1 14.10.2013

Integrationstheorie in mehr als einer Dimension, Linearität von Integranden

16.10.2013

1.1 Monotonie

$$\int_D f(x)\mathrm{d}x \leqslant \int_D g(x)\,\mathrm{d}x \quad \text{falls} \quad f(x) \leqslant g(x),\, \forall x \in D$$

$$\int_{D_1} g(x)\,\mathrm{d}x \leqslant \int_{D_2} g(x)\,\mathrm{d}x \quad \text{falls} \quad D_1 \subset D_2 \subset \mathbb{R}^2 \quad \text{und} \quad g(x) \geqslant 0,\, \forall x \in D_2$$

1.2 Vertauschbarkeit der Integrationsreihenfolge

$$\int_{a_2}^{b_2}\int_{a_1}^{b_1} g(x_1,x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2 = \int_{a_1}^{b_1}\int_{a_2}^{b_2} g(x_1,x_2)\,\mathrm{d}x_2\,\mathrm{d}x_1,\ -\infty \leqslant a_1 \leqslant b_1 \leqslant \infty, -\infty \leqslant a_2 \leqslant b_2 \leqslant \infty$$

Für Integranden der Form $g(x_1,x_2) = h_1(x_1,x_2)h_2(x_2)$ gilt:

$$\int_{a_2}^{b_2}\int_{a_1}^{b_1} h_1(x_1,x_2)h_2(x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2 = \int_{a_2}^{b_2} h_2(x_2)\int_{a_1}^{b_1} h_1(x_1,x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2$$

Für Integranden der Form $h_1(x_1)h_2(x_2)$ gilt:

$$\int_{a_2}^{b_2}\int_{a_1}^{b_1} h_1(x_1)h_2(x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2 = \int_{a_2}^{b_2} h_2(x_2)\int_{a_1}^{b_1} h_1(x_1)\,\mathrm{d}x_1\,\mathrm{d}x_2 = \int_{a_1}^{b_1} h_1(x_1)\,\mathrm{d}x_1\int_{a_2}^{b_2} h_2(x_2)\,\mathrm{d}x_2$$

Analoge Regeln und Bezeichnungen gelten für höherdimensionale Integrale. Zum Beispiel gilt für $g : \mathbb{R}^3 \to \mathbb{R}$ und $a_i \leqslant b_i, i = 1,2,3,\ldots$

$$\int_{a_1}^{b_1}\int_{a_2}^{b_2}\int_{a_3}^{b_3} g(x_1,x_2,x_3)\,\mathrm{d}x_1\,\mathrm{d}x_2\,\mathrm{d}x_3 = \int_{a_2}^{b_2}\int_{a_3}^{b_3}\int_{a_1}^{b_1} g(x_1,x_2,x_3)\,\mathrm{d}x_2\,\mathrm{d}x_3\,\mathrm{d}x_1$$

Die Integration kann in beliebiger Reihenfolge durchgeführt werden.

Für $h_1, \ldots, h_p : \mathbb{R}^1 \to \mathbb{R}^1$ gilt:

$$\int_{a_p}^{b_p} \cdots \int_{a_1}^{b_1} h_1(x_1) \cdots h_p(x_p) \, \mathrm{d}x_1 \cdots \mathrm{d}x_p = \int_{a_1}^{b_1} h_1(x_1) \, \mathrm{d}x_1 \cdots \int_{a_p}^{b_p} h_p(x_p) \, \mathrm{d}x_p$$

In Beispiel 2:

$$\int_0^2 \int_0^1 xy \, \mathrm{d}x \, \mathrm{d}y = \int_0^1 x \, \mathrm{d}x \cdot \int_0^2 y \, \mathrm{d}y = \frac{x^2}{2}\bigg|_0^1 \frac{y^2}{2}\bigg|_0^2 = \frac{1}{2} \cdot 2 = 1$$

Ist $X = (X_1, \ldots, X_p)^T$ ein Zufallsvektor mit Dichte $f : \mathbb{R}^p \to \mathbb{R}$, dann gilt:

$$P(X \in D) = \int_D f(x) \, \mathrm{d}x \quad \text{für} \quad D \subset R^p \quad \text{und} \quad \int f(x) \, \mathrm{d}x = 1$$

Beispiel 3: Sei $X = (X_1, X_2)^T$ Zufallsvektor mit Dichte

$$f(x_1, x_2) = \begin{cases} 1 & \text{falls} \quad 0 \leqslant x_1 \leqslant 1,\, 0 \leqslant x_2 \leqslant 1 \\ 0 & \text{sonst} \end{cases}$$

also Gleichverteilung auf $[0, 1]^2$.

Berechne $P(X_1 \leqslant X_2)$.

Setze

$$D := \{(X_1, X_2)^T \in R^2 : x_1 \leqslant x_2\}$$

Daher gilt

$$P(X_1 \leqslant X_2) = P(X \in D) = \int_D f(x) \, \mathrm{d}x = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathbb{1}_D(x_1, x_2) f(x_1, x_2) \, \mathrm{d}x_1 \, \mathrm{d}x_2$$

($\mathbb{1}$ bezeichnet die Indikatorfunktion).

$$\int_0^1 \int_0^1 \mathbb{1}_D(x_1, x_2) f(x_1, x_2) \, \mathrm{d}x_1 \, \mathrm{d}x_2 = \int_0^1 x_2 dx_2 = \frac{x_2^2}{2}\bigg|_0^1 = \frac{1}{2}$$

Die Indikatorfunktion kann nur die Werte 1 oder 0 annehmen und ist gleich 1, falls $0 \leqslant x_1 \leqslant x_2 \leqslant 1$ und gleich 0 sonst.

Sei X ein p-dimensionaler Zufallsvektor mit Dichte $f : \mathbb{R}^p \to \mathbb{R}$. Betrachte den partitionierten Vektor $X = (x_1, x_2)^T$ wobei X_1 k-dimensional und X_2 $p - k$-dimensionale Zufallsvektoren sind, $1 \leqslant k \leqslant p - 1$.

X_1 hat die sogenannte **Randdichte**

$$f_{X_1}(x_1) = \int f(x_1, x_2) \, \mathrm{d}x_2,\ x_1 \in R^k.$$

2 21.10.2013

X p-dimensional. Dichte f.

$$X = (X_1, X_2)^T$$

, X_1 k-dimensional, X_2 p-k-dimensional. X_1 hat die Randdichte:

$$f_{X_1}(x_1) = \int f(x_1, x_y)\,\mathrm{d}x_2,\ x_1 \in \mathbb{R}^k.$$

X_2 hat die Randdichte

$$f_{X_2}(x_2) = \int f(x_1, x_2)\,\mathrm{d}x_1,\ x_2 \in \mathbb{R}^{p-k}.$$

f heisst dann auch *gemeinsame Dichte*. Die Verteilungen von X_1 und X_2 heißen *Randverteilungen* (*Marginalverteilungen*).

