



# Multivariate Datenanalyse

MSc Psychologie WiSe 2022/23

Prof. Dr. Dirk Ostwald

## (5) Multivariate Wahrscheinlichkeitstheorie

---

Vorbemerkungen

Multivariate Verteilungen

Marginalverteilungen und Bedingte Verteilungen

Erwartungswert, Kovarianzmatrizen und Korrelationsmatrizen

Selbstkontrollfragen

---

## **Vorbemerkungen**

Multivariate Verteilungen

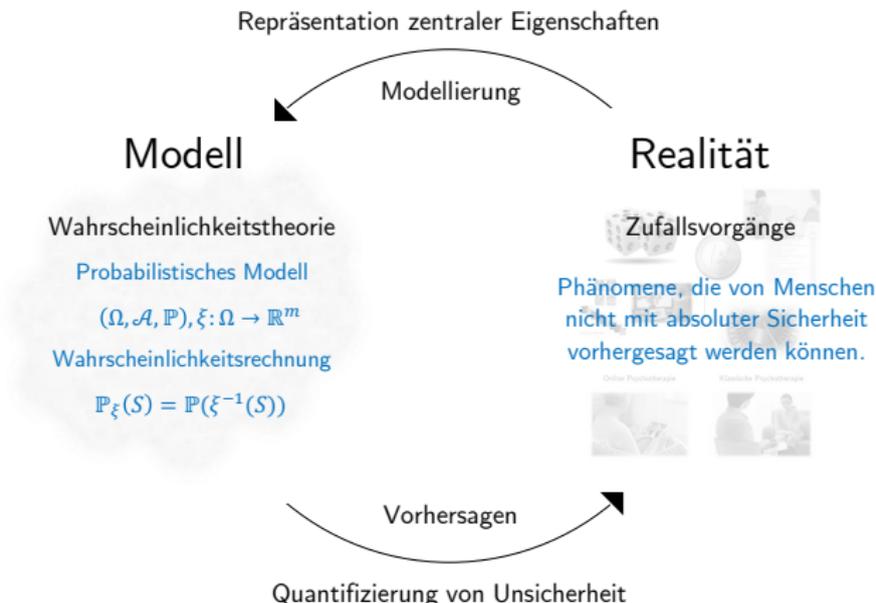
Marginalverteilungen und Bedingte Verteilungen

Erwartungswert, Kovarianzmatrizen und Korrelationsmatrizen

Selbstkontrollfragen

## Modul A1/A3 Forschungsmethoden: Multivariate Verfahren | Themen

Datum	Einheit	Thema
14.10.2022	Grundlagen	(1) Einführung
21.10.2022	Grundlagen	(2) Vektoren
28.10.2022	Grundlagen	(3) Matrizen
04.11.2022	Grundlagen	(4) Eigenanalyse
11.11.2022	Grundlagen	(5) Multivariate Wahrscheinlichkeitstheorie
18.11.2022	Grundlagen	(6) Multivariate Normalverteilungen
25.11.2022	Frequentistische Inferenz	(7) Kanonische Korrelation
02.12.2022	Frequentistische Inferenz	(8) $T^2$ -Tests
09.12.2022	Frequentistische Inferenz	(9) Einfaktorielle MANOVA
16.12.2022	Latente Variablenmodelle	(10) Hauptkomponentenanalyse
	Weihnachtspause	
13.01.2023	Latente Variablenmodelle	(12) Lineare Normalverteilungsmodelle
20.01.2023	Latente Variablenmodelle	(13) Konfirmatorische Faktorenanalyse
27.01.2023	Latente Variablenmodelle	(14) Exploratorische Faktorenanalyse



BSc Psychologie | Grundlagen der Mathematik und Informatik

BSc Psychologie | Wahrscheinlichkeitstheorie und Frequentistische Inferenz

## Realisierungen von Zufallsvariablen

```
# Univariate Normalverteilungsparameter
mu      = 2.0                # Erwartungswertparameter
sigsqr  = 0.5                # Varianzparameter
n       = 10                 # Anzahl von u.i.v. Realisierungen

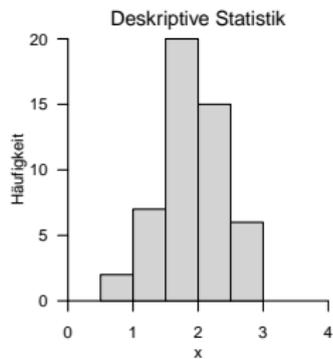
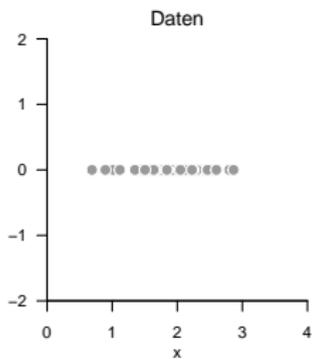
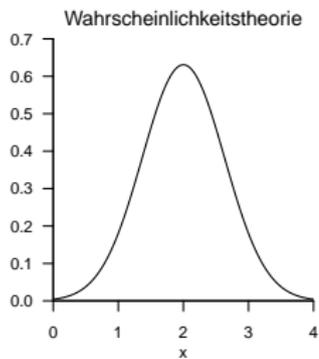
# 10 Realisierungen
X       = rnorm(n,mu,sqrt(sigsqr)) #  $x_{i,i} \sim N(\mu, \sigma^2)$ ,  $i = 1, \dots, n$ 
print(X)

> [1] 1.010 2.181 0.277 1.996 2.440 2.812 0.712 1.825 1.827 1.800

# 10 Realisierungen
X       = rnorm(n,mu,sqrt(sigsqr)) #  $x_{i,i} \sim N(\mu, \sigma^2)$ ,  $i = 1, \dots, n$ 
print(X)

> [1] 1.608 2.445 3.460 0.847 2.362 0.683 1.631 1.963 2.384 1.354
```

