Fubini liefern alle $n!$ Integrationsreihenfolgen dasselbe Ergebnis!)

In diesem Fall heißt f eine *gemeinsame Wahrscheinlichkeitsdichte* oder auch *gemeinsame Dichte* von $X_1, \ldots, X_n$ und wird häufig mit f_X oder $f_{(X_1,\ldots,X_n)}$ bezeichnet. Allgemeiner heißt jede integrierbare Funktion $f : \mathbb{R}^n \to \mathbb{R}$ mit den Eigenschaften (a) und $\int_{\mathbb{R}^n} f(x_1, \ldots, x_n) dx_n \ldots dx_1 = 1$ eine *n-dimensionale Wahrscheinlichkeitsdichte.*

Bemerkung 0.5. Die Aussagen von Bemerkung 4.21 gelten entsprechend auch im n-dimensionalen Fall. Wir gehen hier nur auf Teil (d) (bzw. (iv)?) ein: Ist $X = (X_1, \ldots, X_n)$ eine n-dimensionale stetige Zufallsvariable mit gemeinsamer Wahrscheinlichkeitsdichte f, so gilt für alle („schönen“, d.h. *Borelschen*) Telmengen $B \subseteq \mathbb{R}^n$, dass

$$\begin{aligned} P\big((X_1, \ldots, X_n) \in B\big) &= \int_B f(x_1, \ldots, x_n) dx_n \ldots dx_1 \\ &:= \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} 1_B(x_1, \ldots, x_n) f(x_1, \ldots, x_n) dx_n \ldots dx_1 \,. \end{aligned}$$

Die Zuordnung $B \mapsto P_{(X_1,\ldots,X_n)}(B) := P((X_1, \ldots, X_n) \in B)$ nennt man die *gemeinsame Verteilung* von $X_1, \ldots, X_n$.

Bemerkung 0.6. Auch für n-dimensionale Zufallsvariablen führt man den Begriff der *Rand- bzw. Marginalverteilungen* ein. Hier nennt man für jedes $k \in \{1, \ldots, n\}$ und alle $1 \leq i_1 < i_2 < \ldots < i_k \leq n$, die gemeinsame Verteilung von $(X_{i_1}, X_{i_2}, \ldots, X_{i_n})$ eine *k-te Randverteilung* von $P_{(X_1,\ldots,X_n)}$. Für diskrete bzw. stetige n-dimensionale Zufallsvariablen gelten zu den Sätzen 5.4 und 5.11 analoge Aussagen, dass man die Zähl- bzw. Wahrscheinlichkeitsdichte der Randverteilung von $(X_{i_1}, X_{i_2}, \ldots, X_{i_n})$ erhält, indem man über die zu den komplementären Indizes gehörigen Variablen, also die aus $\{1, \ldots, n\} \setminus \{i_1, \ldots, i_k\}$, summiert bzw. integriert. Da wir diese Aussage im Folgenden nicht benötigen werden, soll dies hier nicht präzisiert werden.

Definition 0.7. Sei $X := (X_1, \dots, X_n) : \Omega \to \mathbb{R}^n$ eine n-dimensionale Zufallsvariable. Dann heißt die Funktion $F_X := F_{(X_1,\dots,X_n)} : \mathbb{R}^n \to \mathbb{R}$ mit $F_{(X_1,\dots,X_n)}(t_1, \dots, t_n) := P(X_1 \le t_1, \dots, X_n \le t_n)$ die *gemeinsame Verteilungsfunktion* von $X_1, \dots, X_n$.

Bemerkung 0.8. (a) Wie im eindimensionalen Fall gilt, dass die gemeinsame Verteilung $P_{(X_1,\dots,X_n)}$ eindeutig durch die gemeinsame Verteilungsfunktion $F_{(X_1,\dots,X_n)}$ charakterisiert ist.

(b) Ist $X = (X_1, \dots, X_n)$ n-dimensionale diskrete Zufallsvariable, so gilt die Formel

$$F_{(X_1,\dots,X_n)}(t_1, \dots, t_n) = \sum_{x_1 \le t_1, \dots, x_n \le t_n} \varrho_{(X_1,\dots,X_n)}(x_1, \dots, x_n)\,.$$

In diesem Fall möchte man meistens die gemeinsame Zähldichte $\varrho_{(X_1,\dots,X_n)}$ direkt kennen und ist nicht so sehr an $F_{(X_1,\dots,X_n)}$ interessiert.

