11.2 Unabhängigkeit von Zufallsgrößen

Def. 42 *Es seien X_1 und X_2 zwei zufällige Variablen auf dem Wahrscheinlichkeitsraum $(\Omega, \mathcal{E}, P)$. Diese beiden zufälligen Variablen X_1 und X_2 heißen stochastisch unabhängig, wenn für alle $A, B \in \mathcal{B}^1$ gilt:*

- $P(X_1 \in A, X_2 \in B) = P(X_1 \in A) \cdot P(X_2 \in B)$;
- $P(A \cap B) = P(A) \cdot P(B)$.

Erinnerung:

$$P(X_1 \in A, X_2 \in B) = P(\{\omega\colon X_1(\omega) \in A\} \cap \{\omega\colon X_2(\omega) \in B\}).$$

Es sei $X = (X_1, X_2)^T$ ein zufälliger Vektor, für den gilt, daß die zufälligen Variablen X_1 und X_2 stochastisch unabhängig sind. Dann gilt:

$$\begin{aligned}
F_{X_1,X_2}(x_1,x_2) &= P(X_1 < x_1, X_2 < x_2) \\
&= P(X_1 \in \underbrace{(-\infty, x_1)}_{A \in \mathcal{B}^1}, X_2 \in \underbrace{(-\infty, x_2)}_{B \in \mathcal{B}^1}) \\
&= P(X_1 \in (-\infty, x_1)) \cdot P(X_2 \in (-\infty, x_2)) \\
&= F_{X_1}(x_1) \cdot F_{X_2}(x_2)
\end{aligned}$$

- Ist der zufällige Vektor $X = (X_1, X_2)^T$ außerdem stetig, so folgt:

$$\underbrace{f_{X_1,X_2}(x_1, x_2)}_{\text{zweidimensio–nale Dichte}} = \underbrace{f_{X_1}(x_1) \cdot f_{X_2}(x_2)}_{\text{Randdichten}}.$$

- Ist der zufällige Vektor $X = (X_1, X_2)^T$ außerdem diskret, so folgt für alle $i, j = 1, \ldots$:

$$p_{ij} = p_{i.} \cdot p_{.j}.$$

Bsp. 67 *In der folgenden Tabelle sind die Einzelwahrscheinlichkeiten p_{ij} einer diskreten zweidimensionalen Zufallsvariablen eingetragen. Die Komponenten X und Y seien unabhängig. Zunächst seien nur die fett eingezeichneten Einträge bekannt. Bestimmen Sie die restlichen Einträge! Zur Kontrolle sind sie blau eingetragen.*

$X \setminus Y$	1	2	3	$p_{i.}$
-1	0.02	**0.06**	0.12	0.20
0	**0.03**	0.09	0.18	0.30
1	0.05	0.15	0.3	**0.50**
$p_{.j}$	0.1	**0.30**	0.60	1

$$
\begin{aligned}
\mathbf{E}X &= -1 \cdot 0.2 + 0 \cdot 0.3 + 1 \cdot 0.5 = 0.3 \\
\mathbf{E}Y &= 1 \cdot 0.1 + 2 \cdot 0.3 + 3 \cdot 0.6 = 2.5 \\
\mathbf{E}(X \cdot Y) &= -0.02 - 2 \cdot 0.06 - 3 \cdot 0.12 + 0 \cdot (\ldots) \\
&\quad +1 \cdot 0.05 + 2 \cdot 0.15 + 3 \cdot 0.3 = 0.75 \\
cov(X, Y) &= \mathbf{E}(X \cdot Y) - (\mathbf{E}X)(\mathbf{E}Y) = 0.75 - 0.75 = 0.
\end{aligned}
$$

Merkwürdig?

Satz 30 *Es seien X_1 und X_2 zwei zufällige Variablen. φ und ψ seien zwei beliebige $\mathcal{B}^1$–meßbare Transformationen dieser beiden Variablen,*

$$X_1' = \varphi(X_1), \quad X_2' = \psi(X_2).$$

Die zufälligen Variablen X_1 und X_2 sind genau dann stochastisch unabhängig, wenn die Zufallsgrößen X_1' und X_2', für alle Transformationen φ und ψ, unabhängig sind.

Beweis: Die Funktionen φ und ψ seien auf der Menge $\mathbb{R}$ definiert und reellwertig. Dann gilt für die jeweilige

Umkehrfunktion genau dann

$$\begin{aligned}\varphi^{-1}(A) &= \{x\colon \varphi(x)\in A\}\in\mathcal{B}^1, \quad \forall A\in\mathcal{B}^1\\ \psi^{-1}(B) &= \{y\colon \psi(y)\in B\}\in\mathcal{B}^1, \quad \forall B\in\mathcal{B}^1,\end{aligned}$$

wenn φ und ψ $\mathcal{B}^1$–meßbar sind.

