Multivariate Verfahren

MSc Psychologie | MSc Klinische Psychologie und Psychotherapie

WiSe 2023/24

Prof. Dr. Dirk Ostwald

(3) Zufallsvektoren

Definition und multivariate Verteilungen

Marginale und bedingte Verteilungen

Normalverteilungen

Selbstkontrollfragen

Definition und multivariate Verteilungen

Marginale und bedingte Verteilungen

Normalverteilungen

Selbstkontrollfragen

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

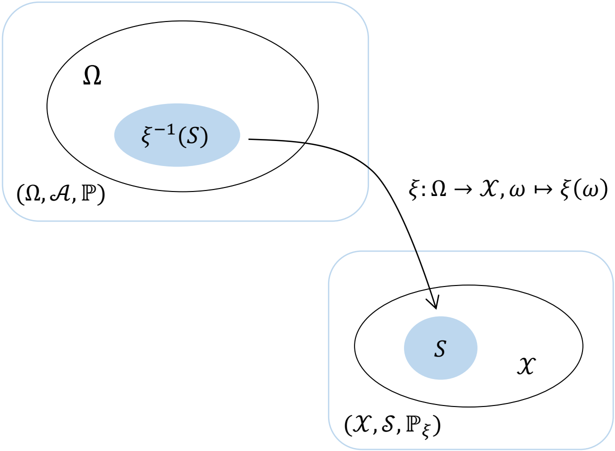
$$\xi : \Omega \to \mathcal{X}, \omega \mapsto \xi(\omega) := \begin{pmatrix} \xi_1(\omega) \\ \vdots \\ \xi_m(\omega) \end{pmatrix} \tag{1}$$

mit der *Messbarkeitseigenschaft*

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

Bemerkungen

- ξ 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 σ-Algebren wie $\mathcal{B}(\mathbb{R}^m)$.
- Ohne Beweis halten wir fest, dass ξ messbar ist, wenn die Funktionen $\xi_1, ..., \xi_m$ 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.



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

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

Definition (Multivariate Verteilung)

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

$$\xi : \Omega \to \mathcal{X}, \omega \mapsto \xi(\omega) \tag{3}$$

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

$$\mathbb{P}_\xi : \mathcal{S} \to [0,1], S \mapsto \mathbb{P}_\xi(S) := \mathbb{P}(\xi^{-1}(S)) = \mathbb{P}\left(\{\omega \in \Omega | \xi(\omega) \in S\}\right) \tag{4}$$

die *multivariate Verteilung des Zufallsvektor* ξ.

Bemerkungen

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

$$\begin{aligned}
\mathbb{P}_\xi(\xi \in S) &:= \mathbb{P}\left(\{\xi \in S\}\right) = \mathbb{P}\left(\{\omega \in \Omega | \xi(\omega) \in S\}\right) \\
\mathbb{P}_\xi(\xi = x) &:= \mathbb{P}\left(\{\xi = x\}\right) = \mathbb{P}\left(\{\omega \in \Omega | \xi(\omega) = x\}\right) \\
\mathbb{P}_\xi(\xi \le x) &:= \mathbb{P}\left(\{\xi \le x\}\right) = \mathbb{P}\left(\{\omega \in \Omega | \xi(\omega) \le x\}\right) \\
\mathbb{P}_\xi(x_1 \le \xi \le x_2) &:= \mathbb{P}\left(\{x_1 \le \xi \le x_2\}\right) = \mathbb{P}\left(\{\omega \in \Omega | x_1 \le \xi(\omega) \le x_2\}\right)
\end{aligned} \tag{5}$$

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

Definition (Diskreter Zufallsvektor, Multivariate WMF)

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

$$p_\xi : \mathcal{X} \to [0,1], x \mapsto p_\xi(x) \tag{6}$$

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 Wahrscheinlichkeitsmassefunktion (WMF)* von ξ.

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.

Dann entspricht der Ergebnisraum von ξ der in untenstehender Tabelle spezifizierten Menge an Tupeln (x_1, x_2)

(x_1, x_2)	$x_2 = 1$	$x_2 = 2$	$x_2 = 3$	$x_2 = 4$
$x_1 = 1$	$(1, 1)$	$(1, 2)$	$(1, 3)$	$(1, 4)$
$x_1 = 2$	$(2, 1)$	$(2, 2)$	$(2, 3)$	$(2, 4)$
$x_1 = 3$	$(3, 1)$	$(3, 2)$	$(3, 3)$	$(3, 4)$

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

$$p_\xi : \{1,2,3\} \times \{1,2,3,4\} \to [0,1], (x_1, x_2) \mapsto p_\xi(x_1, x_2) \quad (7)$$

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 $\sum_{x_1=1}^{3} \sum_{x_2=1}^{4} p_\xi(x_1, x_2) = 1$.

Definition (Kontinuierlicher Zufallsvektor, Multivariate WDF)

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

$$p_\xi : \mathbb{R}^m \to \mathbb{R}_{\geq 0}, x \mapsto p_\xi(x), \tag{8}$$

existiert, für die gilt

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

(2) $\mathbb{P}_\xi(x_1 \leq \xi \leq x_2) = \int_{x_{1_1}}^{x_{2_1}} \cdots \int_{x_{1_m}}^{x_{2_m}} p_\xi(s_1, ..., s_m)\,ds_1 \cdots ds_m$.

Eine entsprechende Funktion p heißt *multivariate Wahrscheinlichkeitsdichtefunktion (WDF)* von ξ.

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

$$\mathbb{P}_\xi(\xi = x) = \mathbb{P}_\xi(x \leq \xi \leq x) = \int_{x_1}^{x_1} \cdots \int_{x_m}^{x_m} p_\xi(s_1, ..., s_m)\,ds_1 \cdots ds_m = 0 \tag{9}$$

- Mit den multivariaten Normalverteilungen diskutieren wir unten ein ausführliches Beispiel. –>

Definition (Erwartungswert eines Zufallsvektors)

ξ sei ein m-dimensionaler Zufallvektor. Der *Erwartungwert von* ξ ist definiert als der m-dimensionale Vektor

$$\mathbb{E}(\xi) := \begin{pmatrix} \mathbb{E}(\xi_1) \\ \vdots \\ \mathbb{E}(\xi_m) \end{pmatrix}. \tag{10}$$

Bemerkung

- Der Erwartungswert eines Zufallsvektors ist der Vektor der Erwartungswerte seiner Komponenten.