## Wahrscheinlichkeitstheorie, Daten, Deskriptive Statistik



## Realisierungen von Zufallsvektoren

```
# R Paket für multivariate Normalverteilungsrealisierung
library(MASS)

# Multivariate Normalverteilungparameter
mu      = c(2.0,5.0)           # Erwartungswertparameter
Sigma   = matrix(c(0.5,0.1,   # Kovarianzmatrixparameter
                  0.1,0.5),
                nrow = 2,
                byrow = TRUE)
n       = 10                  # Anzahl von u.i.v. Realisierungen

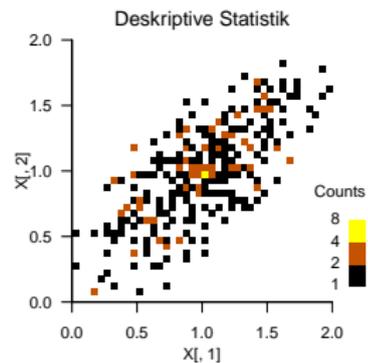
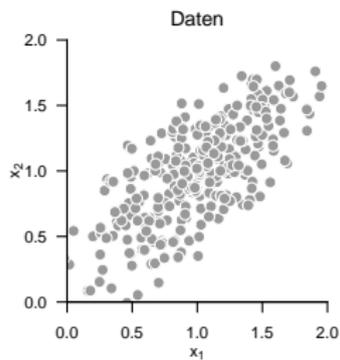
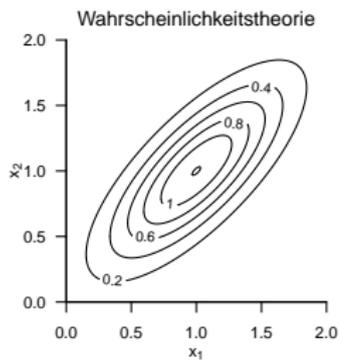
# 10 Realisierungen
X       = t(mvrnorm(n,mu,Sigma)) #  $\{x_{i,j}\} \sim N(\mu, \Sigma), i = 1, \dots, n$ 
print(X)
```

```
>      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10]
> [1,] 1.84 2.12 1.18 1.68 2.98 2.01 2.34 2.12 1.69 0.694
> [2,] 5.67 5.28 4.39 6.13 6.08 4.89 3.63 4.87 4.41 6.649
```

```
# 10 Realisierungen
X       = t(mvrnorm(n,mu,Sigma)) #  $\{x_{i,j}\} \sim N(\mu, \sigma^2), i = 1, \dots, n$ 
print(X)
```

```
>      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10]
> [1,] 1.65 1.76 2.06 1.85 2.23 3.26 3.04 1.12 2.23 0.857
> [2,] 5.42 4.54 4.88 4.87 5.26 6.76 4.01 6.52 4.90 4.048
```

## Multivariate Wahrscheinlichkeitstheorie, Multivariate Daten, Multivariate Deskriptive Statistik



---

Vorbemerkungen

## **Multivariate Verteilungen**

Marginalverteilungen und Bedingte Verteilungen

Erwartungswert, Kovarianzmatrizen und Korrelationsmatrizen

Selbstkontrollfragen

## Definition (Wahrscheinlichkeitsraum)

Ein *Wahrscheinlichkeitsraum* ist ein Triple  $(\Omega, \mathcal{A}, \mathbb{P})$ , wobei

- $\Omega$  eine beliebige nichtleere Menge von *Ergebnissen*  $\omega$  ist und *Ergebnismenge* heißt,
- $\mathcal{A}$  eine  $\sigma$ -*Algebra* auf  $\Omega$  ist und *Ereignissystem* heißt,
- $\mathbb{P}$  eine Abbildung der Form  $\mathbb{P} : \mathcal{A} \rightarrow [0, 1]$  mit den Eigenschaften
  - *Nicht-Negativität*  $\mathbb{P}(A) \geq 0$  für alle  $A \in \mathcal{A}$ ,
  - *Normiertheit*  $\mathbb{P}(\Omega) = 1$  und
  - $\sigma$ -*Additivität*  $\mathbb{P}(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$  für paarweise disjunkte  $A_i \in \mathcal{A}$ist und *Wahrscheinlichkeitsmaß* heißt.

Das Tuple  $(\Omega, \mathcal{A})$  aus Ergebnismenge und Ereignissystem wird als *Messraum* bezeichnet.

Die stillschweigende Mechanik des Wahrscheinlichkeitsraummodells

- Wir stellen uns sequentielle *Durchgänge* eines *Zufallsvorgangs* vor.
- In jedem Durchgang wird genau ein  $\omega$  aus  $\Omega$  mit Wahrscheinlichkeit  $\mathbb{P}(\{\omega\})$  *realisiert*.
- $\mathbb{P}(\{\omega\})$  bestimmt, mit welcher Wahrscheinlichkeit  $\omega$  in einem Durchgang aus  $\Omega$  realisiert wird.
- Wenn  $\omega \in A$  für  $A \in \mathcal{A}$  sagt man, dass das Ereignis  $A$  eingetreten ist.
- Z.B. gilt beim Würfel für  $\omega = 2$  mit  $\omega \in \{2, 4, 6\}$ , dass das Ereignis "Eine gerade Zahl fällt" eingetreten ist.

## Definition (Zufallsvektor)

$(\Omega, \mathcal{A}, \mathbb{P})$  sei ein Wahrscheinlichkeitsraum und  $(\mathcal{X}, \mathcal{S})$  sei ein  $m$ -dimensionaler Messraum. Ein  $m$ -dimensionaler *Zufallsvektor* ist definiert als eine Abbildung

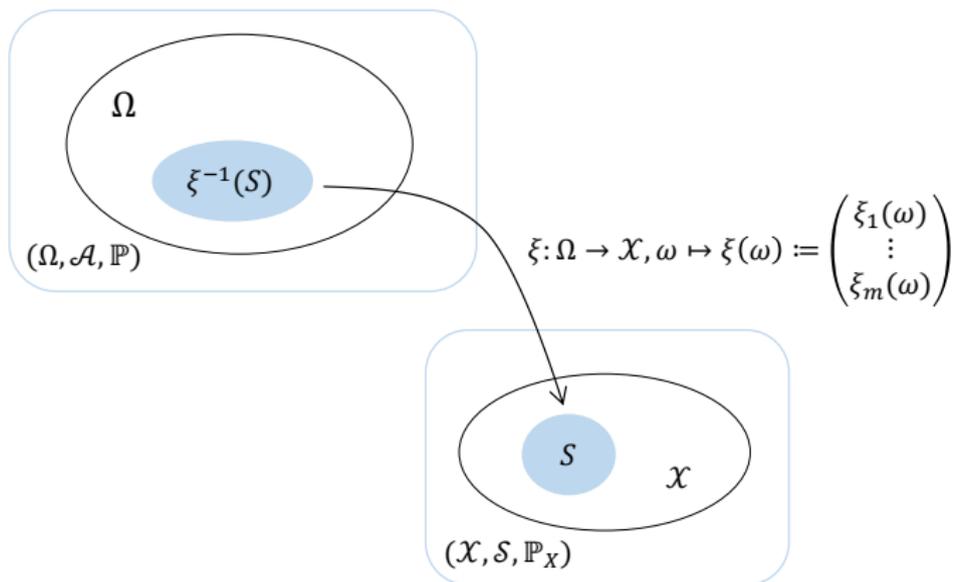
$$\xi : \Omega \rightarrow \mathcal{X}, \omega \mapsto \xi(\omega) := \begin{pmatrix} \xi_1(\omega) \\ \vdots \\ \xi_m(\omega) \end{pmatrix} \quad (1)$$

mit der *Messbarkeitseigenschaft*

$$\{\omega \in \Omega \mid \xi(\omega) \in S\} \in \mathcal{A} \text{ für alle } S \in \mathcal{S}. \quad (2)$$