Beweis f_{X_1} ist eine Dichte von X_1, für den Fall $p = 2$, $k = 1$. Für jedes $x_1 \in \mathbb{R}$ gilt:

$$P(X_1 \leqslant x_1) = P(X_1 \leqslant x_1,\ X_2 \in \mathbb{R}) = P(X \in (-\infty, x_1] \times \mathbb{R}) = \int\limits_{(-\infty, x_1]\times\mathbb{R}} f(u)\,\mathrm{d}u$$

$$= \int\limits_{-\infty}^{x_1} \int\limits_{-\infty}^{\infty} f(u_1, u_2)\,\mathrm{d}u_2\,\mathrm{d}u_1 = \int\limits_{-\infty}^{x_1} f_{X_1}(u_1)\,\mathrm{d}u_1$$

Beispiel 4 $X = (X_1, X_2)^T$ habe die Dichte f aus Beispiel 3. $Y = (Y_1, Y_2)^T$ habe die Dichte $g(y_1, y_2) = \begin{cases} 1 + (2y_1 - 1)(2y_2 - 1),\ (y_1, y_2) \in v \in [0,1]^2 \\ 0 \quad \text{sonst} \end{cases}.$

Bestimme Randdichten f_{X_i}, g_{y_i}

$$f_{X_1}(x_1) = \int\limits_{-\infty}^{\infty} f(x_1, x_2)\,\mathrm{d}x_2 = \begin{cases} 1,\ x_1 \in [0,1] \\ 0 \quad \text{sonst.} \end{cases}$$

$$f_{X_2}(x_2) = \begin{cases} 1,\ x_1 \in [0,1] \\ 0 \quad \text{sonst.} \end{cases}$$

$$g_{y_1}(y_1) = \int\limits_{-\infty}^{\infty} g(y_1, y_2)\,\mathrm{d}y_2$$

$$= \begin{cases} \int_0^1 1 + (2y_1 - 1)(2y_2 - 1)\, \mathrm{d}y_2 = 1 + (2y_1 - 1)[y_2^2 - y_2]_0^1 = 1, \; y_1 \in [0,1] \\ 0 \quad \text{sonst} \end{cases}$$

analog folgt

$$g_{y_2}(y_2) = \begin{cases} 1, \; y_2 \in [0,1] \\ 0 \quad \text{sonst.} \end{cases}$$

$\rightarrow$ X_1, X_2, Y_1, Y_2 haben dieselben Randverteilungen, aber X und Y haben verschiedene Verteilungen.

Sei wieder $X = (X_1, X_2)^T$ p-dimensionaler Zufallsvektor mit Dichte $f : \mathbb{R}^p \rightarrow \mathbb{R}$. X_1 k-dimensionaler und X_2 $p-k$-dimensionaler Zufallsvektor. Beschreibe die Verteilung von X_1 unter der Annahme, dass X_2 den festgehaltenen Wert x_2 annimmt - X bezeichnet hier eine Zufallsvariable, x ist deterministisch.

Für $x_2 \in \mathbb{R}^{p-k}$ ist die *bedingte Dichte* von X_1 gegeben $X_2 = x_2$ definiert durch

$$f_{X_1}(x_1 \,|\, X_2 = x_2) = \frac{f(x_1, x_2)}{f_{X_2}(x_2)}, \; x_1 \in \mathbb{R}^k, \quad \text{falls} \quad f_{X_2}(x_2) > 0.$$

Die bedingte Dichte von X_2 gegeben $X_1 = x_1$ ist

$$f_{X_2}(x_2 \,|\, X_1 = x_1) = \frac{f(x_1, x_2)}{f_{X_1}(x_1)}, \; x_2 \in \mathbb{R}^p - k, \quad \text{falls} \quad f_{X_1}(x_1) > 0.$$

X_1 und X_2 heißen *(stochastisch) unabhängig*, falls

$$f(x_1, x_2) = f_{X_1}(x_1) f_{X_2}(x_2), \; \forall x_1 \in \mathbb{R}^k, \; x_2 \in \mathbb{R}^{p-k}.$$

Falls X_1 und X_2 unabhängig sind, gilt für alle möglichen Werte die hier miteinander verglichen werden können, also für alle $x_1 \in \mathbb{R}^k$ mit der Eigenschaft, dass die bedingte Dichte definiert ist, also $f_{X_1}(x_1) > 0$:

$$f_{X_2}(x_2 \,|\, X_1 = x_1) = \frac{f(x_1, x_2)}{f_{X_1}(x_1)} = \frac{f_{X_1}(x_1) f_{X_2}(x_2)}{f_{X_1}(x_1)} = f_{X_2}(x_2)$$

Das heisst die bedingte Dichte von X_2 gegeben $X_1 = x_1$ hängt nicht von x_1 ab und ist gleich der Randdichte von X_2.

Sind $X_1, \ldots, X_p$ Zufallsvariablen und f eine Dichte von $(X_1, \ldots, X_p)^T$, dann sind $(X_1, \ldots, X_p)$ unabhängig, falls

$$f(x_1, \ldots, x_p) = f_{X_1}(x_1) \cdots f_{X_p}(x_p), \; \forall x_1, \ldots, x_p \in \mathbb{R}.$$

Beispiel 5

$X = (X_1, X_2)^T$ sei Zufallsvektor mit Dichte

$$f(x_1,x_2) = \begin{cases} x_1 + x_2, & \text{für} \quad 0 \leqslant x_1, x_2 \leqslant 1 \\ 0, & \text{sonst} \quad . \end{cases}$$

Randdichte:

$$f_{X_1}(x_1) = \begin{cases} \int_0^1 x_1 + x_2 \,\mathrm{d}x_2 = x_1 + [\frac{x_2^2}{2}]_0^1 = x_1 + \frac{1}{2}, \, x_1 \in [0,1] \\ 0, \, , \quad \text{sonst.} \end{cases}$$

$$f_{X_2}(x_2) = \begin{cases} x_2 + \frac{1}{2}, \, x_2 \in [0,1] \\ 0, \, , \quad \text{sonst.} \end{cases}$$

Es gilt also nicht $f(x_1,x_2) = f_{X_1}(x_1) f_{X_2}(x_2) \forall x_1, x_2 \in \mathbb{R}$. das heisst X_1 und X_2 sind nicht unabängig.

