

Multivariate Verfahren

Multivariate Verteilungen

Annika Hoyer

Sommersemester 2021

Multivariate Verteilungen - Inhalt

Wiederholung: Eindimensionale Zufallsvariablen

- Verteilungsfunktion

- Wahrscheinlichkeitsfunktion, Dichte

- Kennzahlen von Zufallsvariablen

Multivariate Zufallsvariablen

- Definition, Gemeinsame Verteilung

- Erwartungswertvektor

- Kovarianzmatrix, Korrelationsmatrix

Multivariate Normalverteilung

- Beispiele

Weitere multivariate Verteilungen

Wiederholung: Eindimensionale Zufallsvariablen

Zufallsvariablen

In der Statistik werden Daten mit Hilfe von Zufallsvariablen modelliert:

- ▶ **Diskrete Zufallsvariablen** nehmen Werte in endlicher oder höchstens abzählbar unendlicher Menge an (endliche Menge, natürliche Zahlen, ganze Zahlen, ...)
- ▶ **Stetige Zufallsvariablen** nehmen Werte auf einem Intervall oder den gesamten reellen Zahlen an. Wegen des überabzählbaren Wertebereichs gilt $P(X = x) = 0$ für alle $x \in \mathbb{R}$, d.h. die Wahrscheinlichkeit, dass eine stetige Zufallsvariable X einen Wert x genau annimmt, ist Null.

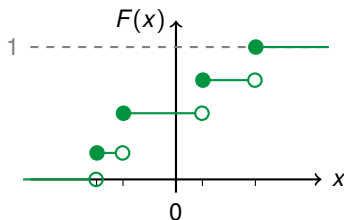
Verteilungsfunktion

Definition: Verteilungsfunktion

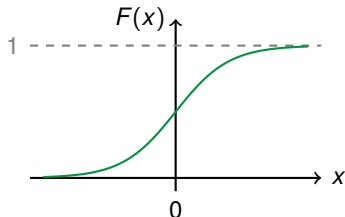
Für eine Zufallsvariable X heißt die Funktion

$$F(x) = P(X \leq x), \quad x \in \mathbb{R}$$

Verteilungsfunktion von X . Durch sie ist X eindeutig bestimmt.



Verteilungsfunktion einer
diskreten Zufallsvariable



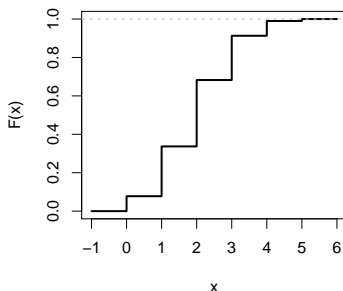
Verteilungsfunktion einer
stetigen Zufallsvariable

Beispiel 1 – Diskrete Zufallsvariablen

Binomialverteilte Zufallsvariablen: $X \sim \text{Bin}(n, p)$

- ▶ Wertebereich: $0, 1 \dots n$ (endliche Menge)
- ▶ Parameter: $n \in \mathbb{N}, p \in [0, 1]$
- ▶ Modellierung: Anzahl von Treffern (Münzwurf, Würfeln,...)

Verteilungsfunktion $\text{Bin}(5, 0.4)$

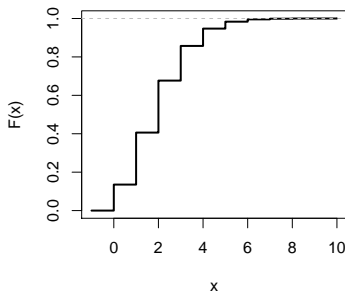


Beispiel 1 – Diskrete Zufallsvariablen

Poissonverteilte Zufallsvariablen: $X \sim \text{Poi}(\lambda)$

- ▶ Wertebereich: $0, 1, 2, \dots$ (abzählbar unendliche Menge)
- ▶ Parameter: $\lambda > 0$
- ▶ Modellierung: Zählvorgänge (Krankheitsfälle, Druckfehler,...)

Verteilungsfunktion $\text{Poi}(2)$

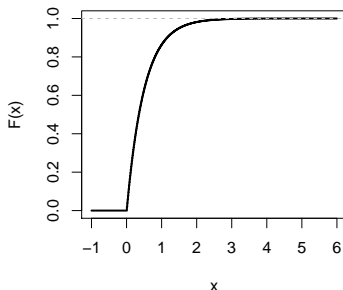


Beispiel 1 – Stetige Zufallsvariablen

Exponentialverteilte Zufallsvariablen: $X \sim \text{Exp}(\lambda)$

- ▶ Wertebereich: $[0, \infty)$ (Intervall)
- ▶ Parameter: $\lambda > 0$
- ▶ Modellierung: Zeitdauer (Wartezeit, Lebensdauer eines Produkts,...)

Verteilungsfunktion $\text{Exp}(2)$

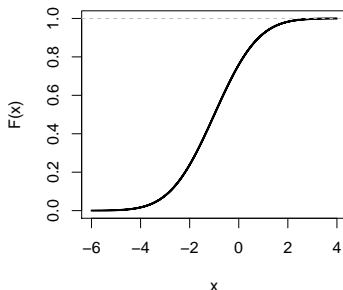


Beispiel 1 – Stetige Zufallsvariablen

Normalverteilte Zufallsvariablen: $X \sim N(\mu, \sigma^2)$

- ▶ Wertebereich: \mathbb{R} (reelle Zahlen)
- ▶ Parameter: $\mu \in \mathbb{R}, \sigma^2 > 0$
- ▶ Modellierung: Abweichung von Messwerten von einem Mittelwert (Gewicht, Größe, Alter,...)

Verteilungsfunktion $N(-1,2)$



Wahrscheinlichkeitsfunktion, Dichte

Statt über die Verteilungsfunktion F werden diskrete Zufallsvariablen häufig über ihre Wahrscheinlichkeitsfunktion, stetige Zufallsvariablen über ihre Dichte beschrieben.

