

11 Mehrdimensionale Zufallsvariablen

11.1 Begriffe

Def. 34 *Es seien X_i , $i = 1, \dots, p$ reellwertige, zufällige Variablen auf dem Wahrscheinlichkeitsraum (Ω, \mathcal{E}, P) . Dann heißt*

$$\mathbf{X} = (X_1, \dots, X_p)^T : \Omega \longrightarrow \mathbb{R}^p,$$

zufälliger Vektor.

Er transformiert den Wahrscheinlichkeitsraum (Ω, \mathcal{E}, P) in den Wahrscheinlichkeitsraum $(\mathbb{R}^p, \mathcal{B}^p, P_X)$, wobei \mathcal{B}^p die σ -Algebra der p -dimensionalen Borelmengen ist.

Def. 35 *Die Funktion*

$$F_{\mathbf{X}}(x_1, \dots, x_p) := P(\{\omega : X_1(\omega) < x_1, \dots, X_p(\omega) < x_p\})$$

heißt Verteilungsfunktion des zufälligen Vektors \mathbf{X} . Sie wird auch mit $F_{X_1, \dots, X_p}(x_1, \dots, x_p)$ bezeichnet.

Es gilt:

$$F_{X_1, \dots, X_p}(x_1, \dots, x_p) = P \left(\bigcap_{i=1}^p \{\omega \in \Omega : X_i(\omega) < x_i\} \right).$$

Die Verteilungsfunktion für zufällige Vektoren besitzt folgende Eigenschaften:

1. Invarianz gegenüber Permutationen, d.h.

$$F_{X_1, \dots, X_p}(x_1, \dots, x_p) = F_{X_{i_1}, \dots, X_{i_p}}(x_{i_1}, \dots, x_{i_p})$$

2. $\lim_{x_p \rightarrow \infty} F_{\mathbf{X}}(x_1, \dots, x_p) = F_{X_1, \dots, X_{p-1}}(x_1, \dots, x_{p-1});$

denn

$$F_{\mathbf{X}}(x_1, \dots, x_p) = P(\underbrace{\{\omega : X_1(\omega) < x_1\} \cap \dots \cap \{\omega : X_{p-1}(\omega) < x_{p-1}\}}_{=: A} \cap \underbrace{\{\omega : X_p(\omega) < x_p\}}_{\xrightarrow{x_p \rightarrow \infty} \Omega}).$$

Damit gilt also:

$$\begin{aligned}\lim_{x_p \rightarrow \infty} F_{\mathbf{X}}(x_1, \dots, x_p) &= P(A \cap \Omega) = P(A) \\ &= F_{X_1, \dots, X_{p-1}}(x_1, \dots, x_{p-1}).\end{aligned}$$

3. $\lim_{x_p \rightarrow -\infty} F_{\mathbf{X}}(x_1, \dots, x_p) = 0;$

Bem.: Man kann wegen 1. auch jede beliebige Komponente wählen!

4. $\lim_{(x_1, \dots, x_p) \rightarrow (\infty, \dots, \infty)} F_{\mathbf{X}}(x_1, \dots, x_p) = 1;$

5. $F_{\mathbf{X}}(x_1, \dots, x_p)$ ist in jedem Argument monoton wachsend;

6. $F_{\mathbf{X}}(x_1, \dots, x_p)$ ist in jedem Argument linksseitig stetig.

Def. 36 Ein zufälliger Vektor $\mathbf{X} = (X_1, \dots, X_p)^T$ heißt stetig verteilt, wenn seine Verteilungsfunktion charakterisiert ist durch:

$$F(x_1, \dots, x_p) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_p} f(t_1, \dots, t_p) dt_p \dots dt_1,$$

wobei für die Funktion f gilt:

1. $f(x_1, \dots, x_p) \geq 0, \forall x_1, \dots, x_p;$
2. $\int_{\mathbb{R}^p} f(x_1, \dots, x_p) dx_1 \dots dx_p = 1.$

Die Funktion $f(x_1, \dots, x_p)$ heißt dann Dichtefunktion des zufälligen Vektors \mathbf{X} .