Definition (Erwartungswert einer Zufallsmatrix)

Ξ sei eine durch die spaltenweise Konkatenation von $\xi^1, ..., \xi^n$ m-dimensionalen Zufallsvektoren gebildete Zufallsmatrix. Dann ist der *Erwartungwert von* Ξ ist definiert als die $m \times n$ Matrix

$$\mathbb{E}(\Xi) := \begin{pmatrix} \mathbb{E}\left(\xi_1^1\right) & \cdots & \mathbb{E}\left(\xi_1^n\right) \\ \vdots & \ddots & \vdots \\ \mathbb{E}\left(\xi_m^1\right) & \cdots & \mathbb{E}\left(\xi_m^n\right) \end{pmatrix}. \tag{11}$$

Bemerkung

- Der Erwartungswert einer Zufallsmatrix ist von Ξ ist die Matrix der spaltenweise konkatenierten Erwartungswerte $\mathbb{E}\left(\xi^1\right), ..., \mathbb{E}\left(\xi^n\right)$.

Theorem (Eigenschaften von Erwartungswerten)

(1) (Linear-affine Transformation einer Zufallsmatrix) Ξ sei ein $m \times n$-dimensionale Zufallsmatrix und es seien $A \in \mathbb{R}^{l\times m}, B \in \mathbb{R}^{n\times p}$ und $C \in \mathbb{R}^{l\times p}$. Dann gilt

$$\mathbb{E}(A\xi B + C) = A\mathbb{E}(\xi)B + C. \tag{12}$$

(2) (Linear-affine Transformation eines Zufallsvektors) ξ sei ein m-dimensionaler Zufallvektor und es seien $A \in \mathbb{R}^{n\times m}$ und $b \in \mathbb{R}^n$. Dann gilt

$$\mathbb{E}(A\xi + b) = A\mathbb{E}(\xi) + b. \tag{13}$$

(3) (Lineare Kombination zweier Zufallsvektoren) ξ und υ seien m-dimensionale Zufallvektoren und es seien $A, B \in \mathbb{R}^{n\times m}$. Dann gilt

$$\mathbb{E}\left(A\xi + B\upsilon\right) = A\mathbb{E}(\xi) + B\mathbb{E}(\upsilon). \tag{14}$$

Bemerkungen

- Die Aussagen sind im Wesentlichen analog zu den Eigenschaften des Erwartungswerts bei Zufallsvariablen.
- Eigenschaft (2) ist ein Spezialfall von Eigenschaft (1).

Beweis (fortgeführt)

Wir halten zunächst fest, dass für $i = 1, \dots, l$ und \$j = 1,…pm it den Regeln von Matrixmultiplikation und Matrixaddition und den Eigenschaften des Erwartungswertes von Zufallsvariablen gilt, dass

$$
\mathbb{E}(A\xi B + C)_{ij} = \mathbb{E}\left(\sum_{r=1}^{m}\sum_{s=1}^{n} a_{ir}\xi_r^s b_{sj} + c_{ij}\right) = \sum_{r=1}^{m}\sum_{s=1}^{n} a_{ir}\mathbb{E}(\xi_r^s) b_{sj} + c_{ij} \tag{15}
$$

. Dann aber gilt mit der Definition des Erwartungswerts einer Zufallsmatrix, dass

$$
\begin{aligned}
\mathbb{E}(A\xi B + C) &= \left(\mathbb{E}(A\xi B + C)_{ij}\right)_{1\le i\le l, 1\le j\le p} \\
&= \left(\sum_{r=1}^{m}\sum_{s=1}^{n} a_{ir}\mathbb{E}(\xi_r^s) b_{sj} + c_{ij}\right)_{1\le i\le l, 1\le j\le p} \\
&= \left(\sum_{r=1}^{m}\sum_{s=1}^{n} a_{ir}\mathbb{E}(\xi_r^s) b_{sj}\right)_{1\le i\le l, 1\le j\le p} + \left(c_{ij}\right)_{1\le i\le l, 1\le j\le p} \\
&= A\mathbb{E}(\xi)B + C.
\end{aligned} \tag{16}
$$

Beweis (fortgeführt)

(2) Für $i = 1,...,n$ gilt mit der Definition des Erwartungswert eines Zufallsvektors und den Eigenschaften des Erwartungswert einer Zufallsvariable, dass

$$\mathbb{E}(A\xi + b)_i = \mathbb{E}\left(\sum_{j=1}^m a_{ij}\xi_j + b_i\right) = \sum_{j=1}^m a_{ij}\mathbb{E}(\xi_j) + b_i. \tag{17}$$

Also gilt

$$\mathbb{E}(A\xi + b) = \begin{pmatrix} \sum_{j=1}^m a_{1j}\mathbb{E}(\xi_j) + b_1 \\ \vdots \\ \sum_{j=1}^m a_{nj}\mathbb{E}(\xi_j) + b_n \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^m a_{1j}\mathbb{E}(\xi_j) \\ \vdots \\ \sum_{j=1}^m a_{nj}\mathbb{E}(\xi_j) \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = A\mathbb{E}(\xi) + b. \tag{18}$$

(3) Für $i = 1,...,n$ gilt mit der Definition des Erwartungswert eines Zufallsvektors und den Eigenschaften des Erwartungswert einer Zufallsvariable, dass

$$\mathbb{E}\left(A\xi + B\upsilon\right)_i = \mathbb{E}\left(\sum_{j=1}^m a_{ij}\xi_j + \sum_{j=1}^m b_{ij}\upsilon_j\right) = \sum_{j=1}^m a_{ij}\mathbb{E}\left(\xi_j\right) + \sum_{j=1}^m b_{ij}\mathbb{E}\left(\upsilon_j\right) \tag{19}$$