Bedingte Dichten: Für $x_2 \in [0,1]$ ist

$$f_{X_1}(x_1 \mid X_2 = x_2) = \frac{f(x_1,x_2)}{f_{X_2}(x_2)} = \begin{cases} \frac{x_1+x_2}{0.5+x_2}, \, x_1 \in [0,1] \\ 0 \quad \text{sonst.} \end{cases}$$

Für $x_1 \in [0,1]$ ist

$$f_{X_2}(x_2 \mid X_1 = x_1) = \begin{cases} \frac{x_1+x_2}{x_1+\frac{1}{2}}, \, x_2 \in [0,1] \\ 0, \quad \text{sonst.} \end{cases}$$

Für $0 \leqslant \alpha \leqslant \beta \leqslant 1$ und für $0 < \delta \leqslant \frac{1}{2}$ gilt:

$$P(\alpha \leqslant X_1 \leqslant \beta \mid \frac{1}{2} \leqslant X_2 \leqslant \frac{1}{2+\delta}) = \frac{P(\alpha \leqslant X_1 \leqslant \beta, \, \frac{1}{2} \leqslant X_2 \leqslant \frac{1}{2} + \delta)}{P(0.5 \leqslant X_2 \leqslant 0.5 + \delta)} = \frac{\int_\alpha^\beta \int_{\frac{1}{2}}^{\frac{1}{2}+\delta} f(x_1,x_2) \,\mathrm{d}x_2 \,\mathrm{d}x_1}{\int_{\frac{1}{2}}^{\frac{1}{2}+\delta} f_{X_2}(x_2) \,\mathrm{d}x_2}$$

$$= \frac{\int_\alpha^\beta \int_{\frac{1}{2}}^{\frac{1}{2}+\delta} x_1 + x_2 \,\mathrm{d}x_2 \,\mathrm{d}x_1}{\int_{\frac{1}{2}}^{\frac{1}{2}+\delta} 0.5 + x_2 \,\mathrm{d}x_2} = \frac{\int_\alpha^\beta \delta x_1 + [\frac{x_2^2}{2}]_{x_2=0.5}^{0.5+\delta} \,\mathrm{d}x_1}{\frac{\delta}{2} + [\frac{x_2^2}{2}]_{0.5}^{0.5+\delta}}$$

$$= \frac{\int_\alpha^\beta \delta x_1 + \frac{\delta}{2} + \frac{\delta^2}{2}\,\mathrm{d}x_1}{\delta + \frac{\delta^2}{2}} = \frac{\int_\alpha^\beta x_1 + \frac{1}{2}\,\mathrm{d}x_1 + \frac{\delta}{2}(\beta-\alpha)}{1+\frac{\delta}{2}}$$

$$\to \lim_{\delta\to 0^+} P(\alpha \leqslant X_1 \leqslant \beta \,|\, 0.5 \leqslant X_2 \leqslant 0.5+\delta) = \int_\alpha^\beta x_1 + 0.5\,\mathrm{d}x_1 = \int_\alpha^\beta f_{X_1}(x_1|X_2 = 0.5)\,\mathrm{d}x_1 \quad \blacksquare$$

Wichtige Parameter für univariate Verteilungen: Erwartungswert und Varianz. Entsprechende Parameter für multivariate Verteilungen ist der Vektor der Erwartungswerte und Kovarianzmatrix. Sei $X = (X_1, \dots X_p)^T$ p-dimensionaler Zufallsvektor mit Dichte f.

$$E(X_1) = \int_{-\infty}^{\infty} x_1 f_{X_1}(x_1)\,\mathrm{d}x_1 = \int_{-\infty}^{\infty} x_1 \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1,\dots,x_p)\,\mathrm{d}x_2 \dots \mathrm{d}x_p\,\mathrm{d}x_1 = \int x_1 f(X)\,\mathrm{d}x.$$

Entsprechend ist der Erwartungswert von X_j

$$E(X_j) = \int_{-\infty}^{\infty} x_j f_{X_j}\,\mathrm{d}x_j = \int x_j f(x_j)\,\mathrm{d}x,\ j = 1,\dots,p$$

Erwartungswert(-vektor) ist definiert durch $E(X) = (E(X_1),\dots,E(X_p))^T$.

Für $g \cdot \mathbb{R}^p \to \mathbb{R}$ ist $E[g(X)] = \int g(x)f(x)\,\mathrm{d}x$. Hierbei sind die Komponenten Vektoren, also:

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(x_1,\dots,x_p)f(x_1,\dots,x_p)\,\mathrm{d}x_1 \dots \mathrm{d}x_p.$$

Allgemeiner: Für matrixwertige Funktionen $G : \mathbb{R}^p \to \mathbb{R}^{n\times m}$:

$$G(x) = \begin{pmatrix} g_{11}(x) & \dots & g_{1m}(x) \\ \vdots & & \vdots \\ g_{n1}(x) & \dots & g_{nm}(x) \end{pmatrix}$$

ist $E[G(x)]$ komponentenweise definiert.

3 23.10.13

$$E[G(X)] = \begin{pmatrix} E[g_{11}(X)] & \dots & E[g_{1m}(X)] \\ \vdots & & \vdots \\ E[g_{n1}(X)] & \dots & E[g_{nm}(X)] \end{pmatrix} = (E[g_{ij}(X)])_{i=1,\dots,n,\ j=1,\dots,m}$$

Für matrixwertige Funktion $H : \mathbb{R}^p \to \mathbb{R}^{n\times m}$, $H(x) = (h_{ij}(x))_{i,j}$ ist $\int H(x)\,\mathrm{d}x$ ebenfalls komponentenweise definiert.

$$\int H(x)\,\mathrm{d}x = \begin{pmatrix} \int h_{11}(x)\,\mathrm{d}x] & \dots & \int h_{1m}(x)\,\mathrm{d}x \\ \vdots & & \vdots \\ \int h_{n1}(x)\,\mathrm{d}x & \dots & \int h_{nm}(x)\,\mathrm{d}x \end{pmatrix}$$

Damit ergibt (G matrixwertig) sich

$$E[G(X)] = E[g_{ij}(X)]_{i,j} = \left(\int g_{ij}(x)f(x)\,\mathrm{d}x\right)_{i,j} = \int G(x)f(x)\,\mathrm{d}x$$

und insbesondere

$$E(X) = \int xf(x)\,\mathrm{d}x$$

Beispiel 6

Berechne $E(X)$ für $X = (X_1, X_2)^T$ aus Beispiel 5.

$$E(X_1) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 f(x_1,x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2 = \int_0^1\int_0^1 x_1(x_1+x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2$$

$$= \int_0^1 x_1^2\,\mathrm{d}x_1 + x_2\int_0^1 x_1\,\mathrm{d}x_1\,\mathrm{d}x_2 = \int_0^1 \frac{1}{3} + x_2\frac{1}{2}\,\mathrm{d}x_2 = \frac{1}{3} + \frac{1}{2}\frac{1}{2} = \frac{7}{12}$$

Daher

$$E(X_2) = \dots = \frac{7}{12}$$

$$E(X) = (E(X_1), E(X_2))^T = (\frac{7}{12}, \frac{7}{12})^T$$

.