Definition: Wahrscheinlichkeitsfunktion

Für eine diskrete Zufallsvariable X mit Wertebereich $\{x_1, x_2, \dots\}$ ist die **Wahrscheinlichkeitsfunktion** oder **Zähldichte** f an einem Wert $x \in \mathbb{R}$ definiert durch

$$f(x) = \begin{cases} P(X = x_i) & x = x_i \text{ mit } i = 1, 2, \dots \\ 0 & \text{sonst.} \end{cases}$$

Eigenschaften von Wahrscheinlichkeitsfunktionen:

- ▶ $0 \leq f(x) \leq 1$ für alle $x \in \mathbb{R}$.
- ▶ $\sum_{i=1}^{\infty} f(x_i) = 1$.

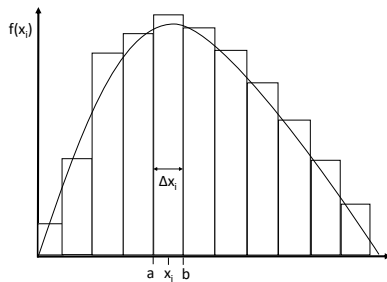
Definition: Dichte

Die **Dichte** f einer stetigen Zufallsvariablen X ist nichtnegativ (d.h. $f(x) \geq 0$ für alle $x \in \mathbb{R}$) und es gilt

$$P(a \leq X \leq b) = \int_a^b f(t) dt$$

für alle Intervalle $[a, b] \subset \mathbb{R}$.

Wahrscheinlichkeitsfunktion, Dichte



- ▶ Variable stetig, falls zu zwei Werten $a < b$ auch jeder Zwischenwert im Intervall $[a, b]$ möglich ist
- ▶ Wie werden die Wahrscheinlichkeiten bestimmt?
- ▶ Diskret: Summe über Wahrscheinlichkeiten für x_i aus $[a, b]$
- ▶ Stetig: ausgehend von diskreter Größe
- Fläche von Rechteck ($a \cdot b$): $P(X = x_i) = f(x_i)\Delta x_i$
- Δx_i immer kleiner machen
- $P(a \leq x \leq b) =$ Fläche zwischen Intervall $[a, b]$ und $f(x)$

Wahrscheinlichkeitsfunktion, Dichte

Definition: Dichte

Die **Dichte** f einer stetigen Zufallsvariablen X ist nichtnegativ (d.h. $f(x) \geq 0$ für alle $x \in \mathbb{R}$) und es gilt

$$P(a \leq X \leq b) = \int_a^b f(t) dt$$

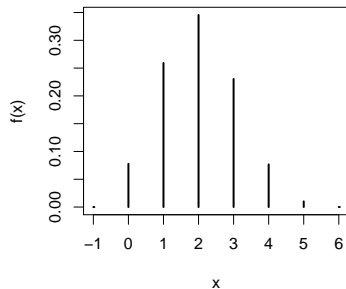
für alle Intervalle $[a, b] \subset \mathbb{R}$.

Eigenschaften von Dichtefunktionen:

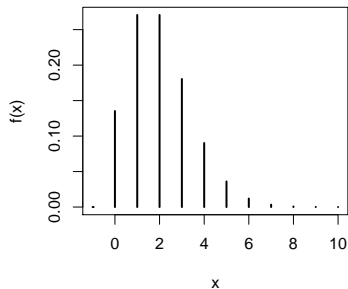
- ▶ $\int_{-\infty}^{\infty} f(t) dt = 1$
- ▶ Zusammenhang mit der Verteilungsfunktion:
 $F(x) = P(-\infty \leq X \leq x) = \int_{-\infty}^x f(t) dt$
- Die Verteilungsfunktion F ist Stammfunktion der Dichte f .
Anders formuliert: Die Dichte f ist die Ableitung der Verteilungsfunktion F .

Beispiel 1 – Diskrete Zufallsvariablen

Zähldichte Bin(5, 0.4)

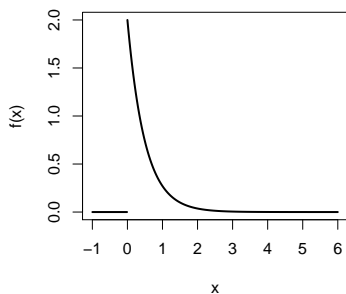


Zähldichte Poi(2)

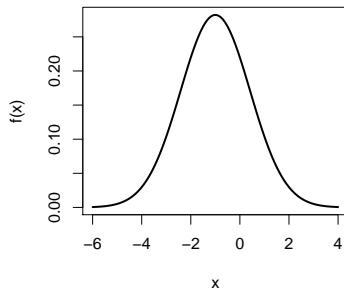


Beispiel 1 – Stetige Zufallsvariablen

Dichte Exp(2)



Dichte N(-1,2)



Kennzahlen von Zufallsvariablen

Definition: Erwartungswert

Der **Erwartungswert** gibt an, welchen Wert eine Zufallsvariable *im Mittel* annimmt.

Für eine *diskrete* Zufallsvariable mit Wahrscheinlichkeitsfunktion f und Wertebereich $\{x_1, x_2, \dots\}$ gilt

$$\mathbb{E}(X) = \sum_{i=1}^{\infty} x_i \cdot f(x_i).$$

Für eine *stetige* Zufallsvariable mit Dichte f gilt

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x \cdot f(x) dx.$$

→ $\mathbb{E}(X)$ ist ein fester Wert (**nicht** zufällig!)