($\Longrightarrow$) Es seien die zufälligen Variablen X_1 und X_2 stochastisch unabhängig. Wir zeigen, daß die zufälligen Variablen $\varphi(X_1)$ und $\psi(X_2)$ unabhängig sind. Da die Funktionen φ und ψ $\mathcal{B}^1$–meßbar sind, gilt

$$
\begin{aligned}
&P(\varphi(X_1) \in A, \psi(X_2) \in B) \\
&\qquad = \quad P(X_1 \in \varphi^{-1}(A), X_2 \in \psi^{-1}(B)) \\
&\qquad = \quad P(X_1 \in \varphi^{-1}(A)) \cdot P(X_2 \in \psi^{-1}(B)) \\
&\qquad = \quad P(\varphi(X_1) \in A) \cdot P(\psi(X_2) \in B)
\end{aligned}
$$

D.h. die zufälligen Variablen $\varphi(X_1)$ und $\psi(X_2)$ sind unabhängig.

(⟸) Es gelte also, daß für alle $\mathcal{B}^1$–meßbaren Funktionen φ und ψ die zufälligen Variablen $\varphi(X_1)$ und $\psi(X_2)$ unabhängig sind. Insbesondere ist das dann auch der Fall für die Funktionen $\varphi(x) \equiv \psi(x) \equiv x$. Das heißt aber nichts anderes, als daß gilt:

$$X_1 = \varphi(X_1), \quad X_2 = \psi(X_2).$$

Folglich sind auch die zufälligen Variablen X_1 und X_2 unabhängig.

$\square$

Bsp. 68 *Sei* $X \sim \mathcal{N}(0,1)$.

- X *und* $Y = X^2$ *sind nicht unabhängig, sogar funktional abhängig*

- X *und* Y *sind unkorreliert, wegen*

  $$\mathbf{E}X = 0, \; cov(X,Y) = \mathbf{E}(X \cdot X^2) - \mathbf{E}X \cdot \mathbf{E}Y = \mathbf{E}X^3 = 0,$$

  da X *symmetrisch ist.*

Die Aussage gilt also für beliebige symmetrische Zufallsvariablen X *mit endlicher Varianz.*

11.3 Transformationssatz für Zufallsvektoren

Es sei $\mathbf{X} = (X_1, \ldots, X_p)^T$ ein zufälliger Vektor. Er habe die Dichtefunktion $f(x_1, \ldots, x_p)$. Es sei $g\colon \mathbb{R}^p \longrightarrow \mathbb{R}^p$ eine umkehrbar eindeutige Abbildung. Sie ordnet einem Vektor $\mathbf{x} = (x_1, \ldots, x_p)^T$ einen Vektor $\mathbf{y} = (y_1, \ldots, y_p)^T$ zu und besteht aus Teilabbildungen $g_1, \ldots, g_p$ mit $g_i\colon \mathbb{R}^p \longrightarrow \mathbb{R}$ (für alle $i = 1, \ldots, p$).

Bsp. 69 $\mathbf{y} = g(\mathbf{x}) = \mathbf{A} \cdot \mathbf{x}$, *wobei* $\mathbf{A}$ *eine reguläre* (p, p)*–Matrix ist.*

Die Umkehrabbildung $g^{-1}\colon \mathbb{R}^p \longrightarrow \mathbb{R}^p$ ist durch Funktionen ψ_i ($i = 1, \ldots, p$) definiert, die einem Vektor

$\mathbf{y} = (y_1, \ldots, y_p)^T$ eine Zahl x_i zuordnen, d.h. $x_i = \psi_i(y_1, \ldots, y_p)$ ($i = 1, \ldots, p$). Diese Funktionen ψ_i ($i = 1, \ldots, p$) existieren aufgrund der umkehrbaren Eindeutigkeit der Funktion g. Die Funktion g^{-1} ist dann folgendermaßen definiert:

$$g^{-1}(\mathbf{y}) \;=\; g^{-1}\begin{pmatrix} y_1 \\ \vdots \\ y_p \end{pmatrix} = \begin{pmatrix} \psi_1(y_1, \ldots, y_p) \\ \vdots \\ \psi_p(y_1, \ldots, y_p) \end{pmatrix} = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$$

In Kurzform,

$$g^{-1}(\mathbf{y}) = (\psi_1(\mathbf{y}), \ldots, \psi_p(\mathbf{y}))^T = \mathbf{x}.$$

Wir definieren nun einen weiteren zufälligen Vektor $\mathbf{Y} = (Y_1, \ldots, Y_p)^T$ wie folgt:

$$\mathbf{Y} = g(\mathbf{X}) := (g_1(X_1, \ldots, X_p), \ldots, g_p(X_1, \ldots, X_p))^T.$$

Die Abbildung $g = (g_1, \ldots, g_p)$ bildet die Transformation des zufälligen Vektors $\mathbf{X}$ in den zufälligen Vektor $\mathbf{Y}$. Wir nehmen an, die g_i ($i = 1, \ldots, p$) besitzen stetige partielle Ableitungen nach allen Argumenten.