### Bemerkungen

- $\xi$  ist hier eine univariate, vektorwertige Abbildung.
- Das Standardbeispiel für  $(\mathcal{X}, \mathcal{S})$  ist  $(\mathbb{R}^m, \mathcal{B}(\mathbb{R}^m))$ .
- Wir verzichten auf eine explizite Einführung  $m$ -dimensionaler  $\sigma$ -Algebren wie  $\mathcal{B}(\mathbb{R}^m)$ .
- Ohne Beweis halten wir fest, dass  $\xi$  messbar ist, wenn die Funktionen  $\xi_1, \dots, \xi_m$  messbar sind.
- Die Komponentenfunktionen eines Zufallsvektors sind Zufallsvariablen.
- Ein  $m$ -dimensionaler Zufallsvektor ist die Konkatenation von  $m$  Zufallsvariablen.
- Für  $m := 1$  ist ein Zufallsvektor eine Zufallsvariable.
- Für einen Zufallsvektor schreiben wir auch häufig  $\xi := (\xi_1, \dots, \xi_m)$ .



$$\mathbb{P}(\xi^{-1}(S)) = \mathbb{P}(\{\omega \in \Omega \mid \xi(\omega) \in S\}) =: \mathbb{P}_{\xi}(S)$$

## Definition (Multivariate Verteilung)

$(\Omega, \mathcal{A}, \mathbb{P})$  sei ein Wahrscheinlichkeitsraum,  $(\mathcal{X}, \mathcal{S})$  sei ein  $m$ -dimensionaler Messraum und

$$\xi : \Omega \rightarrow \mathcal{X}, \omega \mapsto \xi(\omega) \quad (3)$$

sei ein Zufallsvektor. Dann heißt das Wahrscheinlichkeitsmaß  $\mathbb{P}_\xi$ , definiert durch

$$\mathbb{P}_\xi : \mathcal{S} \rightarrow [0, 1], S \mapsto \mathbb{P}_\xi(S) := \mathbb{P}(\xi^{-1}(S)) = \mathbb{P}(\{\omega \in \Omega \mid \xi(\omega) \in S\}) \quad (4)$$

die *multivariate Verteilung des Zufallsvektor*  $\xi$ .

### Bemerkungen

- Der Einfachheit halber spricht man oft auch nur von "der Verteilung des Zufallsvektors  $\xi$ ".
- Die Notationskonventionen für Zufallsvariablen gelten für Zufallsvektoren analog, z.B.

$$\begin{aligned} \mathbb{P}_\xi(\xi \in S) &:= \mathbb{P}(\{\xi \in S\}) = \mathbb{P}(\{\omega \in \Omega \mid \xi(\omega) \in S\}) \\ \mathbb{P}_\xi(\xi = x) &:= \mathbb{P}(\{\xi = x\}) = \mathbb{P}(\{\omega \in \Omega \mid \xi(\omega) = x\}) \\ \mathbb{P}_\xi(\xi \leq x) &:= \mathbb{P}(\{\xi \leq x\}) = \mathbb{P}(\{\omega \in \Omega \mid \xi(\omega) \leq x\}) \end{aligned} \quad (5)$$

$$\mathbb{P}_\xi(x_1 \leq \xi \leq x_2) := \mathbb{P}(\{x_1 \leq \xi \leq x_2\}) = \mathbb{P}(\{\omega \in \Omega \mid x_1 \leq \xi(\omega) \leq x_2\})$$

- Relationsoperatoren wie  $\leq$  werden hier *komponentenweise* verstanden.
- Zum Beispiel heißt  $x \leq y$  für  $x, y \in \mathbb{R}^m$ , dass  $x_i \leq y_i$  für alle  $i = 1, \dots, m$ .

## Definition (Multivariate kumulative Verteilungsfunktionen)

$\xi$  sei ein Zufallsvektor mit Ergebnisraum  $\mathcal{X}$ . Dann heißt eine Funktion der Form

$$P_\xi : \mathcal{X} \rightarrow [0, 1], x \mapsto P_\xi(x) := \mathbb{P}_\xi(\xi \leq x) \quad (6)$$

multivariate kumulative Verteilungsfunktion von  $\xi$ .

### Bemerkung

- Multivariate kumulative Verteilungsfunktionen können zur Definition von multivariaten Verteilungen genutzt werden, häufiger ist allerdings die Definition multivariater Verteilungen durch multivariate Wahrscheinlichkeitsmasse- oder Wahrscheinlichkeitsdichtefunktionen.

## Definition (Diskreter Zufallsvektor, Multivariate WMF)

$\xi$  sei ein Zufallsvektor mit Ergebnisraum  $\mathcal{X}$ .  $\xi$  heißt *diskreter Zufallsvektor* wenn der Ergebnisraum  $\mathcal{X}$  endlich oder abzählbar ist und eine Funktion der Form

$$p_\xi : \mathcal{X} \rightarrow [0, 1], x \mapsto p_\xi(x) \quad (7)$$

existiert, für die gilt

- (1)  $\sum_{x \in \mathcal{X}} p_\xi(x) = 1$  und
- (2)  $\mathbb{P}_\xi(\xi = x) = p_\xi(x)$  für alle  $x \in \mathcal{X}$ .

Ein entsprechende Funktion  $p$  heißt *multivariate Wahrscheinlichkeitsmassenfunktion (WMF)* von  $\xi$ .

### Bemerkungen

- Der Begriff der multivariaten WMF ist analog zum Begriff der WMF.
- Man spricht oft einfach von der WMF eines Zufallsvektors.
- Wie univariate WMFen sind multivariate WMFen nicht-negativ und normiert.

## Beispiel (Multivariate Wahrscheinlichkeitsmassefunktion)

Wir betrachten einen zweidimensionalen Zufallsvektor  $\xi := (\xi_1, \xi_2)$  der Werte in  $\mathcal{X} := \mathcal{X}_1 \times \mathcal{X}_2$  annimmt, wobei  $\mathcal{X}_1 := \{1, 2, 3\}$  und  $\mathcal{X}_2 = \{1, 2, 3, 4\}$  seien.