Rechenregeln

Für p-dimensionale Zufallsvektoren X, Y, $\alpha, \beta \in \mathbb{R}$, $g, h : \mathbb{R}^p \to \mathbb{R}^n$ gilt:

$$E[\alpha g(X) + \beta h(Y)] = \alpha E[g(X)] + \beta E[h(Y)]$$

$$E[AX+b] = AE(X)+b, \quad \text{wobei} \quad A \in \mathbb{R}^{n\times p},\, b \in \mathbb{R}^n$$

Ist $X = (X_1, X_2)^T$ mit X_1 k-dimensional und X_2 $p-k$-dimensional und ist $g : \mathbb{R}^k \to \mathbb{R}$, dann

$$E[g(X_1)] = \int g(X_1) f(x)\, \mathrm{d}x = \int g(X_1) f_{X_1}(x_1)\, \mathrm{d}x_1.$$

Sind X_1 und X_2 unabhängig und $g : \mathbb{R}^k \to \mathbb{R},\ h : \mathbb{R}^{p-k} \to \mathbb{R}$, dann

$$E[g(X_1)h(X_2)] = E[g(X_1)]E[h(X_2)]$$

Nachweis für den Fall $p = 2$, $k = 1$:

$$E[g(X_1)h(X_2)] = \int\int g(x_1)h(x_2)f(x_1,x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2$$

$$\int\int g(x_1)h(x_2)f_{X_1}(x_1)f_{X_2}(x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2 = \int g(x_1)f_{X_1}(x_1)\,\mathrm{d}x_1 \int h(x_2)f_{X_2}(x_2)\,\mathrm{d}x_2$$
$$= E[g(X_1)]E[h(X_2)]\blacksquare$$

Sind X, Y p-dimensionale unabhängige Zufallsvektoren, dann gilt

$$E(XY^T) = E(X)E(Y)^T \quad \text{und} \quad E(X^TY) = E(X)^TE(Y)$$

denn mit $X = (X_1, \ldots, X_p)^T$, $Y = (Y_1, \ldots, Y_p)^T$ ist

$$E(XY^T) = \begin{pmatrix} (X_1) & & & \\ \vdots & (Y_1 & \ldots & Y_n) \\ (X_p) & & & \end{pmatrix} = (E(X_jY_k))_{j,k=1,\ldots,p} = (E(X_j)E(Y_k))_{j,k=1,\ldots,p}$$

$$\begin{pmatrix} E(X_1) \\ \vdots \\ E(X_p) \end{pmatrix} (E(X_1), \ldots, E(Y_p)) = E(X)E(Y)^T$$

$$E(X^TY) = E(\sum_{j=1}^{p} X_jY_j) = \sum_{j=1}^{p} E(X_j)E(Y_j) = E(X)^TE(Y)\blacksquare$$

Einschub für Übungsblatt, Kovarianz:

$$\mathrm{Cov}(X_j, X_k) := E[X_j - E(X_j)(X_k - E(X_k))]$$

4 28.10.13

Sei $X = (X_1, \dots, X_p)^T$ p-dimensionaler Zufallsvektor und $\mu = (\mu_1, \dots, \mu_p) = E(X)$. (Alle ps unten sind eigentlich müs μ)

Varianz von X_j:

$$\sigma_j^2 = Var(X_j) = E[(X_j - \mu_j)^2]$$

ist ein Maß für die Streuung von X_j.

Kovarianz zwischen X_j und X_k

$$\sigma_{jk}^2 = \mathrm{Cov}(X_j, X_k) = \mathrm{E}[(X_j - \mu_j)(X_k - \mu_k)]$$

ist ein Maß für den Zusammenhang zwischen X_j und X_k, $j \neq k$, $(\sigma_{jj} = \sigma_j^2)$.

$$\mathrm{Cov}(X_j, X_k) \geqslant 0 :$$

X_j und X_k weisen tendenziell einen gleichsinnigen linearen Zusammenhang auf.

$$\mathrm{Cov}(X_j, X_k) \leqslant 0 :$$

Sind X_j und X_k unabhängig gilt:

$$\mathrm{Cov}(X_j, X_k) = \mathrm{E}(X_j - \mu_j)\mathrm{E}(X_k - \mu_k) = 0.$$

Aus $\mathrm{Cov}(X_j, X_k) = 0$ folgt im allg. nicht, dass X_j und X_k unabhängig sind, siehe Aufgabe 2, Blatt 1.

Die **Kovarianzmatrix** Σ der Zufallsvariable X ist definiert durch

$$\Sigma = Var(X) = \mathrm{E}[(X - \mu)(X - \mu)^T]$$

$$\Sigma = (\mathrm{E}[(X_j - \mu_j)(X_k - \mu_k)^T])_{j,k=1,\dots,p}$$

$$= \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \dots & \sigma_{1p} \\ \sigma_{21} & \sigma_2^2 & \dots & \sigma_{2p} \\ \vdots & & \ddots & \\ \sigma_{p1} & & \dots & \sigma_p^2 \end{pmatrix}$$

Allgemeiner: Ist $X = (X_1, \dots, X_p)^T$ ein p-dimensionaler und $Y = (Y_1, \dots, Y_q)^T$ ein q-dimensionaler Zufallsvektor dann heißt

$$\mathrm{Cov}(X,Y) = \mathrm{E}[(X - \mathrm{E}(x))(Y - \mathrm{E}(Y))^T]$$

die Kovarianzmatrix von X und Y. In Position (j,k) von $\mathrm{Cov}(X,Y)$ steht

$$\mathrm{E}[(X_j - \mathrm{E}(X_j))(Y_k - \mathrm{E}(Y_k))] = \mathrm{Cov}(X_j, X_k)$$

also

$$\mathrm{Cov}(X,Y) = (\mathrm{Cov}(X_j, Y_k))_{j=1,\dots,p,\ k=1,\dots,q}$$

Lemma 7 Sei X ein p-dimensionaler, Y ein q-dimensionanler Zufallsvektor.

1. $\mathrm{Cov}(X,X) = \mathrm{Var}(X)$
2. $\mathrm{Cov}(X,Y) = \mathrm{E}(XY^T) - \mathrm{E}(X)[\mathrm{E}(Y)]^T$, $\mathrm{Var}(X) = \mathrm{E}(XX^T) - \mathrm{E}(X)(\mathrm{E}(X))^T$
3. $\mathrm{Var}(a^TX) = a^T\mathrm{Var}(X)a$ für $a \in \mathbb{R}^p$
4. $\mathrm{Var}(AX+b) = A\mathrm{Var}(X)A^T$ für $a \in \mathbb{R}^{n\times p}$, $b \in \mathbb{R}^n$
5. Sei $p = q$. Dann gilt $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + \mathrm{Cov}(X,Y) + \mathrm{Cov}(Y,X)$. Falls X und Y unabhängig sind, dann $\mathrm{Cov}(X,Y) = \mathrm{Cov}(Y,X) = 0$ und $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$.
6. $\mathrm{Var}(X)$ ist symmetrisch und nicht-negativ definit, d.h. $a^T\mathrm{Var}(X)a \geqslant 0,\ \forall a \in \mathbb{R}^p$