Kennzahlen von Zufallsvariablen

Rechenregeln:

Für Zufallsvariablen X, Y und Skalare (Zahlen) $a, b \in \mathbb{R}$ gilt



$$\begin{aligned}\mathbb{E}(aX + b) &= \sum_{i=1}^{\infty} ax_i f(x_i) + \sum_{i=1}^{\infty} bf(x_i) \\ &= a \sum_{i=1}^{\infty} x_i f(x_i) + b \sum_{i=1}^{\infty} f(x_i) \\ &= a \cdot \mathbb{E}(X) + b\end{aligned}$$

► $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$

Kennzahlen von Zufallsvariablen

Definition: Varianz

Die **Varianz**

$$\text{Var}(X) = \mathbb{E} \left([X - \mathbb{E}(X)]^2 \right)$$

beschreibt die erwartete quadratische Abweichung vom Erwartungswert.

Varianzverschiebungssatz

$$\text{Var}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2$$

Rechenregeln:

Für Zufallsvariablen X, Y und Skalare (Zahlen) $a, b \in \mathbb{R}$ gilt

- ▶ $\text{Var}(aX + b) = a^2 \text{Var}(X)$
- ▶ $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y)$

Beispiel 1

Erwartungswert und Varianz lassen sich häufig in Abhängigkeit der Verteilungsparameter darstellen.

Beispiele

- ▶ $X \sim \text{Bin}(n, p)$: $\mathbb{E}(X) = np$, $\text{Var}(X) = np(1 - p)$
- ▶ $X \sim \text{Poi}(\lambda)$: $\mathbb{E}(X) = \lambda$, $\text{Var}(X) = \lambda$
- ▶ $X \sim \text{Exp}(\lambda)$: $\mathbb{E}(X) = \frac{1}{\lambda}$, $\text{Var}(X) = \frac{1}{\lambda^2}$
- ▶ $X \sim \text{N}(\mu, \sigma^2)$: $\mathbb{E}(X) = \mu$, $\text{Var}(X) = \sigma^2$

Multivariate Zufallsvariablen

Definition Zufallsvektor

Definition: Zufallsvektor

Betrachte p Zufallsvariablen $X_1 \dots X_p$. Der p -dimensionale Vektor

$$\mathbf{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_p \end{pmatrix} = (X_1, \dots, X_p)^\top \in \mathbb{R}^p$$

heißt **Zufallsvektor** oder **p -dimensionale Zufallsvariable**.

- ▶ Der Zufallsvektor \mathbf{X} heißt *diskret*, falls alle Komponenten X_i von \mathbf{X} diskrete Zufallsvariablen sind.
- ▶ Der Zufallsvektor \mathbf{X} heißt *stetig*, falls alle Komponenten X_i von \mathbf{X} stetige Zufallsvariablen sind. Dann gilt $P(X_1 = x_1, \dots, X_p = x_p) = 0$ für alle $\mathbf{x} = (x_1 \dots x_p)^\top \in \mathbb{R}^p$.

Gemeinsame Verteilungsfunktion

Definition: Gemeinsame Verteilungsfunktion

Für einen Zufallsvektor $\mathbf{X} \in \mathbb{R}^p$ heißt die Funktion

$$F(\mathbf{x}) = F(x_1, \dots, x_p) = P(X_1 \leq x_1 \dots X_p \leq x_p), \quad \mathbf{x} \in \mathbb{R}^p$$

gemeinsame Verteilungsfunktion der Zufallsvariablen $X_1 \dots X_p$.
Durch sie ist \mathbf{X} eindeutig bestimmt.

Analog zum eindimensionalen Fall lassen sich für diskrete bzw. stetige Zufallsvektoren gemeinsame Wahrscheinlichkeitsfunktionen bzw. Dichten definieren.

Gemeinsame Wahrscheinlichkeitsfunktion / Dichte

Definition: Gemeinsame Wahrscheinlichkeitsfunktion

Für einen diskreten Zufallsvektor $\mathbf{X} \in \mathbb{R}^p$ mit Wertebereich $\{\mathbf{x}_1, \mathbf{x}_2, \dots\}$ ist die (gemeinsame)

Wahrscheinlichkeitsfunktion oder (gemeinsame) **Zähldichte** f an $\mathbf{x} \in \mathbb{R}^p$ definiert durch

$$f(\mathbf{x}) = \begin{cases} P(\mathbf{X} = \mathbf{x}_i) & \mathbf{x} = \mathbf{x}_i \text{ mit } i = 1, 2, \dots \\ 0 & \text{sonst.} \end{cases}$$

Definition: Gemeinsame Dichte

Die (gemeinsame) **Dichte** f eines stetigen Zufallsvektors $\mathbf{X} \in \mathbb{R}^p$ ist nichtnegativ (d.h. $f(\mathbf{x}) \geq 0$ für alle $\mathbf{x} \in \mathbb{R}^p$) und es gilt

$$\begin{aligned} P(\mathbf{a} \leq \mathbf{X} \leq \mathbf{b}) &= P(a_1 \leq X_1 \leq b_1, \dots, a_p \leq X_p \leq b_p) \\ &= \int_{a_p}^{b_p} \cdots \int_{a_1}^{b_1} f(t_1, \dots, t_p) dt_1 \dots dt_p. \end{aligned}$$

für alle $\mathbf{a}, \mathbf{b} \in \mathbb{R}^p$ mit $a_j < b_j$, $j = 1 \dots p$.

Gemeinsame Wahrscheinlichkeitsfunktion / Dichte

Eigenschaften der gemeinsamen Dichte:

- Wie im eindimensionalen Fall gilt:

$$\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f(t_1, \dots, t_p) dt_1 \dots dt_p = 1$$

- Zusammenhang mit der gemeinsamen Verteilungsfunktion:

$$\begin{aligned} F(\mathbf{x}) &= F(x_1, \dots, x_p) = P(X_1 \leq x_1, \dots, X_p \leq x_p) \\ &= \int_{-\infty}^{x_p} \cdots \int_{-\infty}^{x_1} f(t_1, \dots, t_p) dt_1 \dots dt_p = \int_{-\infty}^{\mathbf{x}} f(\mathbf{t}) d\mathbf{t} \end{aligned}$$