Für den zufälligen Vektor $\mathbf{X}$ gilt umgekehrt:

$$\begin{aligned}
\mathbf{X} &= (X_1, \ldots, X_p)^T \\
&= (\psi_1(Y_1, \ldots, Y_p), \ldots, \psi_p(Y_1, \ldots, Y_p))^T \\
&= g^{-1}(Y_1, \ldots, Y_p) = g^{-1}(\mathbf{Y}).
\end{aligned}$$

Wir suchen nun die Dichtefunktion h_Y des zufälligen Vektors $\mathbf{Y}$. Analog zur Transformation zufälliger Variablen (vgl. Satz 29 auf Seite 370) gilt der folgende Satz (ohne Beweis)

Satz 31 *Die Zufallsvariable* $\mathbf{X}$ *habe die Dichte* f. *Die Dichte der Zufallsvariablen* $\mathbf{Y} = g(\mathbf{X})$ *ist*

$$h_Y(y_1, \ldots, y_p) = f(\psi_1(y_1, \ldots, y_p), \ldots, \psi_p(y_1, \ldots, y_p)) \cdot |J|,$$

wobei

$$J = \det \left(\frac{\partial \psi_i(y_1, \ldots, y_p)}{\partial y_j} \right)_{i,j=1,\ldots,p}$$

die sogenannte Jacobi-Determinante ist.

J ist also die Determinante der Matrix der partiellen Ableitungen der Funktion g^{-1}.

Bsp. 70 *Es sei* $\mathbf{X} = (X_1, X_2)^T$ *ein zufälliger Vektor* ($p = 2$)*,* *mit unabhängigen Komponenten* X_1 *und* X_2. *Die Dichte* f_{X_1,X_2} *von* $\mathbf{X}$ *sei*

$$f_{X_1,X_2}(x_1, x_2) = f_{X_1}(x_1) \cdot f_{X_2}(x_2).$$

Es sei $\mathbf{Y} = g(\mathbf{X})$ *ein weiterer zufälliger Vektor,*

$$\mathbf{Y} = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} g_1(X_1, X_2) \\ g_2(X_1, X_2) \end{pmatrix} = \begin{pmatrix} X_1 + X_2 \\ X_2 \end{pmatrix}.$$

Wir suchen nun die Dichte des zufälligen Vektors $\mathbf{Y}$.
Zunächst wissen wir, daß die Funktion g *aus zwei*

Teilfunktionen g_1 und g_2 besteht:

$$\begin{aligned} g_1(X_1, X_2) &= Y_1 = X_1 + X_2 \\ g_2(X_1, X_2) &= Y_2 = X_2 \end{aligned}$$

bzw.

$$\begin{aligned} g_1(x_1, x_2) &= y_1 = x_1 + x_2 \\ g_2(x_1, x_2) &= y_2 = x_2 \end{aligned}$$

Die Umkehrfunktion g^{-1} besteht aus den beiden Teilfunktionen:

$$\begin{aligned} \psi_1(y_1, y_2) &= x_1 = y_1 - y_2 \\ \psi_2(y_1, y_2) &= x_2 = y_2 \end{aligned}$$

Wir bestimmen nun die Zahl $|J|$:

$$|J| = \left|\det\begin{pmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{pmatrix}\right| = \left|\det\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}\right| = 1$$

Dichte des zufälligen Vektors $\mathbf{Y} = (X_1 + X_2, X_2)$:

$$\begin{aligned} h_{\mathbf{Y}}(y_1, y_2) &= f_{X_1,X_2}(\psi_1(y_1, y_2), \psi_2(y_1, y_2)) \cdot |1| \\ &= f_{X_1,X_2}(y_1 - y_2, y_2) \\ &= \underline{f_{X_1}(y_1 - y_2) \cdot f_{X_2}(y_2)} \end{aligned}$$

Randdichte für Y_1:

$$\begin{aligned} h_{\mathbf{Y}_1}(y_1) &= \int\limits_{-\infty}^{+\infty} h_{\mathbf{Y}}(y_1, y_2)\, dy_2 = \int\limits_{-\infty}^{+\infty} f_{X_1}(y_1 - y_2) \cdot f_{X_2}(y_2)\, dy_2 \\ &=: f_{X_1} * f_{X_2}(y) \end{aligned}$$

Def. 43 *Die Verknüpfung $f_{X_1} * f_{X_2}$ zweier Funktionen f_1 und f_2 heißt <u>Faltung</u> aus f_1 und f_2.*

Bem.: Die Dichte der Summe zweier unabhängiger Zufallsvariablen ist Faltung der beiden Einzeldichten.