Also gilt

$$\mathbb{E}\left(A\xi + B\upsilon\right) = \begin{pmatrix} \sum_{j=1}^m a_{1j}\mathbb{E}\left(\xi_j\right) + \sum_{j=1}^m b_{1j}\mathbb{E}\left(\upsilon_j\right) \\ \vdots \\ \sum_{j=1}^m a_{nj}\mathbb{E}\left(\xi_j\right) + \sum_{j=1}^m b_{nj}\mathbb{E}\left(\upsilon_j\right) \end{pmatrix} = A\mathbb{E}(\xi) + B\mathbb{E}(\upsilon). \tag{20}$$

Definition (Kovarianzmatrix eines Zufallsvektors)

ξ sei ein m-dimensionaler Zufallvektor. Dann ist die *Kovarianzmatrix von* ξ definiert als die $m \times m$ Matrix

$$\mathbb{C}(\xi) := \mathbb{E}\left((\xi - \mathbb{E}(\xi))(\xi - \mathbb{E}(\xi))^T\right). \tag{21}$$

Bemerkungen

- Die Kovarianzmatrix ist formal analog zur Kovarianz zweier Zufallsvariablen definiert.
- Der äußere Erwartungswert ist der Erwartungswert einer Matrix, die inneren Erwartungswerte von Vektoren.

Theorem (Eigenschaften der Kovarianzmatrix)

ξ sei ein m-dimensionaler Zufallsvektor und $\mathbb{C}(\xi)$ sei seine Kovarianzmatrix. Dann gelten

(1) (Elemente) Die Elemente von $\mathbb{C}(\xi)$ sind die Kovarianzen der Komponenten von ξ,

$$\mathbb{C}(\xi) = \left(\mathbb{C}(\xi_i, \xi_j)\right)_{1\le i,j\le m}. \tag{22}$$

(2) (Kovarianzmatrixverschiebungssatz) Es gilt

$$\mathbb{C}(\xi) = \mathbb{E}\left(\xi\xi^T\right) - \mathbb{E}(\xi)\mathbb{E}(\xi)^T. \tag{23}$$

(3) (Linear-affine Transformation) Für $A \in \mathbb{R}^{n\times m}$ und $b \in \mathbb{R}^n$ gilt

$$\mathbb{C}(A\xi + b) = A\mathbb{C}(\xi)A^T. \tag{24}$$

(4) (Matrixeigenschaften) $\mathbb{C}(\xi)$ ist symmetrisch und positiv-semidefinit.

Bemerkungen

- Die Diagonalelemente von $\mathbb{C}(\xi)$ sind die Varianzen der Komponenten von ξ, da

$$\mathbb{V}(\xi_i) = \mathbb{C}(\xi_i, \xi_i) \text{ für } i = 1, ...m. \tag{25}$$

- Eigenschaften (2) und (3) sind im Wesentlichen analog zu den Eigenschaften der Varianz.

Beweis

(1) Es gilt

$$
\begin{aligned}
\mathbb{C}(\xi) &:= \mathbb{E}\left((\xi - \mathbb{E}(\xi))(\xi - \mathbb{E}(\xi))^T\right) \\
&= \mathbb{E}\left(\left(\begin{pmatrix}\xi_1 \\ \vdots \\ \xi_n\end{pmatrix} - \begin{pmatrix}\mathbb{E}(\xi_1) \\ \vdots \\ \mathbb{E}(\xi_n)\end{pmatrix}\right)\left(\begin{pmatrix}\xi_1 \\ \vdots \\ \xi_n\end{pmatrix} - \begin{pmatrix}\mathbb{E}(\xi_1) \\ \vdots \\ \mathbb{E}(\xi_n)\end{pmatrix}\right)^T\right) \\
&= \mathbb{E}\left(\begin{pmatrix}\xi_1 - \mathbb{E}(\xi_1) \\ \vdots \\ \xi_n - \mathbb{E}(\xi_n)\end{pmatrix}\begin{pmatrix}\xi_1 - \mathbb{E}(\xi_1) \\ \vdots \\ \xi_n - \mathbb{E}(\xi_n)\end{pmatrix}^T\right) \\
&= \mathbb{E}\left(\begin{pmatrix}\xi_1 - \mathbb{E}(\xi_1) \\ \vdots \\ \xi_n - \mathbb{E}(\xi_n)\end{pmatrix}\begin{pmatrix}\xi_1 - \mathbb{E}(\xi_1) & \ldots & \xi_n - \mathbb{E}(\xi_n)\end{pmatrix}\right) \\
&= \mathbb{E}\begin{pmatrix}(\xi_1 - \mathbb{E}(\xi_1))(\xi_1 - \mathbb{E}(\xi_1)) & \ldots & (\xi_1 - \mathbb{E}(\xi_1))(\xi_n - \mathbb{E}(\xi_n) \\ \vdots & \ddots & \vdots \\ (\xi_n - \mathbb{E}(\xi_n))(\xi_1 - \mathbb{E}(\xi_1)) & \ldots & (\xi_n - \mathbb{E}(\xi_n))(\xi_n - \mathbb{E}(\xi_n))\end{pmatrix} \\
&= \left(\mathbb{E}\left((\xi_i - \mathbb{E}(\xi_i))(\xi_j - \mathbb{E}(\xi_j))\right)\right)_{1 \le i,j \le n} \\
&= \left(\mathbb{C}(\xi_i, \xi_j)\right)_{1 \le i,j \le n}.
\end{aligned}
\tag{26}
$$

Beweis (fortgeführt)

(2) Mit den Eigenschaften von Erwartungswerten gilt

$$
\begin{aligned}
\mathbb{C}(\xi) &= \mathbb{E}\left((\xi - \mathbb{E}(\xi))(\xi - \mathbb{E}(\xi))^T\right) \\
&= \mathbb{E}\left(\xi\xi^T - \xi\mathbb{E}(\xi)^T - \mathbb{E}(\xi)\xi^T + \mathbb{E}(\xi)\mathbb{E}(\xi)^T\right) \\
&= \mathbb{E}\left(\xi\xi^T\right) - \mathbb{E}(\xi)\mathbb{E}(\xi)^T - \mathbb{E}(\xi)\mathbb{E}(\xi)^T + \mathbb{E}(\xi)\mathbb{E}(\xi)^T \\
&= \mathbb{E}\left(\xi\xi^T\right) - \mathbb{E}(\xi)\mathbb{E}(\xi)^T.
\end{aligned}
\tag{27}
$$