(c) Ist $X = (X_1, \dots, X_n)$ n-dimensionale stetige Zufallsvariable mit gemeinsamer Wahrscheinlichkeitsdichte f, so gilt für alle $(t_1, \dots, t_n) \in \mathbb{R}^n$

$$F_{(X_1,\dots,X_n)}(t_1, \dots, t_n) = \int_{-\infty}^{t_1} \dots \int_{-\infty}^{t_n} f(x_1, \dots, x_n) dx_n \dots dx_1\,.$$

Insbesondere gilt in jedem Stetigkeitspunkt $(x_1, \dots, x_n)$ von f:

$$\frac{\partial^n F_{(X_1,\dots,X_n)}}{\partial x_1 \dots \partial x_n}(x_1, \dots, x_n) = f(x_1, \dots, x_n)$$

Die folgende Aussage verallgemeinert Satz 5.18 auf den n-dimensionalen Fall:

Satz **0.9.** *Sei $X = (X_1, \dots, X_n) : \Omega \to \mathbb{R}^n$ eine n-dimensionale Zufallsvariable und sei $g : \mathbb{R}^n \to \mathbb{R}$ eine („schöne“, d.h. messbare) Funktion.*

(a) Falls X diskret ist, so existiert $E[g(X_1, \dots, X_n)]$ genau dann, wenn $\sum_{(x_1,\dots,x_n) \in \mathbb{R}^n} |g(x_1, \dots, x_n)| \varrho_{(X_1,\dots,X_n)}(x_1, \dots, x_n) < \infty$ ist und in diesem Fall gilt

$$E[g(X_1, \dots, X_n)] = \sum_{(x_1,\dots,x_n) \in \mathbb{R}^n} g(x_1, \dots, x_n) \varrho_{(X_1,\dots,X_n)}(x_1, \dots, x_n)\,.$$

(b) Falls X stetig ist, so existiert $E[g(X_1, \dots, X_n)]$ genau dann, wenn $\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} |g(x_1, \dots, x_n)| f_{(X_1,\dots,X_n)}(x_1, \dots, x_n) dx_n \dots dx_1 < \infty$ ist und in diesem Fall gilt

$$E[g(X_1, \dots, X_n)] = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} g(x_1, \dots, x_n) f_{(X_1,\dots,X_n)}(x_1, \dots, x_n) dx_n \dots dx_1\,.$$

Definition 0.10. Sei $X = (X_1, \ldots, X_n) : \Omega \to \mathbb{R}^n$ eine n-dimensionale Zufallsvariable. Dann heißen $X_1, \ldots, X_n$ *(stochastisch) unabhängig*, wenn für alle (Borelschen) Teilmengen $B_1, \ldots, B_n \subset \mathbb{R}^n$ gilt:

$$P(X_1 \in B_1, \ldots, X_n \in B_n) = P(X_1 \in B_1) \cdot \ldots \cdot P(X_n \in B_n)$$

Andernfalls heißen sie *abhängig*.

Der folgende Satz verallgemeinert Satz 5.15 auf den n-dimensionalen Fall.

Satz **0.11.** *Sei $X = (X_1, \ldots, X_n) : \Omega \to \mathbb{R}^n$ eine n-dimensionale Zufallsvariable.*

(a) Falls X diskret ist, so sind $X_1, \ldots, X_n$ genau dann unabhängig, wenn für alle $x_1 \in S_1 = X_1(\Omega), \ldots, x_n \in S_n = X_n(\Omega)$ gilt:

$$\varrho_{(X_1,\ldots,X_n)}(x_1, \ldots, x_n) = \varrho_{X_1}(x_1) \cdot \ldots \cdot \varrho_{X_n}(x_n)$$

(b) Falls X stetig ist und $f_{X_1}, \ldots, f_{X_n}$ Randichten von $X_1, \ldots, X_n$ sind, so sind $X_1, \ldots, X_n$ genau dann unabhängig, wenn die Funktion $\mathbb{R}^n \ni (x_1, \ldots, x_n) \mapsto f_{X_1}(x_1) \cdot \ldots \cdot f_{X_n}(x_n) \in \mathbb{R}$ eine gemeinsame Wahrscheinlichkeitsdichte von $X_1, \ldots, X_n$ ist.

Bemerkung 0.12. Die Aussage von Bemerkung 5.16 gilt sinngemäß auch im n-dimensionalen Fall.

Satz **0.13.** *Sei $X = (X_1, \ldots, X_n) : \Omega \to \mathbb{R}^n$ eine n-dimensionale Zufallsvariable, so dass die Varianzen $\mathrm{Var}(X_1), \ldots, \mathrm{Var}(X_n)$ existieren. Dann existiert auch die Varianz von $\sum_{j=1}^n X_j$ und es gilt*

$$\mathrm{Var}\left(\sum_{j=1}^n X_j\right) = \sum_{j=1}^n \mathrm{Var}(X_j) + 2 \sum_{1 \le i < j \le n} \mathrm{Cov}(X_i, X_j)\,.$$

Sind $X_1, \ldots, X_n$ insbesondere paarweise unkorreliert, d.h. gilt $\mathrm{Cov}(X_i, X_j) = 0$ für alle $i \neq j$, dann gilt

$$\mathrm{Var}\left(\sum_{j=1}^n X_j\right) = \sum_{j=1}^n \mathrm{Var}(X_j)\,.$$

Korollar **0.14.** *Seien $X_1, \ldots, X_n$ unabhängige, reelle Zufallsvariablen mit existierenden Varianzen $\mathrm{Var}(X_1), \ldots, \mathrm{Var}(X_n)$. Dann gilt*

$$\mathrm{Var}\left(\sum_{j=1}^n X_j\right) = \sum_{j=1}^n \mathrm{Var}(X_j)\,.$$