Bsp. 71 *Es seien* $X_1, X_2 \sim R(0,1)$*, d.h.*

$$f_{X_i}(x) = \begin{cases} 1 & \text{, falls } 0 \le x < 1 \\ 0 & \text{, sonst} \end{cases}$$

Sei Y *der zufällige Vektor aus Beispiel 70. Für die Dichtefunktion der zufälligen Variablen* $Y_1 = X_1 + X_2$ *gilt* *(mittels Faltung):*

$$h_{Y_1}(y) = \int\limits_0^1 f_{X_1}(y-x) \cdot \underbrace{f_{X_2}(x)}_{\equiv 1}\, dx$$

$$= \int\limits_0^1 f_{X_1}(y-x)\, dx$$

Nun gilt: $0 \leq X_i < 1, \quad i = 1, 2.$, *d.h.*

$$0 \leq X_1 + X_2 = Y_1 < 2.$$

und für die Funktion f_{X_1}*:*

$$
\begin{aligned}
f_{X_1}(y-x) &= \begin{cases} 1 & , \textit{falls } 0 \le y-x \le 1 \\ 0 & , \textit{sonst} \end{cases} \\
&= \begin{cases} 1 & , \textit{falls } 0 \le y-1 \le x \le 1 < y \\ 1 & , \textit{falls } 0 \le x < y \le 1 \\ 0 & , \textit{falls } y-x \notin [0,1[\end{cases}
\end{aligned}
$$

Randdichte Y_1 von $\mathbf{Y}$:

$$\begin{aligned}
h_{Y_1}(y) &= \int_0^1 f_{X_1}(y-x)\,dx \\
&= \begin{cases} \int_{y-1}^{1} dx & \text{, \textit{falls} } 1 < y < 2 \\ \int_0^y dx & \text{, \textit{falls} } 0 \le y \le 1 \\ 0 & \text{, \textit{falls} } y \notin [0,2[\end{cases} \\
&= \begin{cases} 2-y & \text{, \textit{falls} } 1 < y < 2 \\ y & \text{, \textit{falls} } 0 \le y \le 1 \\ 0 & \text{, \textit{falls} } y \notin [0,2[\end{cases}
\end{aligned}$$

Wir addieren nun drei zufällige Variablen X_1, X_2, X_3, $X_i \sim R(0,1)$,

$$Y_3 = (X_1 + X_2) + X_3.$$

Für die Dichtefunktion der Zufallsgröße Y_3 gilt dann nach der Faltungsformel:

$$\begin{aligned} h_{Y_3}(z) &= h_{Y_1} * f_{X_3}(z) = \int\limits_{-\infty}^{+\infty} h_{Y_1}(z-x) \cdot f_{X_3}(x)\, dx \\ &= \int\limits_{0}^{1} h_{Y_1}(z-x) \cdot f_{X_3}(x)\, dx = \int\limits_{0}^{1} h_{Y_1}(z-x)\, dx \end{aligned}$$

Wegen $0 \leq X_i < 1$ *folgt dann:*

$$0 \leq (X_1 + X_2) + X_3 = Y_1 + X_3 = Y_3 < 3.$$

Für die Dichte der Summe der drei Zufallsgrößen X_1, X_2 *und* X_3 *gilt also:*

$$h_{Y_3}(z) = \begin{cases} 0 & \text{, \textit{falls} } z \notin [0,3[\\ \frac{z^2}{2} & \text{, \textit{falls} } 0 \leq z \leq 1 \\ 1 - \frac{(z-1)^2}{2} - \frac{(2-z)^2}{2} & \text{, \textit{falls} } 1 < z \leq 2 \\ \frac{(3-z)^2}{2} & \text{, \textit{falls} } 2 < z < 3 \end{cases}$$

`/sasuser/Stochastik/ZGWS.sas`

Vermutung: Die Summe unabhängiger Zufallsgrößen nähert sich bei wachsender Zahl der Zufallsgrößen einer Normalverteilung.

Diese Vermutung ist richtig. Sie gilt sogar (unter sehr allgemeinen Voraussetzungen) unabhängig davon, welche Verteilung diese Zufallsgrößen vorher hatten (Normal–, Gleich–, Exponentialverteilung oder diskret). Wir kommen später beim Zentralen Grenzwertsatz noch einmal darauf zurück.