(3) Mit den Eigenschaften von Erwartungswerten gilt

$$
\begin{aligned}
\mathbb{C}(A\xi + b) &= \mathbb{E}\left((A\xi + b - \mathbb{E}(A\xi + b))(A\xi + b - \mathbb{E}(A\xi + b))^T\right) \\
&= \mathbb{E}\left((A\xi + b - A\mathbb{E}(\xi) - b)(A\xi + b - A\mathbb{E}(\xi) - b)^T\right) \\
&= \mathbb{E}\left((A(\xi - \mathbb{E}(\xi)))(A(\xi - \mathbb{E}(\xi)))^T\right) \\
&= \mathbb{E}\left(A(\xi - \mathbb{E}(\xi))(\xi - \mathbb{E}(\xi))^T A^T\right) \\
&= A\mathbb{E}\left((\xi - \mathbb{E}(\xi))(\xi - \mathbb{E}(\xi))^T\right) A^T \\
&= A\mathbb{C}(\xi)A^T.
\end{aligned}
\tag{28}
$$

Beweis (fortgeführt)

(4) Die Symmetrie von $\mathbb{C}(\xi)$ folgt aus der Symmetrie der Kovarianz einer Zufallsvariable mit

$$\mathbb{C}(\xi_i, \xi_j) = \mathbb{C}(\xi_j, \xi_i) \text{ für alle } i = 1, ..., m, j = 1, ..., m. \tag{29}$$

Um die positive Semidefinitheit von $\mathbb{C}(\xi)$ nachzuweisen, ist zu zeigen, dass $a^T\mathbb{C}(\xi)a \geq 0$ für alle $a \in \mathbb{R}^m$ mit $a \neq 0_m$. Sei also $a \in \mathbb{R}^m$ mit $a \neq 0$. Dann gilt mit Aussage (3) für $A := a^T \in \mathbb{R}^{1\times m}$, dass

$$a^T\mathbb{C}(\xi)a = \mathbb{C}(a^T\xi). \tag{30}$$

Weiterhin gilt mit der Definition der Kovarianzmatrix aber, dass

$$\mathbb{C}(a^T\xi) = \mathbb{E}\left(\left(a^T\xi - \mathbb{E}(a^T\xi)\right)^2\right) = \mathbb{V}\left(a^T\xi\right). \tag{31}$$

Da mit den Eigenschaften der Varianz die Varianz der Zufallsvariable $a^T\xi$ aber immer nichtnegativ ist, folgt

$$a^T\mathbb{C}(\xi)a = \mathbb{V}(a^T\xi) \geq 0 \tag{32}$$

und damit die positive Semidefinitheit von $\mathbb{C}(\xi)$.

Definition (Korrelationsmatrix)

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

$$\mathbb{R}(\xi) := \left(\rho_{ij}\right)_{1 \le i,j \le m} = \left(\frac{\mathbb{C}(\xi_i, \xi_j)}{\sqrt{\mathbb{V}(\xi_i)}\sqrt{\mathbb{V}(\xi_j)}}\right)_{1 \le i,j \le m} \tag{33}$$

Bemerkungen

- Mit $\mathbb{V}(\xi_i) = \mathbb{C}(\xi_i, \xi_i), i = 1, ..., m$ ist die Korrelationsmatrix in der Kovarianzmatrix implizit.
- Es gelten $\rho_{ij} \in [-1, 1]$ für $1 \le i, j \le m$ und $\rho_{ii} = 1$ für $1 \le i \le m$.

Definition und multivariate Verteilungen

Marginale und bedingte Verteilungen

Normalverteilungen

Selbstkontrollfragen

Definition (Univariate Marginalverteilung)

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

$$\mathbb{P}_{\xi_i} : \mathcal{S}_i \to [0,1], S \mapsto \mathbb{P}_\xi \left(\mathcal{X}_1 \times \cdots \times \mathcal{X}_{i-1} \times S \times \mathcal{X}_{i+1} \times \cdots \times \mathcal{X}_m\right) \text{ für } S \in \mathcal{S}_i \tag{34}$$

definierte Verteilung die *$i$te univariate Marginalverteilung von ξ*.

Bemerkungen

- Univariate Marginalverteilungen sind die Verteilungen der Komponenten eines Zufallsvektors.
- Univariate Marginalverteilungen sind Verteilungen von Zufallsvariablen.
- Die Festlegung der multivariaten Verteilung von ξ legt auch die Verteilungen der ξ_i fest.

Theorem (Marginale Wahrscheinlichkeitsmasse- und dichtefunktionen)

(1) $\xi = (\xi_1, ..., \xi_m)^T$ sei ein m-dimensionaler diskreter Zufallsvektor mit Wahrscheinlichkeitsmassefunktion p_ξ und Komponentenergebnisräumen $\mathcal{X}_1, ..., \mathcal{X}_m$. Dann ergibt sich die Wahrscheinlichkeitsmassefunktion der $i$ten Komponente ξ_i von ξ als

$$p_{\xi_i} : \mathcal{X}_i \to [0,1], x_i \mapsto p_{\xi_i}(x_i) := \sum_{x_1} \cdots \sum_{x_{i-1}} \sum_{x_{i+1}} \cdots \sum_{x_m} p_\xi(x_1, ..., x_{i-1}, x_i, x_{i+1}, ..., x_m). \quad (35)$$

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

$$p_{\xi_i} : \mathbb{R} \to \mathbb{R}_{\geq 0}, x_i \mapsto p_{\xi_i}(x_i) :=$$

$$\int_{x_1} \cdots \int_{x_{i-1}} \int_{x_{i+1}} \cdots \int_{x_m} p_\xi(x_1, ..., x_{i-1}, x_i, x_{i+1}, ..., x_m)\, dx_1 ... dx_{i-1}\, dx_{i+1} ... dx_m. \quad (36)$$

Bemerkungen

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

Beispiel (Marginale Wahrscheinlichkeitsmassefunktionen)

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_{ξ_1} und p_{ξ_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 $\sum_{x_2=1}^{3} p_{\xi_2}(x_2) = 1$ gilt.