Eine exemplarische bivariate WMF der Form

$$p_\xi : \{1, 2, 3\} \times \{1, 2, 3, 4\} \rightarrow [0, 1], (x_1, x_2) \mapsto p_\xi(x_1, x_2) \quad (8)$$

ist dann durch nachfolgende Tabelle definiert:

$p_\xi(x_1, x_2)$	$x_2 = 1$	$x_2 = 2$	$x_2 = 3$	$x_2 = 4$
$x_1 = 1$	0.1	0.0	0.2	0.1
$x_1 = 2$	0.1	0.2	0.0	0.0
$x_1 = 3$	0.0	0.1	0.1	0.1

Man beachte, dass  $\sum_{x \in \mathcal{X}} p_\xi(x) = \sum_{x_1=1}^3 \sum_{x_2=1}^4 p_\xi(x_1, x_2) = 1$ .

## Definition (Kontinuierlicher Zufallsvektor, Multivariate WDF)

Ein Zufallsvektor  $\xi$  heißt *kontinuierlich*, wenn  $\mathbb{R}^m$  der Ergebnisraum von  $\xi$  ist und eine Funktion

$$p_\xi : \mathbb{R}^m \rightarrow \mathbb{R}_{\geq 0}, x \mapsto p_\xi(x), \quad (9)$$

existiert, für die gilt

$$(1) \int_{\mathbb{R}^m} p_\xi(x) dx = 1 \text{ und}$$

$$(2) \mathbb{P}_\xi(x_1 \leq \xi \leq x_2) = \int_{x_{11}}^{x_{21}} \cdots \int_{x_{1m}}^{x_{2m}} p_\xi(s_1, \dots, s_m) ds_1 \cdots ds_m.$$

Eine entsprechende Funktion  $p$  heißt *multivariate Wahrscheinlichkeitsdichtefunktion (WDF)* von  $\xi$ .

### Bemerkungen

- Der Begriff der multivariaten WDF ist analog zum Begriff der WDF.
- Man spricht häufig auch einfach von der WDF eines Zufallsvektors
- Wie univariate WDFen sind multivariate WDFen nicht-negativ und normiert.
- Wie für kontinuierliche Zufallsvariablen gilt für kontinuierliche Zufallsvektoren

$$\mathbb{P}_\xi(\xi = x) = \mathbb{P}_\xi(x \leq \xi \leq x) = \int_{x_1}^{x_1} \cdots \int_{x_m}^{x_m} p_\xi(s_1, \dots, s_m) ds_1 \cdots ds_m = 0 \quad (10)$$

- In (6) Multivariate Normalverteilungen befassen wir uns mit dem Standardbeispiel für multivariate WDFen.

---

Vorbemerkungen

Multivariate Verteilungen

**Marginalverteilungen und Bedingte Verteilungen**

Erwartungswert, Kovarianzmatrizen und Korrelationsmatrizen

Selbstkontrollfragen

## Definition (Univariate Marginalverteilung)

$(\Omega, \mathcal{A}, \mathbb{P})$  sei ein Wahrscheinlichkeitsraum,  $(\mathcal{X}, \mathcal{S})$  sei ein  $m$ -dimensionaler Messraum,  $\xi : \Omega \rightarrow \mathcal{X}$  sei ein Zufallsvektor,  $\mathbb{P}_\xi$  sei die Verteilung von  $\xi$ ,  $\mathcal{X}_i \subset \mathcal{X}$  sei der Ergebnisraum der  $i$ ten Komponente  $\xi_i$  von  $\xi$ , und  $\mathcal{S}_i$  sei eine  $\sigma$ -Algebra auf  $\mathcal{X}_i$ . Dann heißt die durch

$$\mathbb{P}_{\xi_i} : \mathcal{S}_i \rightarrow [0, 1], S \mapsto \mathbb{P}_\xi (\mathcal{X}_1 \times \cdots \times \mathcal{X}_{i-1} \times S \times \mathcal{X}_{i+1} \times \cdots \times \mathcal{X}_m) \text{ für } S \in \mathcal{S}_i \quad (11)$$

definierte Verteilung die *ite univariate Marginalverteilung* von  $\xi$ .

## Bemerkungen

- Univariate Marginalverteilungen sind die Verteilungen der Komponenten eines Zufallsvektors.
- Univariate Marginalverteilungen sind Verteilungen von Zufallsvariablen.
- Die Festlegung der multivariaten Verteilung von  $\xi$  legt auch die Verteilungen der  $\xi_i$  fest.

## Theorem (Marginale Wahrscheinlichkeitsmasse- und dichtefunktionen)

(1)  $\xi = (\xi_1, \dots, \xi_m)$  sei ein  $m$ -dimensionaler diskreter Zufallsvektor mit Wahrscheinlichkeitsmassenfunktion  $p_\xi$  und Komponentenergebnisräumen  $\mathcal{X}_1, \dots, \mathcal{X}_m$ . Dann ergibt sich die Wahrscheinlichkeitsmassenfunktion der  $i$ ten Komponente  $\xi_i$  von  $\xi$  als

$$p_{\xi_i} : \mathcal{X}_i \rightarrow [0, 1], x_i \mapsto p_{\xi_i}(x_i) := \sum_{x_1} \cdots \sum_{x_{i-1}} \sum_{x_{i+1}} \cdots \sum_{x_m} p_\xi(x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_m). \quad (12)$$

(2)  $\xi = (\xi_1, \dots, \xi_m)$  sei ein  $m$ -dimensionaler kontinuierlicher Zufallsvektor mit Wahrscheinlichkeitsdichtefunktion  $p_\xi$  und Komponentenergebnisraum  $\mathbb{R}$ . Dann ergibt sich die Wahrscheinlichkeitsdichtefunktion der  $i$ ten Komponente  $\xi_i$  von  $\xi$  als

$$p_{\xi_i} : \mathbb{R} \rightarrow \mathbb{R}_{\geq 0}, x_i \mapsto p_{\xi_i}(x_i) := \int_{x_1} \cdots \int_{x_{i-1}} \int_{x_{i+1}} \cdots \int_{x_m} p_\xi(x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_m) dx_1 \cdots dx_{i-1} dx_{i+1} \cdots dx_m. \quad (13)$$

### Bemerkungen

- Wir verzichten auf einen Beweis.
- Die WMFen der univariaten Marginalverteilungen diskreter Zufallsvektoren ergeben sich durch Summation.
- Die WDFen der univariaten Marginalverteilungen kontinuierlicher Zufallsvektoren ergeben sich durch Integration.

## Beispiel (Marginale Wahrscheinlichkeitsmassfunktionen)

Wir betrachten erneut den zweidimensionalen Zufallsvektor  $\xi := (\xi_1, \xi_2)$  der Werte in  $\mathcal{X} := \mathcal{X}_1 \times \mathcal{X}_2$  annimmt, wobei  $\mathcal{X}_1 := \{1, 2, 3\}$  und  $\mathcal{X}_2 = \{1, 2, 3, 4\}$  seien.