Beweis: Setze $\mu = \mathrm{E}(X)$, $\upsilon = \mathrm{E}(Y)$

1. $\mathrm{Cov}(X,Y) = \mathrm{E}[(X-\mu)(X-\mu)^T] = \mathrm{Var}(X)$
2. $\mathrm{Cov}(X,Y) = \mathrm{E}[(X-\mu)(Y-\upsilon)^T] = \mathrm{E}(XY^T - \mu Y^T - X\upsilon^T + \mu\upsilon^T) = \mathrm{E}(XY^T) - \mu\mathrm{E}(Y^T) - \mathrm{E}(X)\upsilon^T + \mu\upsilon^T = \mathrm{E}(XY^T - \mathrm{E}(X)(\mathrm{E}(Y))^T$. Daraus folgt mit (a), dass $\mathrm{Var}(X) = \mathrm{Cov}(X,X) = \mathrm{E}(XX^T) - \mathrm{E}(X)(\mathrm{E}(X))^T$
3. Spezialfall von (d) mit $A = a^T,\ b = 0$
4. $\mathrm{Var}(AX+b) = \mathrm{E}[(AX+b-\mathrm{E}(AX+b))(AX+b-\mathrm{E}(AX+b))^T] = A\mathrm{E}[(X-\mathrm{E}(X))(X-\mathrm{E}(X))^TA^T] = \mathcal{A}\mathrm{Var}(X)A^T$
5. $\mathrm{Var}(X+Y) = \mathrm{E}[(X+Y)(X+Y)]^T - (\mathrm{E}(X+Y))^T = \mathrm{E}(XX^T) + \mathrm{E}(XY^T) + \mathrm{E}(YX^T) + \mathrm{E}(YY^T) - \mathrm{E}(X)(\mathrm{E}(X))^T - \mathrm{E}(X)(\mathrm{E}(Y))^T - \mathrm{E}(Y)(\mathrm{E}(X))^T - \mathrm{E}(Y)(\mathrm{E}(Y))^T = \mathrm{Var}(X) + \mathrm{Cov}(X,Y) + \mathrm{Cov}(Y,X) + \mathrm{Var}(Y)$. Sind $X = (X_1, \dots, X_p)^T$ und $Y = (Y_1, \dots, Y_p)^T$ unabhängig, dann gilt $\mathrm{E}(X_j, Y_k) = \mathrm{E}(X_j)\mathrm{E}(Y_k)$ also $\mathrm{Cov}(X_j, Y_k) = 0$. Daraus folgt $\mathrm{Cov}(X,Y) = 0,\ \mathrm{Cov}(Y,X) = 0$.
6. $\{\mathrm{Var}(X)\}_{jk} = \mathrm{Cov}(X_j, X_k) = \mathrm{Cov}(X_k, X_j) = \{\mathrm{Var}(X)\}_{kj}$ für alle $j,k = 1,\dots,p$. D.h. $\mathrm{Var}(X)$ ist symmetrisch. Für alle Vektoren $a^T \in \mathbb{R}^p$ gilt $a^T\mathrm{Var}(X)a = \mathrm{Var}(a^TX) \geqslant 0$ (wegen c).

Beispiel 8 Sei $X = (X_1, X_2)$ wie in Bsp. 5. Berechne (i) $\mathrm{Var}(X)$, (ii) $\mathrm{Var}(X_1 - 3X_2)$, (iii) $\mathrm{Var}(X_1 + X_2 + 1, 2X_1 - X_2 - 1)^T$. (i) $\mathrm{E}(X_1) = \mathrm{E}(X_2) = \frac{7}{12}$ nach Bsp. 6.
$\mathrm{E}(X_1^2) = \int x_1^2 f_{X_1}(x_1)\,\mathrm{d}x = \int\limits_0^1 x_1^2(x_1 + \frac{1}{2})\,\mathrm{d}x = \frac{x_1^4}{4}\Big|_0^1 + \frac{x_1^3}{6}\Big|_0^1 = \frac{1}{4} + \frac{1}{6} = \frac{5}{12} \implies \mathrm{Var}(X_1) = \mathrm{E}(X_1^2 - (\mathrm{E}(X_1))^T = \frac{5}{12} - (\frac{5}{12})^T = \frac{11}{144}$ Ebenso $\mathrm{Var}X_2 = \frac{11}{144}$.

$\mathrm{E}(X_1X_2) = \iint x_1x_2 f(x_1,x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2 = \int\limits_0^1\int\limits_0^1 x_1x_2 f(x_1,x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2 = \int\limits_0^1 x_1^2\,\mathrm{d}x \int\limits_0^1 x_2\,\mathrm{d}x_2 + \int\limits_0^1 x_1\,\mathrm{d}x_1 \int\limits_0^1 x_2^2\,\mathrm{d}x_2 = \frac{1}{3}\frac{1}{2} + \frac{1}{2}\frac{1}{3} = \frac{1}{3}$

$\mathrm{Cov}(X_1, X_2) = \mathrm{E}(X_1X_2) - \mathrm{E}(X_1)\mathrm{E}(X_2) = \frac{1}{3} - \frac{7}{12} \cdot \frac{7}{12} = -\frac{1}{144} \implies \mathrm{Var}(X) =$

$$\begin{pmatrix} \mathrm{Var}(X_1) & \mathrm{Cov}(X_1, X_2) \\ \mathrm{Cov}(X_2, X_1) & \mathrm{Var}(X_2) \end{pmatrix} = \begin{pmatrix} \frac{11}{144} & -\frac{11}{144} \\ -\frac{11}{144} & \frac{11}{144} \end{pmatrix}$$

(ii) Mit $a = (1, -3)^T$ ist $X_1 - 3X_3 = a^TX \implies \mathrm{Var}(X_1 - 3X_3) = \mathrm{Var}(a^TX) = a^T\mathrm{Var}(X)a$. Einsetzen liefert $\frac{29}{36}$. (iii) Mit $A = \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix}$ und $b = (1, -1)^T$ ist $AX + b = \begin{pmatrix} X_1 + X_2 + 1 \\ 2X_1 - X_2 - 1 \end{pmatrix} \implies \mathrm{Var}(AX + b) = A\mathrm{Var}(X)A^T$

5 Übungsblatt 1

Aufgabe 1

a)

$$f_{X_1}(x_1) = \int\limits_0^1 f(x_1, x_2)\,\mathrm{d}x_2 = 4x_1e^{-x_1^2}\int\limits_0^1 x_2\,\mathrm{d}x_2 = x_1e^{-x_1^2}, \quad x_1 \geqslant 0.$$

$$f_{X_1}x_1 = 0, \quad x_1 < 0$$

b)

$$f_{X_1}(x_2) = \int\limits_0^\infty f(x_1, x_2)\,\mathrm{d}x_1 = 4x_2\int\limits_0^\infty x_1e^{-x_1^2}\,\mathrm{d}x_1 = -\frac{e^{-x_1^2}}{2}\Big|_0^\infty = 2x_2, \quad x_2 \in [0, 1]$$

$$\implies f_{X_1}(x_1|X_2 = \frac{1}{2}) = \frac{f(x_1, \frac{1}{2})}{f_{X_2}(\frac{1}{2})} = 2x_1e^{-x_1^2}, \quad x_1 \geqslant 0, x_1 < 0.$$

c)