Erwartungswertvektor

Definition: Erwartungswertvektor

Gegeben sei der Zufallsvektor $\mathbf{X} = (X_1, \dots, X_p)^\top \in \mathbb{R}^p$. Dann ist der **Erwartungswertvektor** von \mathbf{X} definiert als

$$\mathbb{E}(\mathbf{X}) = \begin{pmatrix} \mathbb{E}(X_1) \\ \vdots \\ \mathbb{E}(X_p) \end{pmatrix}.$$

Rechenregeln:

Für Zufallsvektoren $\mathbf{X}, \mathbf{Y} \in \mathbb{R}^p$, eine Matrix $\mathbf{A} \in \mathbb{R}^{q \times p}$ und einen Vektor $\mathbf{b} \in \mathbb{R}^q$ gilt

- ▶ $\mathbb{E}(\mathbf{X} + \mathbf{Y}) = \mathbb{E}(\mathbf{X}) + \mathbb{E}(\mathbf{Y})$
- ▶ $\mathbb{E}(\mathbf{A} \cdot \mathbf{X} + \mathbf{b}) = \mathbf{A} \cdot \mathbb{E}(\mathbf{X}) + \mathbf{b}$

Definition: Kovarianz

Die **Kovarianz** von zwei Zufallsvariablen X und Y ist definiert als

$$\text{Cov}(X, Y) = \mathbb{E}[(X - \mathbb{E}(X)) \cdot (Y - \mathbb{E}(Y))].$$

Verschiebungssatz

$$\text{Cov}(X, Y) = \mathbb{E}(X \cdot Y) - \mathbb{E}(X)\mathbb{E}(Y)$$

Sind X und Y stetig, so gilt

$$\text{Cov}(X, Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mathbb{E}(X))(y - \mathbb{E}(Y)) f(x, y) dx dy$$

$$\mathbb{E}(X \cdot Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x \cdot y \cdot f(x, y) dx dy$$

Definition: Korrelation

Die **Korrelation** von zwei Zufallsvariablen X und Y ist definiert als

$$\text{Corr}(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X) \cdot \text{Var}(Y)}}.$$

Eigenschaften:

- ▶ $\text{Cov}(X, Y) = \text{Cov}(Y, X)$
 $\text{Corr}(X, Y) = \text{Corr}(Y, X)$
- ▶ $\text{Cov}(X, X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = \text{Var}(X)$
 $\Rightarrow \text{Corr}(X, X) = 1$
- ▶ $-1 \leq \text{Corr}(X, Y) \leq 1$
- ▶ X, Y unabhängig $\Rightarrow \text{Cov}(X, Y) = 0$ und $\text{Corr}(X, Y) = 0$
Umkehrung (\Leftarrow) gilt i.A. nicht!

Rechenregeln:

Für Zufallsvariablen X, Y und Skalare a, b, c, d gilt

- ▶ $\text{Cov}(aX + c, bY + d) = ab\text{Cov}(X, Y)$
- ▶ $\text{Corr}(aX + c, bY + d) = \text{Corr}(X, Y)$

Kovarianzmatrix

Definition: Kovarianzmatrix

Gegeben sei der Zufallsvektor $\mathbf{X} = (X_1, \dots, X_p)^\top \in \mathbb{R}^p$. Dann ist die Kovarianzmatrix $\text{Cov}(\mathbf{X})$ von \mathbf{X} definiert durch

$$\begin{aligned}\text{Cov}(\mathbf{X}) = \Sigma &= \mathbb{E} \left[(\mathbf{X} - \mathbb{E}(\mathbf{X}))(\mathbf{X} - \mathbb{E}(\mathbf{X}))^\top \right] \\ &= \begin{pmatrix} \text{Cov}(X_1, X_1) & \dots & \text{Cov}(X_1, X_p) \\ \vdots & \ddots & \vdots \\ \text{Cov}(X_p, X_1) & \dots & \text{Cov}(X_p, X_p) \end{pmatrix} \in \mathbb{R}^{p \times p}.\end{aligned}$$

Verschiebungssatz

$$\text{Cov}(\mathbf{X}) = \mathbb{E}(\mathbf{X}\mathbf{X}^\top) - \mathbb{E}(\mathbf{X})\mathbb{E}(\mathbf{X})^\top.$$

Kovarianzmatrix

Eigenschaften:

- ▶ $\text{Cov}(\mathbf{X})$ ist quadratisch.
- ▶ $\text{Cov}(\mathbf{X})$ ist symmetrisch.
- ▶ Die Diagonalelemente von $\text{Cov}(\mathbf{X})$ sind gegeben durch

$$\text{Cov}(X_i, X_i) = \text{Var}(X_i) \geq 0 \quad \text{für alle } i = 1 \dots, p.$$

- ▶ $\text{Cov}(\mathbf{X})$ ist positiv semidefinit.

Rechenregeln:

Für einen Zufallsvektor $\mathbf{X} \in \mathbb{R}^p$, eine Matrix $\mathbf{A} \in \mathbb{R}^{q \times p}$ und Vektoren $\mathbf{a} \in \mathbb{R}^q$, $\mathbf{b} \in \mathbb{R}^q$ gilt

- ▶ $\text{Var}(\mathbf{a}^\top \mathbf{X}) = \mathbf{a}^\top \text{Cov}(\mathbf{X}) \mathbf{a}$
- ▶ $\text{Cov}(\mathbf{A}\mathbf{X} + \mathbf{b}) = \mathbf{A} \text{Cov}(\mathbf{X}) \mathbf{A}^\top$

Kovarianzmatrix

Welche Matrizen sind gültige Kovarianzmatrizen?

$$\Sigma_1 = \begin{pmatrix} 0.2 & 0.5 \\ 0.2 & 0.3 \\ 0.5 & 0.3 \end{pmatrix} \quad \Sigma_2 = \begin{pmatrix} 0.5 & 0.7 & 0.9 \\ 0.3 & 0.9 & 0.3 \\ 0.9 & 0.7 & 0.5 \end{pmatrix}$$

$$\Sigma_3 = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \quad \Sigma_4 = \begin{pmatrix} 0.5 & 0.7 & -0.9 \\ 0.7 & 0.9 & 0.3 \\ -0.9 & 0.3 & -0.5 \end{pmatrix}$$