Beispiel (Marginale Wahrscheinlichkeitsmassefunktionen)

Ein Realisierungsbeispiel mithilfe relativer Häufigkeiten mag den Begriff der marginalen WMF intuitiv verdeutlichen. Nehmen wir an, wir hätten $n = 100$ (unabhängige) Realisierungen von ξ vorliegen.

Um die Wahrscheinlichkeiten $p_\xi(x_1, x_2)$ zu schätzen, würden wir die Anzahl der Realisierungen von (x_1, x_2) zählen und durch n teilen. Hätten wir beispielsweise 12 Realisierungen von $(3, 2)$ vorliegen, so würden wir $p_\xi(3, 2) \approx 12/100 = 0.12$ schätzen.

Die Frage nach der marginalen Wahrscheinlichkeit von $x_2 = 2$ entspräche dann der Frage, wie oft unter den Realisierungen zu finden sind, bei denen $x_2 = 2$ ist, irrespektive des Wertes von x_1. Dies wäre gerade die Anzahl der Realisierungen der Form $(1, 2), (2, 2)$ und $(3, 2)$. Gäbe es von diesen beispielsweise $0, 22$ und 12 respektive, so würde man die Wahrscheinlichkeit $p_{\xi_2}(2)$ natürlicherweise durch

$$\frac{0 + 22 + 12}{100} = \frac{0}{100} + \frac{22}{100} + \frac{12}{100} = 0.00 + 0.22 + 0.12 = 0.34 \tag{37}$$

schätzen. Anstelle der Wahrscheinlichkeiten $p_\xi(1, 2)$, $p_\xi(2, 2)$, $p_\xi(3, 2)$ addiert man hier also die entsprechenden relativen Häufigkeiten.

Marginale und bedingte Verteilungen

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

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}. \tag{38}$$

Analog wird für zwei Zufallsvariablen ξ_1, ξ_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 ξ_1 gegeben ξ_2 mithilfe der Ereignisse

$$A := \{\xi_1 \in S_1\} \text{ und } B := \{\xi_2 \in S_2\} \tag{39}$$

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

$$\mathbb{P}(\{\xi_1 \in S_1\}|\{\xi_2 \in S_2\}) = \frac{\mathbb{P}(\{\xi_1 \in S_1\} \cap \{\xi_2 \in S_2\})}{\mathbb{P}(\{\xi_2 \in S_2\})}. \tag{40}$$

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 Ergebnisraum $\mathcal{X} := \mathcal{X}_1 \times \mathcal{X}_2$, WMF $p_\xi = p_{\xi_1,\xi_2}$ und marginalen WMFen p_{ξ_1} und p_{ξ_2}. Die bedingte WMF von ξ_1 gegeben $\xi_2 = x_2$ ist dann für $p_{\xi_2}(x_2) > 0$ definiert als

$$p_{\xi_1|\xi_2=x_2} : \mathcal{X}_1 \to [0,1], x_1 \mapsto p_{\xi_1|\xi_2=x_2}(x_1|x_2) := \frac{p_{\xi_1,\xi_2}(x_1,x_2)}{p_{\xi_2}(x_2)} \tag{41}$$

Analog ist für $p_{\xi_1}(x_1) > 0$ die bedingte WMF von ξ_2 gegeben $\xi_1 = x_1$ definiert als

$$p_{\xi_2|\xi_1=x_1} : \mathcal{X}_2 \to [0,1], x_2 \mapsto p_{\xi_2|\xi_1=x_1}(x_2|x_1) := \frac{p_{\xi_1,\xi_2}(x_1,x_2)}{p_{\xi_1}(x_1)} \tag{42}$$

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 ξ_1 gegeben $\xi_2 = x_2$ und ξ_2 gegeben $\xi_1 = x_1$*, respektive.

Bemerkungen

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

$$p_{\xi_1|\xi_2}(x_1|x_2) = \frac{p_{\xi_1,\xi_2}(x_1,x_2)}{p_{\xi_2}(x_2)} = \frac{\mathbb{P}(\{\xi_1 = x_1\} \cap \{\xi_2 = x_2\})}{\mathbb{P}(\{\xi_2 = x_2\})}. \tag{43}$$

- Bedingte Verteilungen sind (lediglich) normalisierte gemeinsame Verteilungen.

Marginale und bedingte Verteilungen

Beispiel (Bedingte Wahrscheinlichkeitsmassefunktionen)

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 WM-Fen ergeben sich folgende bedingte WMFen für $p_{\xi_2|\xi_1=x_1}$

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

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 $\mathbb{R}^2$, WDF $p_\xi = p_{\xi_1,\xi_2}$ und marginalen WDFen p_{ξ_1} und p_{ξ_2}. Die bedingte WDF von ξ_1 gegeben $\xi_2 = x_2$ ist dann für $p_{\xi_2}(x_2) > 0$ definiert als

$$p_{\xi_1|\xi_2=x_2} : \mathbb{R} \to \mathbb{R}_{\geq 0}, x_1 \mapsto p_{\xi_1|\xi_2=x_2}(x_1|x_2) := \frac{p_{\xi_1,\xi_2}(x_1, x_2)}{p_{\xi_2}(x_2)} \tag{44}$$

Analog ist für $p_{\xi_1}(x_1) > 0$ die bedingte WMF von ξ_2 gegeben $\xi_1 = x_1$ definiert als

$$p_{\xi_2|\xi_1=x_1} : \mathbb{R} \to \mathbb{R}_{\geq 0}, x_2 \mapsto p_{\xi_2|\xi_1=x_1}(x_2|x_1) := \frac{p_{\xi_1,\xi_2}(x_1, x_2)}{p_{\xi_1}(x_1)} \tag{45}$$

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 ξ_1 gegeben $\xi_2 = x_2$ und ξ_2 gegeben $\xi_1 = x_1$*, respektive.