$$\mathrm{E}(X_1) = \int_0^\infty x_1 f_{X_1}(x_1)\,\mathrm{d}x_1 \underbrace{=}_{a)} 2\int_0^\infty x_1^2 e^{-x_1^2}\,\mathrm{d}x_1$$

$$= \int_{-\infty}^\infty x_1^2 e^{-x_1^2}\,\mathrm{d}x_1 \underbrace{=}_{x_1=\frac{t^2}{2}} \int_{-\infty}^\infty \frac{t^2}{2} e^{-\frac{t^2}{2}} \frac{1}{\sqrt{2}}\,\mathrm{d}t = \frac{\sqrt{\pi}}{2}\int_{-\infty}^\infty t^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}\,\mathrm{d}t = \frac{\sqrt{\pi}}{2}$$

$$\mathrm{E}(X_2) = \int_0^1 x_2 f_{X_2}(x_2)\,\mathrm{d}x_2 \underbrace{=}_{b)} \int_0^1 2x_2^2\,\mathrm{d}x_2 = \frac{2}{3} \implies \mathrm{E}(X) = \left(\frac{\sqrt{\pi}}{2}, \frac{2}{3}\right)^T$$

d) Nach a), b) gilt

$$f_{X_1} f_{X_2} = f(x_1, x_2)\ \forall x_1, x_2 \in \mathbb{R} \quad \implies X_1, X_2 \quad \text{sind unabhängig.} \quad \mathrm{Cov}(X_1, X_2) = 0.$$

e)

$$P(X_1 > X_2) = \int_0^1 \int_{x_2}^\infty f(x_1, x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2 = \int_0^1 4x_2 \int_{x_2}^\infty x_1 e^{-x_1^2}\,\mathrm{d}x_1\,\mathrm{d}x_2$$

$$= \int_0^1 4x_2 - \frac{e^{-x_1^2}}{2}\Bigg|_{x_1=x_2}^\infty \mathrm{d}x_2 = \int_0^1 4x_2 \frac{e^{-x_2^2}}{2}\,\mathrm{d}x_2 = -e^{-x_2^2}\Big|_0^1 = 1 - e^{-1}$$

Aufgabe 2

a)

$$1 = c\int_{-} 1^1 \int_{-} 1^1 x_1^2 x_2^2\,\mathrm{d}x_1\,\mathrm{d}x_2 = \int_{-} 1^1 \frac{2}{3} + 2x_2^2\,\mathrm{d}x_2$$

$$= c\left(\frac{4}{3} + \frac{4}{3}\right) = \frac{8c}{3} \implies c = \frac{3}{8}$$

b)

Für $x_1 \in [-1, 1]$ gilt

$$f_{X_1}(x_1) = \int_{-1}^1 f(x_1, x_2)\,\mathrm{d}x_2 = c\int_{-1}^1 x_1^2 + x_2^2\,\mathrm{d}x_2 = c\left(2x_1^2 + \frac{2}{3}\right)$$

$$f_{X_2}(x_2|X_1 = x_1) = \frac{f(x_1, x_2)}{f_{X_1}(x_1)} = \frac{c(x_1^2 + x_2^2)}{c(2x_1^2 + \frac{2}{3})} = \frac{x_1^2 + x_2^2}{2x_1 + \frac{2}{3}} \quad \text{falls} \quad x_2 \in [-1, 1]$$

$$f_{X_2}(x_2|X_1 = x_1) = 0 \quad \text{für} \quad x_2 \in \mathbb{R}\backslash[-1,1]$$

c)

$$\mathrm{E}(X_1) = \int_{-1}^{1} x_1 f_{X_1}(x_1)\,\mathrm{d}x_1 \underbrace{=}_{b)} = c\int_{-1}^{1} 2x_1^3 + \frac{2x_1}{3}\,\mathrm{d}x_1 = 0 \quad \mathrm{E}(X_2) = 0 \quad \text{analog.}$$

$$\mathrm{Cov}(X_1, X_2) = \mathrm{E}(X_1X_2) = c\int_{-1}^{1}\int_{-1}^{1} x_1x_2(x_1^2 + x_2^2)\,\mathrm{d}x_1\,\mathrm{d}x_2$$

$$= c\int_{-1}^{1}\int_{-1}^{1} \underbrace{x_1^3x_2 + x_1x_2^3}_{\text{ungerade Funktion von } x_1}\,\mathrm{d}x_1\,\mathrm{d}x_2 = 0$$

d) X_1 und X_2 sind nicht unabhängig, denn nach b) ist die bedingte Dichte der zweiten Variable $f_{X_2}(x_2|X_1 = x_1)$ nicht unabhängig von x_1.

Aufgabe 3 Für $j = 1,\ldots,p$ und alle $x_j \in \mathbb{R}$ gilt:

$$f_{X_j}(x_j) = \int\cdots\int f(x_1,\ldots x_p)\,\mathrm{d}x_1\ldots\mathrm{d}x_{j-1}\,\mathrm{d}x_{j+1}\ldots\mathrm{d}x_p = \int\cdots\int h_1(x_1)\cdots h_p(x_p)\,\mathrm{d}x_1\ldots\mathrm{d}x_{j-1}\,\mathrm{d}x_j$$

$$= h_j(x_j) = \prod_{i\neq j}\int h_i(x_i)\,\mathrm{d}x_i \implies f_{X_j}(x_j) = c_jh_j(x_j) \quad \text{für eine Konstante} \quad c_j$$

Wegen $1 = \int f_{X_j}(x_j)\,\mathrm{d}x_j = c_j\int h_j(x_j)\,\mathrm{d}x_j$ ist $c_j = \frac{1}{\int h_j(x_j)\,\mathrm{d}x_j}$.

Für alle $x_1,\ldots,x_p \in \mathbb{R}$ gilt:

$$f_{X_1}(x_1),\cdots f_{X_p}(x_p) = \frac{h_1(x_1)}{\int h_1(u_1)\,\mathrm{d}u_1}\cdots\frac{h_p(x_p)}{\int h_p(u_p)\,\mathrm{d}u_p}$$

$$= \frac{f(x_1,\ldots,x_p)}{\int\cdots\int h_1(u_1)\cdots h_p(u_p)\,\mathrm{d}u_1\ldots\mathrm{d}u_p} = \frac{f(x_1,\ldots,x_p)}{\int\cdots\int f(u_1,\ldots,u_p)\,\mathrm{d}u_1\ldots\mathrm{d}u_p} = f(x_1,\ldots,x_p)$$

Daher sind $X_1,\ldots,X_p$ unabhängig.

6 30.10.13

$\mathrm{Cov}(X,Y)$ ist ein Maß für den linearen Zusammenhang zwischen den beiden Zufallsvariablen X und Y. Der Wert ist aber schwer interpretierbar, denn er ist maßstabsabhängig.

$$\mathrm{Cov}(aX,bY) = ab\mathrm{Cov}(X,Y) \quad \forall a,b \in \mathbb{R}$$

Man verwendet daher ein standardisiertes Zusammenhangsmaß, nämlich den Korrelationskoeffizienten. Der *Korrelationskoeffizient* von Zufallsvariablen X und Y ist definiert durch $\rho(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}}$, sofern $\mathrm{Var}(X), \mathrm{Var}(Y) > 0$.