Kovarianzmatrix

Für zwei Vektoren $\mathbf{X} = (X_1, \dots, X_p)^\top$ und $\mathbf{Y} = (Y_1, \dots, Y_q)$ gilt

$$\Sigma_{XY} = \text{cov}(\mathbf{X}, \mathbf{Y}) = \begin{pmatrix} \text{cov}(X_1, Y_1) & \dots & \text{cov}(X_1, Y_q) \\ \vdots & \ddots & \vdots \\ \text{cov}(X_p, Y_1) & \dots & \text{cov}(X_p, Y_q) \end{pmatrix}$$

→ Σ_{XY} ist i.A. nicht symmetrisch!

Rechenregeln

► $\mathbb{E}(\mathbf{AX} + \mathbf{b}) = \mathbf{A}\mathbb{E}(\mathbf{X}) + \mathbf{b}$

►

$$\begin{aligned}\text{Cov}(\mathbf{X}) &= \mathbb{E} \left((\mathbf{X} - \boldsymbol{\mu})(\mathbf{X} - \boldsymbol{\mu})^\top \right) \\ &= \mathbb{E} \left(\mathbf{XX}^\top - \mathbf{X}\boldsymbol{\mu}^\top - \boldsymbol{\mu}\mathbf{X}^\top + \boldsymbol{\mu}\boldsymbol{\mu}^\top \right) \\ &= \mathbb{E}(\mathbf{XX}^\top) - \mathbb{E}(\mathbf{X}\boldsymbol{\mu}^\top) - \mathbb{E}(\boldsymbol{\mu}\mathbf{X}^\top) + \mathbb{E}(\boldsymbol{\mu}\boldsymbol{\mu}^\top) \\ &= \mathbb{E}(\mathbf{XX}^\top) - \mathbb{E}(\mathbf{X})\boldsymbol{\mu}^\top - \boldsymbol{\mu}\mathbb{E}(\mathbf{X}^\top) + \boldsymbol{\mu}\boldsymbol{\mu}^\top \\ &= \mathbb{E}(\mathbf{XX}^\top) - \boldsymbol{\mu}\boldsymbol{\mu}^\top - \boldsymbol{\mu}\boldsymbol{\mu}^\top + \boldsymbol{\mu}\boldsymbol{\mu}^\top \\ &= \mathbb{E}(\mathbf{XX}^\top) - 2\boldsymbol{\mu}\boldsymbol{\mu}^\top + \boldsymbol{\mu}\boldsymbol{\mu}^\top \\ &= \mathbb{E}(\mathbf{XX}^\top) - \boldsymbol{\mu}\boldsymbol{\mu}^\top\end{aligned}$$

► $\text{Cov}(\mathbf{AX} + \mathbf{b}) = \mathbf{A}\text{Cov}(\mathbf{X})\mathbf{A}^\top$

► $\text{Cov}(\mathbf{X}, \mathbf{Y}) = \mathbb{E} \left((\mathbf{X} - \mathbb{E}(\mathbf{X}))(\mathbf{Y} - \mathbb{E}(\mathbf{Y}))^\top \right)$

► $\text{Cov}(\mathbf{AX} + \mathbf{a}, \mathbf{BY} + \mathbf{b}) = \mathbf{A}\text{Cov}(\mathbf{X}, \mathbf{Y})\mathbf{B}^\top$

Zerlegung der Kovarianzmatrix

Es gilt:

$$\Sigma = \Sigma^{\frac{1}{2}} \Sigma^{\frac{1}{2}} \quad \text{bzw.} \quad \Sigma = \Sigma^{\frac{1}{2}} \Sigma^{\frac{T}{2}}$$

und

$$\Sigma^{-1} = \Sigma^{-\frac{T}{2}} \Sigma^{-\frac{1}{2}}$$

Daraus resultiert:

$$\begin{aligned} \text{Cov}(\Sigma^{-\frac{1}{2}} \mathbf{X}) &= \Sigma^{-\frac{1}{2}} \text{Cov}(\mathbf{X}) \Sigma^{-\frac{T}{2}} \\ &= \Sigma^{-\frac{1}{2}} \Sigma \Sigma^{-\frac{T}{2}} \\ &= \Sigma^{-\frac{1}{2}} \Sigma^{\frac{1}{2}} \Sigma^{\frac{T}{2}} \Sigma^{-\frac{T}{2}} \\ &= (\Sigma^{\frac{1}{2}})^{-1} \Sigma^{\frac{1}{2}} \Sigma^{\frac{T}{2}} (\Sigma^{\frac{T}{2}})^{-1} \\ &= \mathbf{I} \end{aligned}$$

$\Rightarrow \Sigma^{-\frac{1}{2}}(\mathbf{X} - \mathbb{E}(\mathbf{X}))$ hat Erwartungswert **0** and Varianz **I**
(Multivariate Standardisierung).

Spektralzerlegung

Es gilt:

$$\Sigma = \mathbf{P}\Lambda\mathbf{P}^\top \quad \text{und} \quad \Sigma^{-1} = \mathbf{P}\Lambda^{-1}\mathbf{P}^\top,$$

wobei

$\mathbf{P} = (\mathbf{p}_1, \dots, \mathbf{p}_p)^\top$: Matrix der *orthonormalen* Eigenvektoren
und

$\Lambda = \text{diag}(\lambda_1, \dots, \lambda_p)$: Diagonalmatrix der *geordneten*
Eigenwerte.

