# 11 Mehrdimensionale Zufallsvariablen

## 11.1 Begriffe

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

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

*zufälliger Vektor.*

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

**Def. 35** *Die Funktion*

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

*heißt* <u>*Verteilungsfunktion*</u> *des zufälligen Vektors* **X**. *Sie wird auch mit* $\underline{F_{X_1,\ldots,X_p}(x_1, \ldots, x_p)}$ *bezeichnet.*

Es gilt:

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

Die Verteilungsfunktion für zufällige Vektoren besitzt folgende Eigenschaften:

1. Invarianz gegenüber Permutationen, d.h.

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

2. $\lim\limits_{x_p\to\infty} F_{\mathbf{X}}(x_1,\ldots,x_p) = F_{X_1,\ldots,X_{p-1}}(x_1,\ldots,x_{p-1})$;

$$F_X(x_1,\ldots,x_p) = P(\underbrace{\{\omega\colon X_1(\omega) < x_1\} \cap \ldots \cap \{\omega\colon X_{p-1}(\omega) < x_{p-1}\}}_{=:A} \cap \underbrace{\{\omega\colon X_p(\omega) < x_p\}}_{\longrightarrow_{x_p\to\infty}\Omega}.$$

Damit gilt also:

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

3. $\lim\limits_{x_p \to -\infty} F_{\mathbf{X}}(x_1, \ldots, x_p) = 0$;

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

4. $\lim\limits_{(x_1, \ldots, x_p) \to (\infty, \ldots, \infty)} F_X(x_1, \ldots, x_p) = 1$;

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

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

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

$$F(x_1, \ldots, x_p) = \int\limits_{-\infty}^{x_1} \ldots \int\limits_{-\infty}^{x_p} f(t_1, \ldots, t_p)\, dt_p \ldots dt_1,$$

*wobei für die Funktion* $f$ *gilt:*

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

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

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

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

**Def. 37** *Ein zufälliger Vektor $\mathbf{X} = (X_1, \ldots, X_p)^T$ 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* $\mathbf{X} = (X_1, \ldots, X_p)^T$ *ein diskreter zufälliger Vektor. Für* $i = 1, \ldots, p$ *habe* $X_i$ *den Wertevorrat* $\{x_{i1}, \ldots, x_{ik}, \ldots\}$. *Dann definieren wir:*

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

Für die Verteilungsfunktion des zufälligen Vektors **X** gilt:

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

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

– **X ist diskret.** Dann gilt zunächst:

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

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

Daraus folgt (Definition 38):

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

Weiterhin gilt:

$$\begin{aligned} P(X_1 \in \{x_i \colon i \in \mathbb{N}\}) &= 1 \\ P(X_2 \in \{y_j \colon j \in \mathbb{N}\}) &= 1 \end{aligned}$$

Wir bezeichnen:

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

Der zufällige Vektor **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.

$$
\begin{aligned}
P(X_1 = x_i) &= P(\{X_1 = x_i\} \cap \Omega) \\
&= P(\{X_1 = x_i\} \cap \\
&\underbrace{\{(X_2 = y_1) \vee (X_2 = y_2) \vee \ldots \vee (X_2 = y_n) \vee \ldots\}}_{= \bigcup_{j \in \mathbb{N}} \{X_2 = y_j\} = \Omega}) \\
&= P\Big(\{X_1 = x_i\} \cap \Big(\bigcup_{j \in \mathbb{N}} \{X_2 = y_j\}\Big)\Big) \\
&= P\Big(\bigcup_{j \in \mathbb{N}} \{(X_1 = x_i) \,\&\, (X_2 = y_j)\}\Big) \\
&= \sum_{j \in \mathbb{N}} p_{ij} =: p_{i\cdot}
\end{aligned}
$$

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:

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

**Bsp. 64** *Es wird eine Umfrage zum Thema „Rauchen“ durchgeführt. Dabei werden Männer und Frauen darüber befragt, ob sie Raucher oder Nichtraucher sind. Das ergibt die beiden folgenden Zufallsvariablen:*

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

$$X_2 = \begin{cases} 1 & \text{, falls Raucher} \\ 2 & \text{, falls Nichtraucher} \end{cases}$$

*Folglich erhalten wir für den zufälligen Vektor*

$X = (X_1, X_2)^T$ *das folgende Schema:*

| $X_1 \backslash X_2$ | 1 | 2 |
|---|---|---|
| 1 | $p_{11}$ | $p_{12}$ |
| 2 | $p_{21}$ | $p_{22}$ |

*Wir erstellen nun eine* $2 \times 2$*–Kontingenztafel. Dabei wird für jedes der vier Felder anhand der statistischen Erhebung die Anzahl der Personen ermittelt, für die die beiden*

*Eigenschaften zutreffen, die dieses Feld charakterisieren.*

| $X_1 \backslash X_2$ | 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 Raucherverhalten* $j$*;*

$n_{.1}$ – *die Anzahl der Raucher;*

$n_{.2}$ – *die Anzahl der Nichtraucher;*

$n_{1\cdot}$ *– die Anzahl der Frauen;*

$n_{2\cdot}$ *– die Anzahl der Männer;*

$n_{\cdot\cdot}$ *– die Gesamtzahl der Befragten.*

*Mit* $\hat{p}_{ij} = \frac{n_{ij}}{n_{\cdot\cdot}}$ *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:*

| $X_1 \backslash X_2$ | 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 Würfel 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. $\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} f(x,y)\,dx\,dy = 1;$

2. $f(x,y) \geq 0, \forall (x,y) \in \mathbb{R}^2.$

Die Verteilungsfunktion ergibt sich dann nach Definition 36 durch:

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

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

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

Offenbar

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

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

**Def. 40** *Die Verteilungsfunktionen* $F_{X_1}$ *und* $F_{X_2}$ *bezeichnen wir als* <u>*Randverteilungen*</u> *von* $X_1$ *bzw.* $X_2$.

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

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

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* *<u>Randdichten</u> von* $X_1$ *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`