$(\mathrm{Var}(X) = 0 \iff P(X=c) = 1 \quad \text{für ein} \quad c \in \mathbb{R}$. ρ ist maßstabsunabhängig:

$$\rho(aX,bY) = \frac{ab\mathrm{Cov}(X,Y)}{\sqrt{a^2\mathrm{Var}(X)b^2\mathrm{Var}(Y)}} = \frac{a}{|a|}\frac{b}{|b|}\rho(X,Y), \quad a,b \neq 0$$

Vorzeichen können sich jedoch ändern.

Satz 9 X und Y seien Zufallsvariablen mit $\mathrm{Var}(X) > 0$ und $\mathrm{Var}(Y) > 0 \implies -1 \leqslant \rho(X,Y) \leqslant 1$.

$$\rho(X,Y) = 1 \iff \exists \alpha > 0, \beta \in \mathbb{R} \quad \text{so dass} \quad P(Y = \alpha X + \beta) = 1$$

$$\rho(X,Y) = -1 \iff \exists \alpha < 0, \beta \in \mathbb{R} \quad \text{so dass} \quad P(Y = \alpha X + \beta) = 1$$

Lemma 10 Cauchy-Schwarz Ungleichung.

Für Zufallsvariablen X und Y gilt $(\mathrm{E}(XY))^2 \leqslant \mathrm{E}(X^2)\mathrm{E}(Y^2)$ und

$$(\mathrm{E}(XY))^2 = \mathrm{E}(X^2)\mathrm{E}(Y^2) \iff \exists a,b \neq 0 \in \mathbb{R} : P(aX = bY) = 1$$

Beweis Setze $\alpha = \mathrm{E}(Y^2)\beta = \mathrm{E}(XY)$. Behauptung klar, falls $\alpha = 0$, da dann $P(Y=0) = 1$ und $(\mathrm{E}(XY)) = 0 = \mathrm{E}(X^2)\mathrm{E}(Y^2)$ und es gilt $P(aX = bY) = 1$ mit $a = 0, b = 1$.

Sei nun α nicht degeniert, also $\alpha \geqslant 0$.

$$0 \leqslant \mathrm{E}[(\alpha X - \beta Y)^2] = \alpha^2\mathrm{E}(X^2) - 2\alpha\beta \underbrace{\mathrm{E}(XY)}_{=\beta} + \beta^2 \underbrace{\mathrm{E}(Y^2)}_{=\alpha} = \alpha^2\mathrm{E}(X^2) - \alpha\beta^2 = \alpha(\mathrm{E}(Y^2)\mathrm{E}(X^2) - (\mathrm{E}(XY))^2)$$

$$\underbrace{\implies}_{\alpha > 0} (\mathrm{E}(XY))^2 \leqslant \mathrm{E}(X^2)\mathrm{E}(Y^2)$$

Falls $(\mathrm{E}(XY))^2 = \mathrm{E}(X^2\mathrm{E}(Y^2))$, dann

$$\mathrm{E}[(\alpha X - \beta Y)^2] = 0 \implies P(\alpha X - \beta Y = 0) = 1 \quad \text{also} \quad P(aX = bY) = 1 \quad \text{mit} \quad a = \alpha \neq 0, b = \beta$$

Falls $P(aX = bY) = 1$ mit z.B. $a \neq 0$m dann $P(X = \frac{b}{a}Y) = 1$ und

$$(\mathrm{E}(XY))^2 = (\mathrm{E}(\frac{b}{a}Y^2))^2 = \frac{b^2}{a^2}\mathrm{E}(Y^2)\mathrm{E}(Y^2) = \mathrm{E}(X^2)\mathrm{E}(Y^2)$$

04.11.13

Satz 9

$$-1 \leqslant \rho(X,Y) \leqslant 1$$

Cauchy Schwarz:

$$(\mathrm{E}(XY))^2 \leqslant \mathrm{E}(X^2)\mathrm{E}(Y^2)$$

Lemma 10

X und Y seien Zufallsvariablen mit $\mathrm{Var}(X) > 0, \mathrm{Var}(Y) > 0$

$$(\rho(X,Y))^2 = \frac{\mathrm{Cov}(X,Y)^2}{\mathrm{Var}(X)\mathrm{Var}(Y)} = \frac{(\mathrm{E}[(X-\mathrm{E}(X))(Y-\mathrm{E}(Y))])^2}{\mathrm{Var}(X)\mathrm{Var}(Y)}$$

$$\underbrace{\leqslant}_{\text{Lemma 10}} \frac{\mathrm{E}[(X-\mathrm{E}(X))^2]\mathrm{E}[(Y-\mathrm{E}(Y))]^2}{\mathrm{Var}(X)\mathrm{Var}(Y)} = 1$$

$$|\rho(X,Y)| = 1 \underbrace{\Longrightarrow}_{\text{Lemma 10}} = \exists\, a,b \in \mathbb{R}, \quad \text{nicht beide} \quad = 0, \quad \text{mit}$$

$$(*) P(a - (X - \mathrm{E}(X)) = b(Y - \mathrm{E}(Y))) = 1$$

Hier muss zusätzlich $a \neq 0$ sein, denn sonst $P(Y = \mathrm{E}(Y)) = 1$ im Widerspruch zu $\mathrm{Var}(Y) > 0$. Ebenso muss $b \neq 0$ sein.

Also kann man $(*)$ schreiben als

$$P(Y = \underbrace{\frac{a}{b}}_{\alpha} X \underbrace{- \frac{a}{b}\mathrm{E}(X) + \mathrm{E}(Y))}_{\beta}$$

Damit gezeigt

$$|\rho(X,Y) = 1| \implies \exists \alpha \neq 0, \beta \in \mathbb{R} \quad \text{mit} \quad P(Y = \alpha X + \beta) = 1$$

Bzgl. Vorzeichen von α :

Falls $P(Y = \alpha X + \beta) = 1$ mit $\alpha \neq 0$, dann

$$\rho(X,Y) = \rho(X, \alpha X + \beta) = \rho(X, \alpha X) = \frac{\alpha}{|\alpha|}\frac{\mathrm{Cov}(X,X)}{\mathrm{Var}(X)} = \begin{cases} 1, & \alpha > 0 \\ -1, & \alpha < 0. \end{cases}$$

Beispiel 11 Betrachte zweimaligen Wurf einer Münze mit $P(\text{'Kopf'}) = p \in (0,1)$.