Falls die Dichtefunktion $f(x_1, \dots, x_p)$ stetig ist, so gilt:

$$f(x_1, \dots, x_p) = \frac{\partial^p F_X(x_1, \dots, x_p)}{\partial x_1 \dots \partial x_p}.$$

Def. 37 Ein zufälliger Vektor $\mathbf{X} = (X_1, \dots, X_p)^T$ heißt

- diskret, falls jede Komponente von \mathbf{X} diskret ist, d.h. jedes X_i besitzt höchstens abzählbar viele Argumente.
- mixed(gemischt), falls einige seiner Komponenten diskret, die restlichen dagegen stetig sind.
- stetig, falls alle Komponenten von \mathbf{X} stetige Zufallsgrößen sind.

Def. 38 *Es sei $\mathbf{X} = (X_1, \dots, X_p)^T$ ein diskreter zufälliger Vektor. Für $i = 1, \dots, p$ habe X_i den Wertevorrat $\{x_{i1}, \dots, x_{ik}, \dots\}$. Dann definieren wir:*

$$p_{j\dots k} = P(X_1 = x_{1j}, \dots, X_p = x_{pk}).$$

Für die Verteilungsfunktion des zufälligen Vektors \mathbf{X} gilt:

$$\begin{aligned} F(x_1, \dots, x_p) &= P\left(\bigcap_{i=1}^p \{\omega \in \Omega : X_i(\omega) < x_i\}\right) \\ &= \sum_{\substack{j: x_{1j} < x_1 \\ \dots \\ k: x_{pk} < x_p}} p_{j\dots k} \end{aligned}$$

Es sei nun $p = 2$ und wir betrachten den zufälligen Vektor $\mathbf{X} = (X_1, X_2)^T$. Wir untersuchen zwei Fälle:

– **\mathbf{X} ist diskret.** Dann gilt zunächst:

$$X_1 : \begin{pmatrix} x_1 & x_2 & \dots & x_n & \dots \\ p_1 & p_2 & \dots & p_n & \dots \end{pmatrix}$$

$$X_2 : \begin{pmatrix} y_1 & y_2 & \dots & y_n & \dots \\ q_1 & q_2 & \dots & q_n & \dots \end{pmatrix}$$

Daraus folgt (Definition 38):

$$p_{ij} = P(X_1 = x_i, X_2 = y_j) = P(\mathbf{X} = (x_i, y_j)).$$

Weiterhin gilt:

$$P(X_1 \in \{x_i : i \in \mathbb{N}\}) = 1$$

$$P(X_2 \in \{y_j : j \in \mathbb{N}\}) = 1$$

Wir bezeichnen:

$$\mathcal{X} := \{x_i : i \in \mathbb{N}\}, \quad \mathcal{Y} := \{y_j : j \in \mathbb{N}\}.$$

Der zufällige Vektor \mathbf{X} kann Werte der Form

$(x_i, y_j) \in \mathcal{X} \times \mathcal{Y}$ annehmen. Folglich gilt:

$$P(\mathbf{X} \in \mathcal{X} \times \mathcal{Y}) = P(X_1 \in \mathcal{X}, X_2 \in \mathcal{Y}) = \sum_{i,j \in \mathbb{N}} p_{ij} = 1.$$

Es könnte evtl. interessant sein, die Wahrscheinlichkeit zu untersuchen, daß eine der zufälligen Variablen X_1

bzw. X_2 einen bestimmten Wert x_i bzw. y_j annimmt.