Bemerkung

- Im kontinuierlichen Fall gilt zwar $\mathbb{P}(\xi = x) = 0$, aber nicht notwendig auch $p_\xi(x) = 0$.

Definition und multivariate Verteilungen

Marginale und bedingte Verteilungen

Normalverteilungen

Selbstkontrollfragen

Definition (Normalverteilte Zufallsvariable)

ξ sei eine Zufallsvariable mit Ergebnisraum $\mathbb{R}$ und WDF

$$p : \mathbb{R} \to \mathbb{R}_{>0}, x \mapsto p(x) := \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right). \tag{46}$$

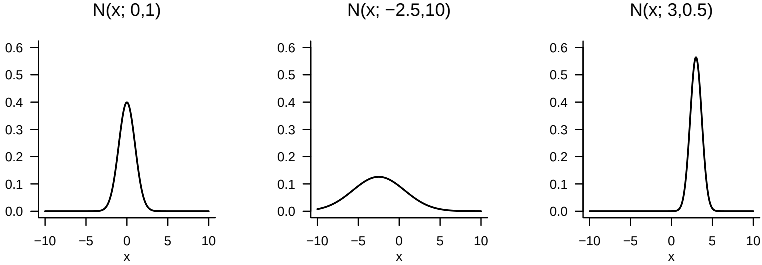
Dann sagen wir, dass ξ einer *Normalverteilung (oder Gauß-Verteilung) mit Erwartungswertparameter* $\mu \in \mathbb{R}$ *und Varianzparameter* $\sigma^2 > 0$ unterliegt und nennen ξ eine *normalverteilte Zufallsvariable*. Wir kürzen dies mit $\xi \sim N(\mu, \sigma^2)$ ab. Die WDF einer normalverteilten Zufallsvariable bezeichnen wir mit

$$N\left(x; \mu, \sigma^2\right) := \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right). \tag{47}$$

Bemerkungen

- Es gelten $\mathbb{E}(\xi) = \mu$ und $\mathbb{V}(\xi) = \sigma^2$.
- Der Parameter μ entspricht dem Wert höchster Wahrscheinlichkeitsdichte.
- Der Parameter σ^2 spezifiziert die Breite der WDF.
- $\xi \sim N(0, 1)$ heißt auch *standardnormalverteilt*.

Visualisierung univariater Normalverteilungsdichtefunktionen



Theorem (Konstruktion bivariater Normalverteilungen)

$\zeta_1 \sim N(0,1)$ und $\zeta_2 \sim N(0,1)$ seien zwei unabhängige standardnormalverteilte Zufallsvariablen. Weiterhin seien $\mu_1, \mu_2 \in \mathbb{R}$, $\sigma_1, \sigma_2 > 0$ und $\rho \in]-1,1[$. Schließlich seien

$$
\begin{aligned}
\xi_1 &:= \sigma_1 \zeta_1 + \mu_1 \\
\xi_2 &:= \sigma_2 \left(\rho \zeta_1 + (1-\rho^2)^{1/2} \zeta_2\right) + \mu_2 .
\end{aligned}
\tag{48}
$$

Dann hat die WDF des Zufallsvektors $\xi := (\xi_1, \xi_2)^T$, also der gemeinsamen Verteilung von ξ_1 und ξ_2, die Form

$$
p : \mathbb{R}^2 \to \mathbb{R}_{>0},\, x \mapsto p(x) := (2\pi)^{-1} |\Sigma|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)\right), \tag{49}
$$

wobei

$$
n := 2, \mu := \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \text{ und } \Sigma := \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_2\sigma_1 & \sigma_2^2 \end{pmatrix} \tag{50}
$$

Bemerkungen

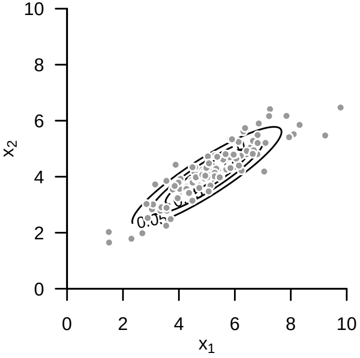
- Man nennt die gemeinsame Verteilung von ξ_1 und ξ_2 *bivariate Normalverteilung*.

Konstruktion bivariater Normalverteilungen

$$\mu_1 := 5.0, \mu_2 := 4.0, \sigma_1 := 1.5, \sigma_2 := 1.0, \rho := 0.9$$

• Realisierungen von $\xi = (\xi_1, \xi_2)^T$

– Isokonturen (Linien gleicher Wahrscheinlichkeitsdichte) von p



Definition (Multivariate Normalverteilung)

ξ sei ein m-dimensionaler Zufallsvektor mit Ergebnisraum $\mathbb{R}^m$ und WDF

$$p_\xi : \mathbb{R}^m \to \mathbb{R}_{>0},\, x \mapsto p_\xi(x) := (2\pi)^{-\frac{m}{2}} |\Sigma|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right). \tag{51}$$

Dann sagen wird, dass ξ einer *multivariaten (oder m-dimensionalen) Normalverteilung* mit *Erwartungswertparameter* $\mu \in \mathbb{R}^m$ und positive-definitem *Kovarianzmatrixparameter* $\Sigma \in \mathbb{R}^{m\times m}$ unterliegt und nennen ξ einen *(multivariat) normalverteilten Zufallsvektor*. Wir kürzen dies mit $\xi \sim N(\mu, \Sigma)$ ab. Die WDF eines multivariat normalverteilten Zufallsvektors bezeichnen wir mit

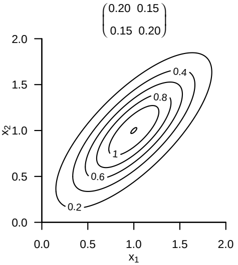
$$N(x; \mu, \Sigma) := (2\pi)^{-\frac{m}{2}} |\Sigma|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right). \tag{52}$$

Bemerkungen

- Der Parameter $\mu \in \mathbb{R}^m$ entspricht dem Wert höchster Wahrscheinlichkeitsdichte
- Die Diagonalelemente von Σ spezifizieren die Breite der WDF bezüglich $\xi_1, ..., \xi_m$.
- Das i, jte Element von Σ spezifiziert die Kovarianz on ξ_i und ξ_j.
- Der Term $(2\pi)^{-m/2}|\Sigma|^{-1/2}$ ist die Normalisierungskonstante für den Exponentialfunktionsterm.