Basierend auf der oben definierten WMF ergeben sich folgende marginale WMFen  $p_{\xi_1}$  und  $p_{\xi_2}$ :

$p_{\xi}(x_1, x_2)$	$x_2 = 1$	$x_2 = 2$	$x_2 = 3$	$x_2 = 4$	$p_{\xi_1}(x_1)$
$x_1 = 1$	0.1	0.0	0.2	0.1	0.4
$x_1 = 2$	0.1	0.2	0.0	0.0	0.3
$x_1 = 3$	0.0	0.1	0.1	0.1	0.3
$p_{\xi_2}(x_2)$	0.2	0.3	0.3	0.2	

Man beachte, dass  $\sum_{x_1=1}^3 p_{\xi_1}(x_1) = 1$  und  $\sum_{x_2=1}^4 p_{\xi_2}(x_2) = 1$  gilt.

## Vorbemerkungen zu Bedingten Verteilungen

Wir erinnern uns, dass für einen Wahrscheinlichkeitsraum  $(\Omega, \mathcal{A}, \mathbb{P})$  und zwei Ereignisse  $A, B \in \mathcal{A}$  mit  $\mathbb{P}(B) > 0$  die bedingte Wahrscheinlichkeit von  $A$  gegeben  $B$  definiert ist als

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}. \quad (14)$$

Analog wird für zwei Zufallsvariablen  $\xi_1, \xi_2$  mit Ereignisräumen  $\mathcal{X}_1, \mathcal{X}_2$  und (messbaren) Mengen  $S_1 \in \mathcal{X}_1, S_2 \in \mathcal{X}_2$  die bedingte Verteilung von  $\xi_1$  gegeben  $\xi_2$  mithilfe der Ereignisse

$$A := \{\xi_1 \in S_1\} \text{ und } B := \{\xi_2 \in S_2\} \quad (15)$$

definiert.

So ergibt sich zum Beispiel die bedingte Wahrscheinlichkeit, dass  $\xi_1 \in S_1$  gegeben dass  $\xi_2 \in S_2$  unter der Annahme, dass  $\mathbb{P}(\{\xi_2 \in S_2\}) > 0$ , zu

$$\mathbb{P}(\{\xi_1 \in S_1\}|\{\xi_2 \in S_2\}) = \frac{\mathbb{P}(\{\xi_1 \in S_1\} \cap \{\xi_2 \in S_2\})}{\mathbb{P}(\{\xi_2 \in S_2\})}. \quad (16)$$

Wir betrachten zunächst durch WMFen/WDFen zweidimensionaler Zufallsvektoren definierte bedingte Verteilungen.

## Definition (Bedingte WMF, diskrete bedingte Verteilung)

$\xi := (\xi_1, \xi_2)$  sei ein diskreter Zufallsvektor mit  $m + k$  dimensionalen Ergebnisraum  $\mathcal{X} := \mathcal{X}_1 \times \mathcal{X}_2$ , WMF  $p_\xi = p_{\xi_1, \xi_2}$  und marginalen WMFen  $p_{\xi_1}$  und  $p_{\xi_2}$ . Die bedingte WMF von  $\xi_1$  gegeben  $\xi_2 = x_2$  ist dann für  $p_{\xi_2}(x_2) > 0$  definiert als

$$p_{\xi_1|\xi_2=x_2} : \mathcal{X}_1 \rightarrow [0, 1], x_1 \mapsto p_{\xi_1|\xi_2=x_2}(x_1|x_2) := \frac{p_{\xi_1, \xi_2}(x_1, x_2)}{p_{\xi_2}(x_2)} \quad (17)$$

Analog ist für  $p_{\xi_1}(x_1) > 0$  die bedingte WMF von  $\xi_2$  gegeben  $\xi_1 = x_1$  definiert als

$$p_{\xi_2|\xi_1=x_1} : \mathcal{X}_2 \rightarrow [0, 1], x_2 \mapsto p_{\xi_2|\xi_1=x_1}(x_2|x_1) := \frac{p_{\xi_1, \xi_2}(x_1, x_2)}{p_{\xi_1}(x_1)} \quad (18)$$

Die bedingten Verteilungen mit WMFen  $p_{\xi_1|\xi_2=x_2}$  und  $p_{\xi_2|\xi_1=x_1}$  heißen dann die *diskreten bedingten Verteilungen von  $\xi_1$  gegeben  $\xi_2 = x_2$  und  $\xi_2$  gegeben  $\xi_1 = x_1$* , respektive.

### Bemerkungen

- In Analogie zur Definition der bedingten Wahrscheinlichkeit von Ereignissen gilt also

$$p_{\xi_1|\xi_2}(x_1|x_2) = \frac{p_{\xi_1, \xi_2}(x_1, x_2)}{p_{\xi_2}(x_2)} = \frac{\mathbb{P}(\{\xi_1 = x_1\} \cap \{\xi_2 = x_2\})}{\mathbb{P}(\{\xi_2 = x_2\})}. \quad (19)$$

- Bedingte Verteilungen sind (lediglich) normalisierte gemeinsame Verteilungen.

## Beispiel (Bedingte Wahrscheinlichkeitsmassenfunktionen)

Wir betrachten erneut den zweidimensionalen Zufallsvektor  $\xi := (\xi_1, \xi_2)$  der Werte in  $\mathcal{X} := \mathcal{X}_1 \times \mathcal{X}_2$  annimmt, wobei  $\mathcal{X}_1 := \{1, 2, 3\}$  und  $\mathcal{X}_2 = \{1, 2, 3, 4\}$  seien.

Basierend auf der oben definierten WMF und den entsprechenden oben evaluierten marginalen WMFen ergeben sich folgende bedingte WMFen für  $p_{\xi_2|\xi_1=x_1}$

$p_{\xi_2 \xi_1}(x_2 x_1)$	$x_2 = 1$	$x_2 = 2$	$x_2 = 3$	$x_2 = 4$
$p_{\xi_2 \xi_1=1}(x_2 x_1 = 1)$	$\frac{0.1}{0.4} = 0.25$	$\frac{0.0}{0.4} = 0.00$	$\frac{0.2}{0.4} = 0.50$	$\frac{0.1}{0.4} = 0.25$
$p_{\xi_2 \xi_1=2}(x_2 x_1 = 2)$	$\frac{0.1}{0.3} = 0.3\bar{3}$	$\frac{0.2}{0.3} = 0.6\bar{6}$	$\frac{0.0}{0.3} = 0.00$	$\frac{0.0}{0.3} = 0.00$
$p_{\xi_2 \xi_1=3}(x_2 x_1 = 3)$	$\frac{0.0}{0.3} = 0.00$	$\frac{0.1}{0.3} = 0.3\bar{3}$	$\frac{0.1}{0.3} = 0.3\bar{3}$	$\frac{0.1}{0.3} = 0.3\bar{3}$

### Bemerkungen

- Man beachte, dass  $\sum_{x_2=1}^4 p_{\xi_2|\xi_1=x_1}(x_2|x_1) = 1$  für alle  $x_1 \in \mathcal{X}_1$ .
- Man beachte die qualitative Ähnlichkeit der WMFen  $p_{\xi_1, \xi_2}(x_1, x_2)$  und  $p_{\xi_2|\xi_1}(x_2|x_1)$ .
- Bedingte Verteilungen sind (lediglich) normalisierte gemeinsame Verteilungen.