$$\begin{aligned}
 P(X_1 = x_i) &= P(\{X_1 = x_i\} \cap \Omega) \\
 &= P(\{X_1 = x_i\} \cap \\
 &\quad \underbrace{\{(X_2 = y_1) \cup (X_2 = y_2) \cup \dots \cup (X_2 = y_n) \cup \dots\}}_{= \bigcup_{j \in \mathbb{N}} \{X_2 = y_j\} = \Omega}) \\
 &= P(\{X_1 = x_i\} \cap \left(\bigcup_{j \in \mathbb{N}} \{X_2 = y_j\} \right)) \\
 &= P\left(\bigcup_{j \in \mathbb{N}} \{(X_1 = x_i) \cap (X_2 = y_j)\} \right) \\
 &= \sum_{j \in \mathbb{N}} p_{ij} =: p_i.
 \end{aligned}$$

Wir erhalten also:

$$p_{i\cdot} = P(X_1 = x_i).$$

Analog erhalten wir, wenn wir die Betrachtung mit der zufälligen Variable X_2 wiederholen:

$$p_{\cdot j} = P(X_2 = y_j).$$

Def. 39 Die Wahrscheinlichkeiten $p_{i\cdot}$ bzw. $p_{\cdot j}$ ($i, j \in \mathbb{N}$) nennen wir die Randwahrscheinlichkeiten des zufälligen Vektors $X = (X_1, X_2)^T$.

Das folgende Schema verdeutlicht noch einmal die Zusammenhänge zwischen den einzelnen Wahrscheinlichkeiten:

$X_1 \setminus X_2$	y_1	y_2	y_3	\dots	y_j	\dots	y_n	\dots	Σ
x_1	p_{11}	p_{12}	p_{13}	\dots	p_{1j}	\dots	p_{1n}	\dots	$p_{1\cdot}$
x_2	p_{21}	p_{22}	p_{23}	\dots	p_{2j}	\dots	p_{2n}	\dots	$p_{2\cdot}$
x_3	p_{31}	p_{32}	p_{33}	\dots	p_{3j}	\dots	p_{3n}	\dots	$p_{3\cdot}$
\vdots	\vdots	\vdots	\vdots		\vdots		\vdots		\vdots
x_i	p_{i1}	p_{i2}	p_{i3}	\dots	p_{ij}	\dots	p_{in}	\dots	$p_{i\cdot}$
\vdots	\vdots	\vdots	\vdots		\vdots		\vdots		\vdots
x_n	p_{n1}	p_{n2}	p_{n3}	\dots	p_{nj}	\dots	p_{nn}	\dots	$p_{n\cdot}$
\vdots	\vdots	\vdots	\vdots		\vdots		\vdots		\vdots
Σ	$p_{\cdot 1}$	$p_{\cdot 2}$	$p_{\cdot 3}$	\dots	$p_{\cdot j}$	\dots	$p_{\cdot n}$	\dots	1

Bsp. 64 *Es wird eine Umfrage zum Thema “Sport” durchgeführt. Dabei werden Männer und Frauen darüber befragt, ob sie Sport treiben oder nicht. Zufallsvariablen:*

$$X_1 = \begin{cases} 1 & , \text{ falls weiblich} \\ 2 & , \text{ falls männlich} \end{cases}$$
$$X_2 = \begin{cases} 1 & , \text{ falls ja} \\ 2 & , \text{ falls nein} \end{cases}$$

Schema für den zufälligen Vektor $X = (X_1, X_2)^T$:

$X_1 \setminus X_2$	1	2
1	p_{11}	p_{12}
2	p_{21}	p_{22}

2×2 -Kontingenztafel:

$X_1 \setminus X_2$	1	2	
1	n_{11}	n_{12}	$n_{1.}$
2	n_{21}	n_{22}	$n_{2.}$
	$n_{.1}$	$n_{.2}$	$n_{..}$

Dabei bedeuten:

n_{ij} – die Anzahl der Personen mit dem Geschlecht i und dem Sportverhalten j ;

$n_{.1}$ – die Anzahl der Sportler;

$n_{.2}$ – die Anzahl der Nichtsportler;

$n_{1.}$ – die Anzahl der Frauen;

$n_{2.}$ – die Anzahl der Männer;

$n_{..}$ – die Gesamtzahl der Befragten.