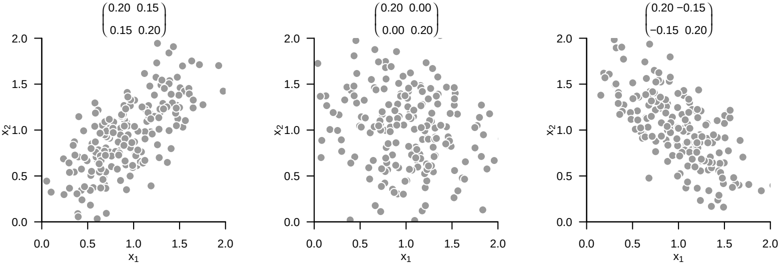
Visualisierung bivariater Normalverteilungsdichtefunktionen

$\mu = (1,1)^T, \Sigma \in \mathbb{R}^{2\times 2}$ wie im Abbildungstitel vermerkt.



Realisierung bivariater normalverteilter Zufallsvektoren

$\mu = (1,1)^T, \Sigma \in \mathbb{R}^{2\times 2}$ wie im Abbildungstitel vermerkt.



Theorem (Erwartungswert und Kovarianzmatrix)

$\xi \sim N(\mu, \Sigma)$ sei ein multivariate normalverteilter Zufallsvektor mit Erwartungswertparameter $\mu \in \mathbb{R}^m$ und Kovarianzmatrixparameter $\Sigma \in \mathbb{R}^{m \times m}$pd . Dann gelten für den Erwartungswert und die Kovarianzmatrix von ξ, dass

$$\mathbb{E}(\xi) = \mu \text{ und } \mathbb{C}(\xi) = \Sigma, \tag{53}$$

respektive.

Bemerkungen

- Wir verzichten auf einen Beweis.
- Das Theorem ist die direkte Generalisierung der Eigenschaften univariater normalverteilter Zufallsvariablen

Einige Eigenschaften von multivariaten Normalverteilungen

Wie bei univariaten Normalverteilungen gilt bei multivariaten Normalverteilungen, dass linear-affine Transformationen wiederrum auf Normalverteilugen führen, deren Parameter anhand der Ausgangsparameter und der Transformationsparameter errechnet werden können.

Multivariate Normalverteilungen haben weiterhin die Eigenschaft, dass auch alle anderen assoziierten Verteilung Normalverteilungen sind und deren Erwartungswert- und Kovarianzmatrixparameter aus den Parametern der jeweils komplementären Verteilung errechnet werden können.

Insbesondere gelten:

(1) Die uni- und multivariaten Marginalverteilungen multivariater Normalverteilungen sind Normalverteilungen.
(2) Wie alle multivariaten Verteilungen lassen sich multivariate Normalverteilungen multiplikativ in eine marginale und eine bedingte Verteilung zerlegen. Insbesondere sind bei multivariaten Normalverteilungen diese Verteilungen auch Normalverteilungen, deren Parameter aus den Parametern der gemeinsame Verteilung errechnet werden können und umgekehrt.

Wir fassen diese Einsichten in den folgenden Theoremen zusammen.

Theorem (Linear-affine Transformation)

$\xi \sim N(\mu, \Sigma)$ sei ein normalverteilter m-dimensionaler Zufallsvektor und es sei

$$\upsilon := A\xi + b \text{ mit } A \in \mathbb{R}^{n \times m} \text{ und } b \in \mathbb{R}^n. \tag{54}$$

Dann gilt

$$\upsilon \sim N\left(A\mu + b, A\Sigma A^T\right). \tag{55}$$

Bemerkung

- Die linear-affine Transformation eines multivariate normalverteilten Zuallsvektors ergibt wieder ein normalverteilten Zufallsvektor. Die Parameter des resultierenden normalverteilten Zufallsvektors ergeben sich dabei aus den Parametern des ursprünglichen Zufallsvektors und der Transformationsparamter.

Theorem (Marginale Normalverteilungen)

Es sei $m := k + l$ und $\xi = (\xi_1, ..., \xi_m)^T$ sei ein m-dimensionaler normalverteilter Zufallsvektor mit Erwartungswertparameter

$$\mu = \begin{pmatrix} \mu_\upsilon \\ \mu_\zeta \end{pmatrix} \in \mathbb{R}^m, \tag{56}$$

mit $\mu_\upsilon \in \mathbb{R}^k$ and $\mu_\zeta \in \mathbb{R}^l$ und Kovarianzmatrixparameter

$$\Sigma = \begin{pmatrix} \Sigma_{\upsilon\upsilon} & \Sigma_{\upsilon\zeta} \\ \Sigma_{\zeta\upsilon} & \Sigma_{\zeta\zeta} \end{pmatrix} \in \mathbb{R}^{m \times m}, \tag{57}$$

mit $\Sigma_{\upsilon\upsilon} \in \mathbb{R}^{k \times k}$, $\Sigma_{\upsilon\zeta} \in \mathbb{R}^{k \times l}$, $\Sigma_{\zeta\upsilon} \in \mathbb{R}^{l \times k}$, und $\Sigma_{\zeta\zeta} \in \mathbb{R}^{l \times l}$. Dann sind $\upsilon := (\xi_1, ..., \xi_k)^T$ und $\zeta := (\xi_{k+1}, ..., \xi_m)^T$ k- und l-dimensionale normalverteilte Zufallsvektoren, respektive, und es gilt

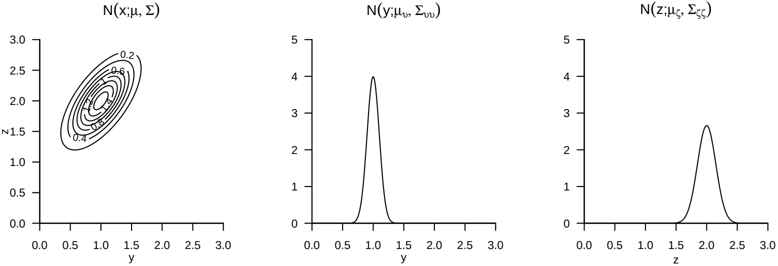
$$\upsilon \sim N(\mu_\upsilon, \Sigma_{\upsilon\upsilon}) \text{ and } \zeta \sim N(\mu_\zeta, \Sigma_{\zeta\zeta}), \tag{58}$$

Bemerkungen

- Die Marginalverteilungen einer multivariaten Normalverteilung sind auch Normalverteilungen.
- Die Parameter der Marginalverteilungen ergeben sich aus den Parametern der gemeinsamen Verteilung.