## Definition (Bedingte WDF, kontinuierliche bedingte Verteilungen)

$\xi := (\xi_1, \xi_2)$  sei ein kontinuierlicher Zufallsvektor mit Ergebnisraum  $\mathbb{R}^{m+k}$ , WDF  $p_\xi = p_{\xi_1, \xi_2}$  und marginalen WDFen  $p_{\xi_1}$  und  $p_{\xi_2}$ . Die bedingte WDF von  $\xi_1$  gegeben  $\xi_2 = x_2$  ist dann für  $p_{\xi_2}(x_2) > 0$  definiert als

$$p_{\xi_1|\xi_2=x_2} : \mathbb{R} \rightarrow \mathbb{R}_{\geq 0}, x_1 \mapsto p_{\xi_1|\xi_2=x_2}(x_1|x_2) := \frac{p_{\xi_1, \xi_2}(x_1, x_2)}{p_{\xi_2}(x_2)} \quad (20)$$

Analog ist für  $p_{\xi_1}(x_1) > 0$  die bedingte WMF von  $\xi_2$  gegeben  $\xi_1 = x_1$  definiert als

$$p_{\xi_2|\xi_1=x_1} : \mathbb{R} \rightarrow \mathbb{R}_{\geq 0}, x_2 \mapsto p_{\xi_2|\xi_1=x_1}(x_2|x_1) := \frac{p_{\xi_1, \xi_2}(x_1, x_2)}{p_{\xi_1}(x_1)} \quad (21)$$

Die Verteilungen mit WDFen  $p_{\xi_1|\xi_2=x_2}$  und  $p_{\xi_2|\xi_1=x_1}$  heißen dann die *kontinuierlichen bedingten Verteilungen* von  $\xi_1$  gegeben  $\xi_2 = x_2$  und  $\xi_2$  gegeben  $\xi_1 = x_1$ , respektive.

### Bemerkung

- Im kontinuierlichen Fall gilt zwar  $\mathbb{P}_\xi(\xi = x) = 0$ , aber nicht notwendig auch  $p_\xi(x) = 0$ .
- In Einheit (6) Multivariate Normalverteilung beschäftigen wir uns eingehend mit einem Beispiel.

---

Vorbemerkungen

Multivariate Verteilungen

Marginalverteilungen und Bedingte Verteilungen

**Erwartungswert, Kovarianzmatrizen und Korrelationsmatrizen**

Selbstkontrollfragen

## Definition (Erwartungswert, Varianz, Kovarianz, Korrelation)

$\xi$  und  $v$  seien Zufallsvariablen auf einem Wahrscheinlichkeitsraum  $(\Omega, \mathcal{A}, \mathbb{P})$ .

- Der *Erwartungswert* von  $\xi$  definiert als

- $\mathbb{E}(\xi) := \sum_{x \in \mathcal{X}} x p_{\xi}(x)$ , wenn  $\xi : \Omega \rightarrow \mathcal{X}$  diskret mit WMF  $p_{\xi}$  ist, und als

- $\mathbb{E}(\xi) := \int_{-\infty}^{\infty} x p_{\xi}(x) dx$ , wenn  $\xi : \Omega \rightarrow \mathbb{R}$  kontinuierlich mit WDF  $p_{\xi}$  ist.

- Die *Varianz* von  $\xi$  ist definiert als

$$\mathbb{V}(\xi) := \mathbb{E} \left( (\xi - \mathbb{E}(\xi))^2 \right). \quad (22)$$

- Die *Kovarianz* von  $\xi$  und  $v$  definiert als

$$\mathbb{C}(\xi, v) := \mathbb{E} \left( (\xi - \mathbb{E}(\xi)) (v - \mathbb{E}(v)) \right). \quad (23)$$

- Die *Korrelation* von  $\xi$  und  $v$  ist definiert als

$$\rho(\xi, v) := \frac{\mathbb{C}(\xi, v)}{\sqrt{\mathbb{V}(\xi)} \sqrt{\mathbb{V}(v)}}. \quad (24)$$

### Bemerkung

- Es gilt  $\mathbb{C}(\xi, \xi) = \mathbb{V}(\xi)$ .
- Für eine ausführliche Diskussion der Begriffe, siehe [Wahrscheinlichkeitstheorie und Frequentistische Inferenz](#).
- Die Begriffe sind nicht identisch mit den Begriffen des Stichprobenmittels und der Stichproben(ko)varianz.
- Die Begriffe sind nicht identisch mit den Normalverteilungsparametern.

## Definition (Erwartungswert eines Zufallsvektors)

$\xi$  sei ein  $m$ -dimensionaler Zufallsvektor. Dann ist der *Erwartungswert* von  $\xi$  definiert als

$$\mathbb{E}(\xi) := \begin{pmatrix} \mathbb{E}(\xi_1) \\ \vdots \\ \mathbb{E}(\xi_m) \end{pmatrix}. \quad (25)$$

### Bemerkungen

- Der Erwartungswert von  $\xi$  ist der Vektor der Erwartungswerte  $\mathbb{E}(\xi_1), \dots, \mathbb{E}(\xi_m)$ .
- Im ALM  $y = X\beta + \varepsilon$  ist zum Beispiel  $\mathbb{E}(\varepsilon) = 0_n$  und  $\mathbb{E}(y) = X\beta \in \mathbb{R}^n$ .

## Definition (Kovarianzmatrix)

$\xi$  sei ein  $m$ -dimensionaler Zufallsvektor. Dann ist die *Kovarianzmatrix* von  $\xi$  definiert als

$$\mathbb{C}(\xi) := \mathbb{E} \left( (\xi - \mathbb{E}(\xi))(\xi - \mathbb{E}(\xi))^T \right) = \begin{pmatrix} \mathbb{C}(\xi_1, \xi_1) & \cdots & \mathbb{C}(\xi_1, \xi_m) \\ \vdots & \ddots & \vdots \\ \mathbb{C}(\xi_m, \xi_1) & \cdots & \mathbb{C}(\xi_m, \xi_m) \end{pmatrix}. \quad (26)$$

### Bemerkungen

- Die Kovarianzmatrix  $\mathbb{C}(\xi)$  ist die  $m \times m$  Matrix der Kovarianzen  $\mathbb{C}(\xi_i, \xi_j)$ ,  $i, j = 1, \dots, m$ .
- Im ALM mit sphärischer Kovarianzmatrix gilt per Definition  $\mathbb{C}(\varepsilon) = \mathbb{C}(y) = \sigma^2 I_n \in \mathbb{R}^{n \times n}$ .
- Die Äquivalenz der Kovarianzmatrixschreibweisen folgt wie im Folgenden dargestellt.