Korrelationsmatrix

Definition: Korrelationsmatrix

Die **Korrelationsmatrix** eines Zufallsvektors $\mathbf{X} \in \mathbb{R}^p$ ist wie folgt definiert:

$$\text{Corr}(\mathbf{X}) = \begin{pmatrix} \text{Corr}(X_1, X_1) & \dots & \text{Corr}(X_1, X_p) \\ \vdots & \ddots & \vdots \\ \text{Corr}(X_p, X_1) & \dots & \text{Corr}(X_p, X_p) \end{pmatrix} \in \mathbb{R}^{p \times p}.$$

Eigenschaften:

- ▶ $\text{Corr}(\mathbf{X})$ ist quadratisch
- ▶ $\text{Corr}(\mathbf{X})$ ist symmetrisch
- ▶ Diagonale: $\text{Corr}(X_i, X_i) = 1$ für alle $i = 1 \dots, p$
- ▶ $\text{Corr}(\mathbf{X})$ ist positiv semidefinit.

Korrelation und Unabhängigkeit

Definition: Unkorrelierte Zufallsvektoren

Seien $\mathbf{Y} \in \mathbb{R}^r$ und $\mathbf{Z} \in \mathbb{R}^q$ zwei Zufallsvektoren mit

$$\begin{aligned} \text{Cov} \left(\begin{pmatrix} \mathbf{Y} \\ \mathbf{Z} \end{pmatrix} \right) &= \begin{pmatrix} \text{Cov}(Y_1, Y_1) & \dots & \text{Cov}(Y_1, Y_r) & \text{Cov}(Y_1, Z_1) & \dots & \text{Cov}(Y_1, Z_q) \\ & \ddots & & & \ddots & \\ \text{Cov}(Y_r, Y_1) & \dots & \text{Cov}(Y_r, Y_r) & \text{Cov}(Y_r, Z_1) & \dots & \text{Cov}(Y_r, Z_q) \\ \text{Cov}(Z_1, Y_1) & \dots & \text{Cov}(Z_1, Y_r) & \text{Cov}(Z_1, Z_1) & \dots & \text{Cov}(Z_1, Z_q) \\ & \ddots & & & \ddots & \\ \text{Cov}(Z_q, Y_1) & \dots & \text{Cov}(Z_q, Y_r) & \text{Cov}(Z_q, Z_1) & \dots & \text{Cov}(Z_q, Z_q) \end{pmatrix} \\ &= \begin{pmatrix} \Sigma_Y & \Sigma_{YZ} \\ \Sigma_{ZY} & \Sigma_Z \end{pmatrix}. \end{aligned}$$

\mathbf{Y} und \mathbf{Z} heißen **unkorreliert**, falls $\Sigma_{YZ} = \mathbf{0}$.

Bemerkung:

- \mathbf{Y} und \mathbf{Z} sind unabhängig \Rightarrow \mathbf{Y} und \mathbf{Z} sind unkorreliert.
Die Umkehrung (\Leftarrow) gilt i.A. nicht!

Multivariate Normalverteilung

Multivariate Normalverteilung

Definition: Multivariate Normalverteilung

Ein Zufallsvektor $\mathbf{X} \in \mathbb{R}^p$ heißt **multivariat normalverteilt** mit Parametern $\boldsymbol{\mu}$ und Σ , falls er die folgende Dichte hat

$$f(\mathbf{x}) = \frac{1}{(\sqrt{2\pi})^p \sqrt{\det(\Sigma)}} \exp \left\{ -\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^\top \Sigma^{-1}(\mathbf{x} - \boldsymbol{\mu}) \right\}$$

mit $\boldsymbol{\mu} \in \mathbb{R}^p$ und positiv definiter Matrix $\Sigma \in \mathbb{R}^{p \times p}$.

Die Schreibweise ist analog zur eindimensionalen Normalverteilung:

$$\mathbf{X} \sim N_p(\boldsymbol{\mu}, \Sigma)$$

Häufig wird der Index p unterdrückt, wenn sich die Dimension aus dem Zusammenhang erschließen lässt.

Zusammenhang zur Spektralzerlegung

Es gilt (quadratische Form):

$$\begin{aligned}(\mathbf{X} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{X} - \boldsymbol{\mu}) &= (\mathbf{X} - \boldsymbol{\mu})^\top \mathbf{P} \boldsymbol{\Lambda}^{-1} \mathbf{P}^\top (\mathbf{X} - \boldsymbol{\mu}) = \\ &= \mathbf{Y} \boldsymbol{\Lambda}^{-1} \mathbf{Y}^\top = \sum_{i=1}^p \frac{y_i^2}{\lambda_i} = \sum_{i=1}^p \left(\frac{y_i}{\sqrt{\lambda_i}} \right)^2\end{aligned}$$

Die Isodensiten (Kurven mit derselben Dichte) sind damit durch folgende Gleichung bestimmt:

$$(\mathbf{X} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{X} - \boldsymbol{\mu}) = \sum_{i=1}^p \left(\frac{y_i}{\sqrt{\lambda_i}} \right)^2 = c$$

→ Dies stellt in den y-Werten ein Ellipsoid mit den Hauptachsenlängen $\sqrt{\lambda_i}$ dar!

Multivariate Normalverteilung

Eigenschaften:

- ▶ $\mathbb{E}(\mathbf{X}) = \boldsymbol{\mu}$, $\text{Cov}(\mathbf{X}) = \Sigma$
- ▶ Für $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \Sigma)$ und $\mathbf{A} \in \mathbb{R}^{q \times p}$, $\mathbf{b} \in \mathbb{R}^q$ gilt

$$\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{b} \sim N_q(\mathbf{A}\boldsymbol{\mu} + \mathbf{b}, \mathbf{A}\Sigma\mathbf{A}^\top).$$

- ▶ $(\mathbf{X} - \boldsymbol{\mu})^\top \Sigma^{-1} (\mathbf{X} - \boldsymbol{\mu}) \sim \chi^2(p)$

Spezialfälle:

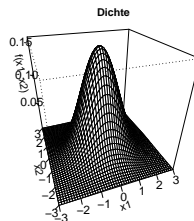
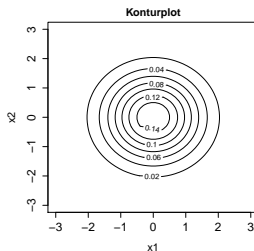
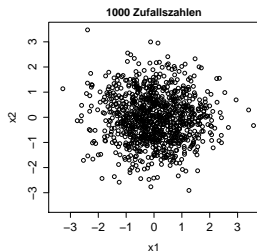
- ▶ Für $p = 1$ ergibt sich die univariate Normalverteilung mit Parametern $\mu = \mathbb{E}(X)$ und $\Sigma = \text{Var}(X)$.
- ▶ Für $\mathbf{X} \sim N_p(\mathbf{0}, \mathbf{I})$, also mit

$$\boldsymbol{\mu} = \mathbf{0} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \quad \text{und} \quad \Sigma = \mathbf{I} = \begin{pmatrix} 1 & \cdots & 0 \\ & \ddots & \\ 0 & \cdots & 1 \end{pmatrix},$$

ergibt sich die multivariate **Standardnormalverteilung**.

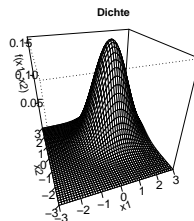
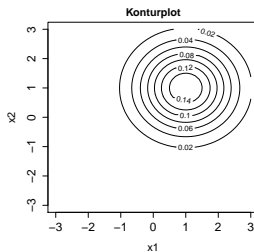
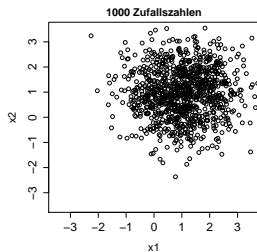
Multivariate Normalverteilung – Beispiele

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N_2 \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right)$$



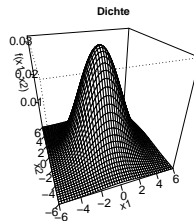
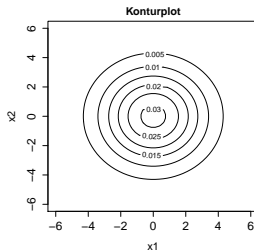
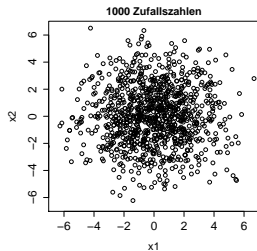
Multivariate Normalverteilung – Beispiele

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N_2 \left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right)$$



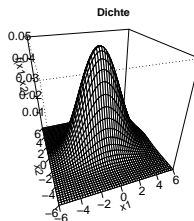
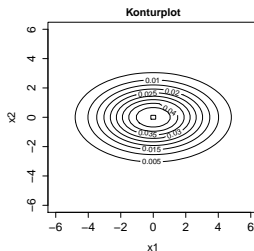
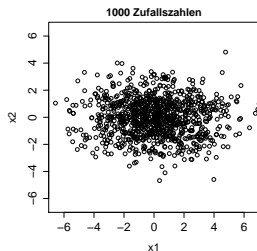
Multivariate Normalverteilung – Beispiele

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N_2 \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix} \right)$$



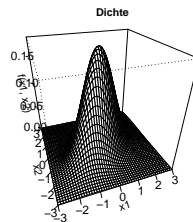
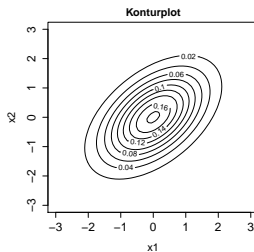
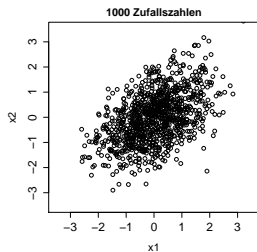
Multivariate Normalverteilung – Beispiele

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 5 & 0 \\ 0 & 2 \end{pmatrix} \right)$$



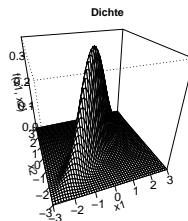
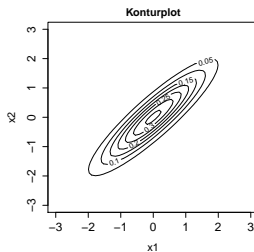
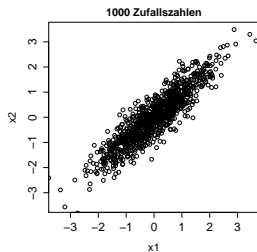
Multivariate Normalverteilung – Beispiele

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N_2 \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix} \right)$$



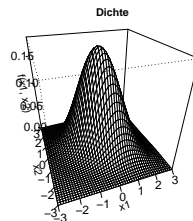
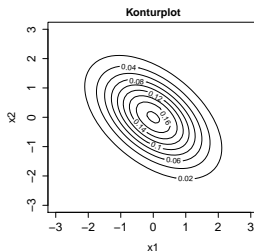
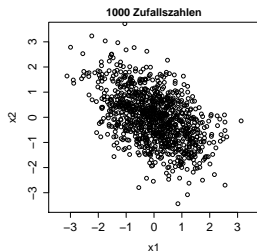
Multivariate Normalverteilung – Beispiele

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N_2 \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0.9 \\ 0.9 & 1 \end{pmatrix} \right)$$



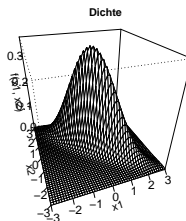
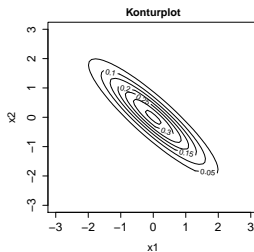
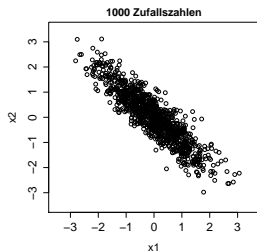
Multivariate Normalverteilung – Beispiele