Mit $\hat{p}_{ij} = \frac{n_{ij}}{n_{..}}$ ergibt sich nun eine Schätzung für die Wahrscheinlichkeit p_{ij} .

Bsp. 65 *Werfen zweier Würfel. Wir betrachten den zufälligen Vektor $X = (X_1, X_2)^T$, wobei X_1 die Augenzahl des ersten Würfels ist und X_2 die des zweiten. Für die zufälligen Variablen X_1 und X_2 gilt:*

$$X_1 : \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{pmatrix}$$

$$X_2 : \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{pmatrix}$$

Da die Würfel voneinander unabhängig sind, gilt (Definition

10):

$$\begin{aligned} p_{ij} &= P(X_1 = i, X_2 = j) \\ &= P(\{X_1 = i\} \cap \{X_2 = j\}) \\ &= \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36} \end{aligned}$$

Damit erhalten wir das folgende Schema:

$X_1 \setminus X_2$	1	2	3	4	5	6
1	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
2	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
3	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
4	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
5	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
6	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$

Wir wollen nun wissen, wie groß die Wahrscheinlichkeit ist, daß der erste Würfel weniger als vier Augen zeigt und der zweite Würfel weniger als drei. Wir wissen bereits, daß die

beiden Würfeln voneinander unabhängig sind. Es gilt:

$$P(X_1 < 4, X_2 < 3) = \sum_{i < 4; j < 3} p_{ij} = \frac{6}{36} = \frac{1}{6}$$

Die hier addierten Wahrscheinlichkeiten sind in dem oben angegebenen Schema eingerahmt.

Die Aussagen zu zweidimensionalen zufälligen Vektoren, die wir bis hierher gemacht haben, gelten analog erweitert auch für höherdimensionale zufällige Vektoren.

– X ist stetig. In diesem Fall erhalten wir eine zweidimensionale Dichtefunktion $f(x, y)$, für die gilt (vgl. Definition 36):

1. $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy = 1;$
2. $f(x, y) \geq 0, \forall (x, y) \in \mathbb{R}^2.$

Die Verteilungsfunktion ergibt sich (vgl. Def. 36):

$$F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(u, v) du dv = P(X_1 < x, X_2 < y).$$

Da die Dichtefunktion $f(x, y)$ stetig ist, gilt weiterhin:

$$f(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y}.$$

Offenbar

$$\lim_{y \rightarrow \infty} F(x, y) = F_{X_1}(x) = P(X_1 < x).$$

$$\lim_{x \rightarrow \infty} F(x, y) = F_{X_2}(y) = P(X_2 < y).$$

Def. 40 Die Verteilungsfunktionen F_{X_1} und F_{X_2} bezeichnen wir als Randverteilungen von \mathbf{X} bzgl. X_1 bzw. X_2 .

Integrieren wir die Dichtefunktion nur nach einer der beiden Variablen, so erhalten wir:

$$\int_{-\infty}^{+\infty} f(x, y) dy = \frac{dF_{X_1}(x)}{dx} =: f_{X_1}(x) \quad (2)$$

Analog gilt:

$$\int_{-\infty}^{+\infty} f(x, y) dx = \frac{dF_{X_2}(y)}{dy} =: f_{X_2}(y).$$

Def. 41 Die Funktionen f_{X_1} und f_{X_2} bezeichnen wir als Randdichten von \mathbf{X} bzgl. X_1 bzw. X_2 .

Offenbar,

$$F_{X_1}(x) = \int_{-\infty}^x f_{X_1}(t) dt$$

$$F_{X_2}(y) = \int_{-\infty}^y f_{X_2}(t) dt$$

Bsp. 66 *Zweidimensionale Normalverteilung*

Descr_normal_2D