# Erwartungswert, Kovarianzmatrizen und Korrelationsmatrizen

Bemerkungen (fortgeführt)

$$\begin{aligned} & \mathbb{E} \left( (\xi - \mathbb{E}(\xi))(\xi - \mathbb{E}(\xi))^T \right) \\ &= \mathbb{E} \left( \left( \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_m \end{pmatrix} - \begin{pmatrix} \mathbb{E}(\xi_1) \\ \vdots \\ \mathbb{E}(\xi_m) \end{pmatrix} \right) \left( \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_m \end{pmatrix} - \begin{pmatrix} \mathbb{E}(\xi_1) \\ \vdots \\ \mathbb{E}(\xi_m) \end{pmatrix} \right)^T \right) \\ &= \mathbb{E} \left( \begin{pmatrix} \xi_1 - \mathbb{E}(\xi_1) \\ \vdots \\ \xi_m - \mathbb{E}(\xi_m) \end{pmatrix} \begin{pmatrix} \xi_1 - \mathbb{E}(\xi_1) \\ \vdots \\ \xi_m - \mathbb{E}(\xi_m) \end{pmatrix}^T \right) \\ &= \mathbb{E} \left( \begin{pmatrix} \xi_1 - \mathbb{E}(\xi_1) \\ \vdots \\ \xi_m - \mathbb{E}(\xi_m) \end{pmatrix} \begin{pmatrix} \xi_1 - \mathbb{E}(\xi_1) & \dots & \xi_m - \mathbb{E}(\xi_m) \end{pmatrix} \right) \tag{27} \\ &= \mathbb{E} \begin{pmatrix} (\xi_1 - \mathbb{E}(\xi_1))(\xi_1 - \mathbb{E}(\xi_1)) & \dots & (\xi_1 - \mathbb{E}(\xi_1))(\xi_m - \mathbb{E}(\xi_m)) \\ \vdots & \ddots & \vdots \\ (\xi_m - \mathbb{E}(\xi_m))(\xi_1 - \mathbb{E}(\xi_1)) & \dots & (\xi_m - \mathbb{E}(\xi_m))(\xi_m - \mathbb{E}(\xi_m)) \end{pmatrix} \\ &= \left( \mathbb{E} \left( (\xi_i - \mathbb{E}(\xi_i))(\xi_j - \mathbb{E}(\xi_j)) \right) \right)_{1 \leq i, j \leq m} \\ &=: \left( \mathbb{C}(\xi_i, \xi_j) \right)_{1 \leq i, j \leq m} \\ &=: \mathbb{C}(\xi). \end{aligned}$$

## Definition (Korrelationsmatrix)

$\xi$  sei ein  $m$ -dimensionaler Zufallsvektor. Dann ist die *Korrelationsmatrix* von  $\xi$  definiert als

$$\mathbb{R}(\xi) := (\rho_{ij})_{1 \leq i, j \leq m} = \left( \frac{\mathbb{C}(\xi_i, \xi_j)}{\sqrt{\mathbb{V}(\xi_i)} \sqrt{\mathbb{V}(\xi_j)}} \right)_{1 \leq i, j \leq m} \quad (28)$$

### Bemerkungen

- Es gelten  $\rho_{ij} \in [-1, 1]$  für  $1 \leq i, j \in m$  und  $\rho_{ii} = 1$  für  $1 \leq i \leq m$ .

## Definition (Stichprobenmittel, -kovarianzmatrix, -korrelationsmatrix)

$v^{(1)}, \dots, v^{(n)}$  seien  $m$ -dimensionale Zufallsvektoren.

- Das *Stichprobenmittel* der  $v^{(1)}, \dots, v^{(n)}$  ist definiert als

$$\bar{v} := \frac{1}{n} \sum_{i=1}^n v^{(i)}. \quad (29)$$

- Die *Stichprobenkovarianzmatrix* der  $v^{(1)}, \dots, v^{(n)}$  ist definiert als

$$C_v := \frac{1}{n-1} \sum_{i=1}^n (v^{(i)} - \bar{v})(v^{(i)} - \bar{v})^T. \quad (30)$$

- Die *Stichprobenkorrelationsmatrix* der  $v^{(1)}, \dots, v^{(m)}$  definiert als

$$R_v := \left( \frac{(C_v)_{ij}}{\sqrt{(C_v)_{ii}} \sqrt{(C_v)_{jj}}} \right)_{1 \leq i, j \leq m}. \quad (31)$$

### Bemerkungen

- Bei u.i.v.  $v^{(1)}, \dots, v^{(n)}$  ist  $\bar{v}$  ein unverzerrter Schätzer von  $\mathbb{E}(v^{(1)})$ .
- Bei u.i.v.  $v^{(1)}, \dots, v^{(n)}$  ist  $C_v$  ein unverzerrter Schätzer von  $\mathbb{C}(v^{(1)})$ .

## Theorem (Datenmatrix und Deskriptivstatistiken)

$$Y := \begin{pmatrix} y^{(1)} & y^{(2)} & \dots & y^{(n)} \end{pmatrix} \in \mathbb{R}^{m \times n} \quad (32)$$

sei eine  $m \times n$  Datenmatrix, die durch die spaltenweise Konkatination von  $n$  Realsierungen  $m$ -dimensionaler Zufallsvektoren  $v^{(1)}, v^{(2)}, \dots, v^{(n)}$  gegeben sei. Dann ergeben sich

- für das Stichprobenmittel und

$$\bar{y} = \frac{1}{n} Y \mathbf{1}_n \quad (33)$$

- für die Stichprobenkovarianzmatrix

$$C_y = \frac{1}{n-1} \left( Y \left( I_n - \frac{1}{n} \mathbf{1}_n \mathbf{1}_n^T \right) Y^T \right). \quad (34)$$

- Schließlich ergibt sich mit

$$D := \text{diag} \left( \sqrt{C_{y_{ii}}}^{-1}, i = 1, \dots, m \right) \quad (35)$$

für die Stichprobenkorrelationsmatrix

$$R_y = D C D \quad (36)$$

### Bemerkungen

- Das Theorem erlaubt eine programmatisch effiziente Berechnung von  $\bar{y}$ ,  $C_y$  und  $R_y$ .