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N_2 \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & -0.5 \\ -0.5 & 1 \end{pmatrix} \right)$$



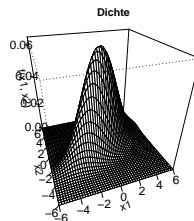
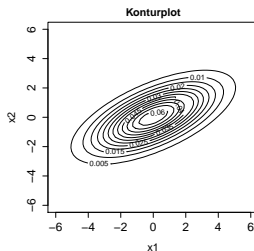
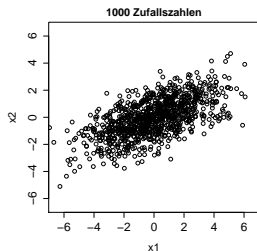
Multivariate Normalverteilung – Beispiele

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N_2 \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & -0.9 \\ -0.9 & 1 \end{pmatrix} \right)$$



Multivariate Normalverteilung – Beispiele

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N_2 \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 5 & 2 \\ 2 & 2 \end{pmatrix} \right)$$



Rand- und bedingte Verteilung

Sei $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \Sigma)$. Betrachte die Partition

$$\mathbf{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_r \\ X_{r+1} \\ \vdots \\ X_p \end{pmatrix} = \begin{pmatrix} \mathbf{Y} \\ \mathbf{Z} \end{pmatrix}$$

und damit

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_r \\ \mu_{r+1} \\ \vdots \\ \mu_p \end{pmatrix} = \begin{pmatrix} \boldsymbol{\mu}_Y \\ \boldsymbol{\mu}_Z \end{pmatrix} \quad \Sigma = \begin{pmatrix} \Sigma_Y & \Sigma_{YZ} \\ \Sigma_{ZY} & \Sigma_Z \end{pmatrix}.$$

Rand- und bedingte Verteilung

Dann gilt:

- ▶ Randverteilungen:

$$\mathbf{Y} \sim N_r(\boldsymbol{\mu}_Y, \Sigma_Y)$$

$$\mathbf{Z} \sim N_q(\boldsymbol{\mu}_Z, \Sigma_Z) \quad \text{mit} \quad q = p - r$$

- ▶ Bedingte Verteilung:

$$\mathbf{Y}|\mathbf{Z} = \mathbf{z}_0 \sim N_r(\boldsymbol{\mu}_{Y|Z}, \Sigma_{Y|Z})$$

$$\text{mit} \quad \boldsymbol{\mu}_{Y|Z} = \boldsymbol{\mu}_Y + \Sigma_{YZ} \cdot \Sigma_Z^{-1}(\mathbf{z}_0 - \boldsymbol{\mu}_Z)$$

$$\Sigma_{Y|Z} = \Sigma_Y - \Sigma_{YZ} \Sigma_Z^{-1} \Sigma_{ZY}$$

- ▶ Unabhängigkeit:

$$\Sigma_{YZ} = 0 \Rightarrow \mathbf{Y} \text{ und } \mathbf{Z} \text{ sind unabhängig}$$

$$\text{also: } \mathbf{Y} \text{ und } \mathbf{Z} \text{ sind unabhängig} \Leftrightarrow \Sigma_{YZ} = 0$$

Weitere multivariate Verteilungen

Wishart-Verteilung

Seien $\mathbf{X}_1, \dots, \mathbf{X}_m \stackrel{i.i.d.}{\sim} N_p(\mathbf{0}, \Sigma)$, dann ist die Matrix

$$\mathbf{M} = \sum_{i=1}^m \mathbf{X}_i \mathbf{X}_i^{\top} = \mathbf{X}^{\top} \mathbf{X} \in \mathbb{R}^{p \times p}$$

wishart-verteilt mit Parametern Σ und m .

- ▶ $\mathbf{M} \sim W_p(\Sigma, m)$
- ▶ Falls $p = 1$, so gilt $\mathbf{M} = \sum_{i=1}^m X_i^2 \sim \sigma^2 \chi^2(m)$, wobei die Komponenten $X_i \sim N(0, \sigma^2)$
- ▶ Multivariate Erweiterung der χ^2 -Verteilung

Hotellings T^2 -Verteilung

Seien $\mathbf{d} \sim N_p(\mathbf{0}, \mathbf{I})$ und $\mathbf{M} \sim W_p(\mathbf{I}, m)$ unabhängig, dann ist

$$u = m \mathbf{d}^\top \mathbf{M}^{-1} \mathbf{d}$$

Hotellings T^2 -verteilt mit Parametern p und m .

- ▶ $u \sim T^2(p, m)$
- ▶ Falls $p = 1$, so ergibt sich die $F(1, m)$ -Verteilung
- ▶ Seien $\mathbf{d} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ und $\mathbf{M} \sim W_p(\boldsymbol{\Sigma}, m)$ unabhängig, dann erhält man:

$$m(\mathbf{d} - \boldsymbol{\mu})^\top \mathbf{M}^{-1} (\mathbf{d} - \boldsymbol{\mu}) \sim T^2(p, m)$$

- ▶ Verwendung bei Testproblemen (multivariate Erweiterung der t-Verteilung bzw. des t-Tests)

Wilks' Λ -Verteilung

Seien $\mathbf{A} \sim W_p(\mathbf{I}, m)$ und $\mathbf{B} \sim W_p(\mathbf{I}, n)$ unabhängig, dann ist

$$\Lambda = \frac{\det(\mathbf{A})}{\det(\mathbf{A} + \mathbf{B})}$$

Wilks' Λ -verteilt mit Parametern p , m und n .

- ▶ $\Lambda \sim \Lambda(p, m, n)$
- ▶ Falls $p = 1$, so gilt $A \sim \chi^2(m)$ und $B \sim \chi^2(n)$ und damit erhält man: $\Lambda \sim B(m/2, n/2)$
- ▶ Verwendung beim Testen im Rahmen der einfaktoriellen Varianzanalyse