Sei X die Anzahl Würfe, in denen Kopf fällt und Y der Rest. Berechnen nun $\rho(X,Y)$. Betrachte Erwartungswerte:

$$\mathrm{E}(X) = 2p,\ \mathrm{Var}(X) = 2p(1-p),\ \mathrm{E}(Y) = 2(1-p),\ \mathrm{Var}(Y) = 2p(1-p),$$
$$\mathrm{E}(XY) = 1 \cdot P(X = 1, Y = 1) = 2p(1-p)$$
$$\mathrm{Cov}(X,Y) = \mathrm{E}(XY) - \mathrm{E}(X)\mathrm{E}(Y) = 2p(1-p) - 4p(1-p) = -2p(1-p)$$
$$\implies \rho(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}} = \frac{-2p(1-p)}{2p(1-p)} = -1$$

Hier ist $P(X = \alpha X + \beta) = 1$ mit $\alpha = -1$ und $\beta = 2$.

Die *Korrelationsmatrix* eines Zufallsvektors $X(X_1,\dots,X_p)^T$ ist definiert durch die Matrix der einzelnen Korrelationskoeffizienten

$$P = (\rho_j k)_{j,k=1,\dots,p}, \rho_{jk} = \rho(X_j, X_k).$$

sofern $\rho_j = \sqrt{\mathrm{Var}(X_j)} > 0$ für alle $j = 1,\dots,p$.

P ist die Kovarianzmatrix des Vektors der standardisierten Zufallsvariablen $Z_j = \frac{X_j - \mathrm{E}(X_j)}{\sigma_j}$.

Mit $\Delta = \mathrm{diag}(\sigma_1,\dots,\sigma_p)$ gilt

$$P = \mathrm{Var}((Z_1,\dots,Z_p)^T) = \mathrm{Var}(\Delta^{-1}(X - \mathrm{E}(X))) \underbrace{=}_{\text{Lemma 7}} \Delta^{-1}\mathrm{Var}(X)\Delta^{-1}.$$

X_j und X_k heißen *unkorreliert*, falls $\rho(X_j, X_k) = 0$.

Satz 12 Transformationssatz für Dichten.

Sei X ein p-dimensionaler Zufallsvektor mit Dichte f_X (nicht Randdichte!).

Sei $T = \{x \in \mathbb{R}^p : f_X(x) > 0\}$ Träger von X und M sei eine offene Teilmenge von T mit $P(X \in M) = 1$.

Sei $h : M \to h(M) \subset \mathbb{R}^p$ eine bijektive Abbildung, stetig differenzierbar mit $J(x) = \det \mathrm{D}h(x) = \det \frac{\partial h_i(x)}{\partial x_j}_{i,j=1,\dots,p} \neq 0\ \forall x \in M$. D ist Differentialoperator.

Dann hat $Y = h(X)$ die Dichte

$$f_Y(y) = \begin{cases} f_X(h^{-1}(y))|\det(\mathrm{D}h^-1(y))| = \frac{f_X(h^-1(y))}{|J(h^-1(y))|}, & y \in h(M) \\ 0 \quad \text{sonst.} \end{cases}$$

Beispiel 13 Log-Normalverteilung.

Sei $X\ N(\mu,\sigma^2)$, also $f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right),\ x \in \mathbb{R}$.

Hier ist $p = 1, T = \mathbb{R}, M = \mathbb{R}, h(x) = e^x, h(M) = (0,\infty), h^-1(y) = \log y$, Funktionaldeterminante $J(x) = h'(x) = e^x \neq 0\ \forall x \in M$.

$\implies Y = h(X)$ hat die Dichte:

$$f_Y(y) \begin{cases} \frac{f_X(h^{-1}(y))}{|J(h-1(y))|} = \frac{f_X(\log y)}{|J(\log y)|} = \frac{1}{\sqrt{2\pi\sigma^2}} \exp(-\frac{(\log y-\mu)^2}{2\sigma^2})\frac{1}{y} \end{cases}$$

Beispiel 14 Sei $X = (X_1, X_2)^T$ ein 2-dimensionaler Zufallsvektor mit Dichte f_X und es sei $T = \{x \in \mathbb{R}^2 : f_X(x) > 0\}$ offen.

Bestimme Dichte von $Y = \begin{pmatrix} X_1 + X_2 \\ X_1 - X_2 \end{pmatrix}$.

Setze $M = T$ und

$$h(x) = \begin{pmatrix} h_1(x) \\ h_2(x) \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ x_1 - x_2 \end{pmatrix} \quad \text{für} \quad x = (x_1, x_2)^T \in M$$

$$J(x) = \det(\frac{\partial h_i(x)}{\partial x_j})_i, j = 1, 2 = \det \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = 2 \neq 0$$

$$h(x) = (Y_1, Y_2)^T \iff \begin{matrix} x_1 + x_2 = y_1 \\ x_1 - x_2 = y_2 \end{matrix} \iff x_1 = \frac{y_1 + y_2}{2},\ x_2 = \frac{y_1 - y_2}{2}$$

Also

$$h^{-1}(y) = \begin{pmatrix} \frac{y_1+y_2}{2} \\ \frac{y_1-y_2}{2} \end{pmatrix}$$

Insbesondere ist $h : M \to h(M)$ bijektiv.

$\implies Y$ hat die Dichte

$$f_Y(y) = \begin{cases} \frac{f_X(h^{-1}(y))}{|J(h-1(y))|} = \frac{1}{2} f_X(\frac{Y_1+Y_2}{2}, \frac{Y_1-Y_2}{2}), & y \in h(M) \\ 0 \quad \text{sonst.} \end{cases}$$

Falls $y \in \mathbb{R}^2 \backslash h(M)$, dann $h^-1(y) \in \mathbb{R}^2 \backslash M$, also $f_X(h^-1(y)) = 0$.

$$f_Y(y) = \frac{1}{2} f_X(\frac{Y_1 + Y_2}{2}, \frac{Y_1 - Y_2}{2}) \;\forall\, y \in \mathbb{R}^2$$

Wicihtiger Spezialfall: Sind X_1 und X_2 unabhängig, also

$$f_X(x) = f_{X_1}(x_1) f_{X_2}(x_2) \quad \text{für alle} \quad x \in \mathbb{R}$$

Dann hat $X_1 + X_2$ die Dichte

$$f_{X_1+X_2}(z) = f_{Y_1}(z) = \int_{-\infty}^{\infty} f_Y(z, y_2)\, \mathrm{d}y_2 = \frac{1}{2} \int_{-\infty}^{\infty} f_X(\frac{t + y_2}{2}, \frac{z - y^2}{2})\, \mathrm{d}y_2$$

$$= \frac{1}{2} \int_{-\infty}^{\infty} f_{X_1}(\frac{z-y^2}{2}) f_{X_2}(\frac{z-y^2}{2}) \, \mathrm{d}y_2 \underbrace{=}_{t=\frac{z+y_2}{2}} \frac{1}{2} \int_{-\infty}^{\infty} f_{X_1}(t) f_{X_2}(\frac{z-(2t-z)}{2}) 2 \, \mathrm{d}t$$

$$= \int_{-\infty}^{\infty} f_{X_1}(t) f_{X_2}(z-t) \, \mathrm{d}t$$

Dies nennt man Faltung von f_{X_1} und f_{X_2}.