## Datenmatrix und Deskriptivstatistiken

```
# Laden einer m x n Datenmatrix
fname = file.path(getwd(), "5_Multivariate_Wahrscheinlichkeitstheorie.csv")
Y      = as.matrix(read.table(fname, sep = ",", header = TRUE))

# Deskriptivstatistiken
n      = ncol(Y)                                # Anzahl Datenvektorealisierungen
I_n    = diag(n)                               # Einheitsmatrix I_n
J_n    = matrix(rep(1,n^2), nrow = n)          # 1_{nn}
y_bar  = (1/n)* Y %*% J_n[,1]                  # Stichprobenmittel
C      = (1/(n-1))*(Y %*% (I_n-(1/n)*J_n) %*% t(Y)) # Stichprobenkovarianzmatrix
D      = diag(1/sqrt(diag(C)))                 # Kov-Korr-Transformationsmatrix
R      = D %*% C %*% D                         # Stichprobenkorrelationsmatrix
```

## Beweis

Die Darstellung des Stichprobenmittels ergibt sich nach Definition und mit  $v^{(i)} = y^{(i)}$  aus

$$\begin{aligned}\bar{y} &:= \frac{1}{n} \sum_{i=1}^n y^{(i)} \\ &= \frac{1}{n} \begin{pmatrix} \sum_{i=1}^n y_1^{(i)} \\ \vdots \\ \sum_{i=1}^n y_m^{(i)} \end{pmatrix} \\ &= \frac{1}{n} \left( \begin{pmatrix} y_1^{(1)} & \cdots & y_1^{(n)} \\ \vdots & \ddots & \vdots \\ y_m^{(1)} & \cdots & y_m^{(n)} \end{pmatrix} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \right) \\ &= \frac{1}{n} Y 1_n\end{aligned}\tag{37}$$

# Erwartungswert, Kovarianzmatrizen und Korrelationsmatrizen

## Beweis

Hinsichtlich der Darstellung der Stichprobenkovarianzmatrix halten wir zunächst fest, dass nach Definition und mit  $v^{(i)} = y^{(i)}$  gilt, dass

$$\begin{aligned}C_y &:= \frac{1}{n-1} \sum_{i=1}^n (y^{(i)} - \bar{y})(y^{(i)} - \bar{y})^T \\&= \frac{1}{n-1} \sum_{i=1}^n \left( y^{(i)} y^{(i)T} - y^{(i)} \bar{y}^T - \bar{y} y^{(i)T} + \bar{y} \bar{y}^T \right) \\&= \frac{1}{n-1} \left( \sum_{i=1}^n y^{(i)} y^{(i)T} - \sum_{i=1}^n y^{(i)} \bar{y}^T - \sum_{i=1}^n \bar{y} y^{(i)T} + \sum_{i=1}^n \bar{y} \bar{y}^T \right) \\&= \frac{1}{n-1} \left( \sum_{i=1}^n y^{(i)} y^{(i)T} - n \bar{y} \bar{y}^T - n \bar{y} \bar{y}^T + n \bar{y} \bar{y}^T \right) \\&= \frac{1}{n-1} \left( \sum_{i=1}^n y^{(i)} y^{(i)T} - n \bar{y} \bar{y}^T \right)\end{aligned} \tag{38}$$

# Erwartungswert, Kovarianzmatrizen und Korrelationsmatrizen

## Beweis

Mit  $\mathbf{1}_n \mathbf{1}_n^T = \mathbf{1}_{nn}$  ergibt sich dann weiterhin

$$\begin{aligned} Y \left( I_n - \frac{1}{n} \mathbf{1}_{nn} \right) Y^T &= \left( Y I_n - \frac{1}{n} Y \mathbf{1}_{nn} \right) Y^T \\ &= Y Y^T - \frac{1}{n} Y \mathbf{1}_{nn} Y^T \\ &= \begin{pmatrix} y^{(1)} & \dots & y^{(n)} \end{pmatrix} \begin{pmatrix} y^{(1)T} \\ \vdots \\ y^{(n)T} \end{pmatrix} - \frac{1}{n} Y \mathbf{1}_n \mathbf{1}_n^T Y^T \\ &= \sum_{i=1}^n y^{(i)} y^{(i)T} - n \left( \frac{1}{n} Y \mathbf{1}_n \right) \left( \frac{1}{n} \mathbf{1}_n^T Y^T \right) \\ &= \sum_{i=1}^n y^{(i)} y^{(i)T} - n \bar{y} \bar{y}^T \\ &= \frac{1}{n-1} \sum_{i=1}^n (y^{(i)} - \bar{y})(y^{(i)} - \bar{y})^T \\ &= C_y \end{aligned} \tag{39}$$

## Beweis

Hinsichtlich der Korrelationsmatrix ergibt sich nach Definition und mit  $v^{(i)} = y^{(i)}$  und für ein beliebiges Indexpaar  $i, j$  mit  $1 \leq i, j \leq m$ , dass

$$\begin{aligned} R_{y_{ij}} &= \frac{(C_y)_{ij}}{\sqrt{(C_y)_{ii}} \sqrt{(C_y)_{jj}}} \\ &= \frac{1}{\sqrt{(C_y)_{ii}}} (C_y)_{ij} \frac{1}{\sqrt{(C_y)_{jj}}} \\ &= (DCD)_{ij} \end{aligned} \tag{40}$$

---

Vorbemerkungen

Multivariate Verteilungen

Marginalverteilungen und Bedingte Verteilungen

Erwartungswert, Kovarianzmatrizen und Korrelationsmatrizen

**Selbstkontrollfragen**

# Selbstkontrollfragen

---

1. Definieren Sie den Begriff des Zufallsvektors.
2. Definieren Sie den Begriff der multivariaten Verteilung eines Zufallsvektors.
3. Definieren Sie den Begriff der multivariaten Wahrscheinlichkeitsmassefunktion.
4. Definieren Sie den Begriff der multivariaten Wahrscheinlichkeitsdichtefunktion.
5. Definieren Sie den Begriff der univariaten Marginalverteilung eines Zufallsvektors.
6. Wie berechnet man die WMF der  $i$ ten Komponente eines diskreten Zufallsvektors?
7. Wie berechnet man die WDF der  $i$ ten Komponente eines kontinuierlichen Zufallsvektors?
8. Definieren Sie die Begriffe der bedingten WMF und der diskreten bedingten Verteilung.
9. Definieren Sie die Begriffe der bedingten WDF und der kontinuierlichen bedingten Verteilung.
10. Geben Sie die Definition des Erwartungswerts eines Zufallsvektors wieder.
11. Geben Sie die Definition der Kovarianzmatrix eines Zufallsvektors wieder.
12. Geben Sie die Definition der Korrelationsmatrix eines Zufallsvektors wieder.
13. Geben Sie die Definition von Stichprobenmittel, -kovarianzmatrix und -korrelationsmatrix wieder.
14. Geben Sie das Theorem zur Datenmatrix und Deskriptivstatistiken wieder.