Marginale Normalverteilungen

$$m := 2, k = 1, l = 1, \mu := \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \Sigma := \begin{pmatrix} 0.10 & 0.08 \\ 0.08 & 0.15 \end{pmatrix}$$



Theorem (Gemeinsame Normalverteilungen)

ξ sei ein m-dimensionaler normalverteilter Zufallsvektor mit WDF

$$p_\xi : \mathbb{R}^m \to \mathbb{R}_{>0},\, x \mapsto p_\xi(x) := N(x; \mu_\xi, \Sigma_{\xi\xi}) \text{ mit } \mu_\xi \in \mathbb{R}^m, \Sigma_{\xi\xi} \in \mathbb{R}^{m \times m}, \tag{59}$$

$A \in \mathbb{R}^{n \times m}$ sei eine Matrix, $b \in \mathbb{R}^n$ sei ein Vektor und υ sei ein n-dimensionaler bedingt normalverteilter Zufallsvektor mit bedingter WDF

$$p_{\upsilon|\xi}(\cdot|x) : \mathbb{R}^n \to \mathbb{R}_{>0},\, y \mapsto p_{\upsilon|\xi}(y|x) := N(y; A\xi + b, \Sigma_{\upsilon\upsilon}) \text{ mit } \Sigma_{\upsilon\upsilon} \in \mathbb{R}^{n \times n}. \tag{60}$$

Dann ist der $m+n$-dimensionale Zufallsvektor $(\xi, \upsilon)^T$ normalverteilt mit (gemeinsamer) WDF

$$p_{\xi,\upsilon} : \mathbb{R}^{m+n} \to \mathbb{R}_{>0},\, \begin{pmatrix} x \\ y \end{pmatrix} \mapsto p_{\xi,\upsilon}\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = N\left(\begin{pmatrix} x \\ y \end{pmatrix}; \mu_{\xi,\upsilon}, \Sigma_{\xi,\upsilon}\right), \tag{61}$$

mit $\mu_{\xi,\upsilon} \in \mathbb{R}^{m+n}$ and $\Sigma_{\xi,\upsilon} \in \mathbb{R}^{m+n \times m+n}$ und insbesondere

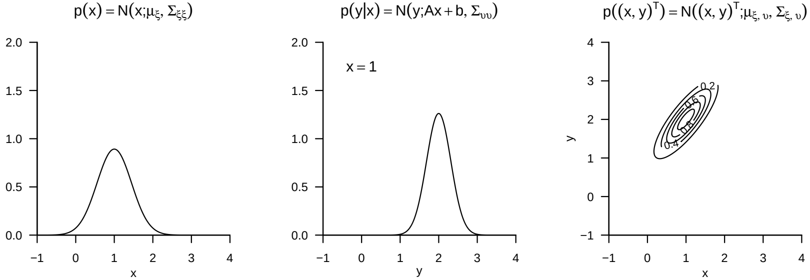
$$\mu_{\xi,\upsilon} = \begin{pmatrix} \mu_\xi \\ A\mu_\xi + b \end{pmatrix} \text{ und } \Sigma_{\xi,\upsilon} = \begin{pmatrix} \Sigma_{\xi\xi} & \Sigma_{\xi\xi} A^T \\ A\Sigma_{\xi\xi} & \Sigma_{\upsilon\upsilon} + A\Sigma_{\xi\xi} A^T \end{pmatrix}. \tag{62}$$

Bemerkungen

- Eine marginale und eine bedingte multivariate Normalverteilung induzieren eine gemeinsame Normalverteilung.
- Die Parameter der gemeinsamen Verteilungen ergeben sich als linear-affine Transformation der Parameter der induzierenden Verteilungen.

Gemeinsame Normalverteilungen

$$m := 1, n := 1, \mu_\xi := 1, \Sigma_{\xi\xi} := 0.2, A := 1, b := 1, \Sigma_{\upsilon\upsilon} := 0.1$$



Theorem (Bedingte Normalverteilungen)

(ξ, υ) sei ein $m+n$-dimensionaler normalverteilter Zufallsvektor mit WDF

$$p_{\xi,\upsilon} : \mathbb{R}^{m+n} \to \mathbb{R}_{>0}, \begin{pmatrix} x \\ y \end{pmatrix} \mapsto p_{\xi,\upsilon}\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) := N\left(\begin{pmatrix} x \\ y \end{pmatrix}; \mu_{\xi,\upsilon}, \Sigma_{\xi,\upsilon}\right), \tag{63}$$

mit

$$\mu_{\xi,\upsilon} = \begin{pmatrix} \mu_\xi \\ \mu_\upsilon \end{pmatrix}, \Sigma_{\xi,\upsilon} = \begin{pmatrix} \Sigma_{\xi\xi} & \Sigma_{\xi\upsilon} \\ \Sigma_{\upsilon\xi} & \Sigma_{\upsilon\upsilon} \end{pmatrix}, \tag{64}$$

mit $x, \mu_\xi \in \mathbb{R}^m, y, \mu_\upsilon \in \mathbb{R}^n$ and $\Sigma_{\xi\xi} \in \mathbb{R}^{m\times m}, \Sigma_{\xi\upsilon} \in \mathbb{R}^{m\times n}, \Sigma_{\upsilon\upsilon} \in \mathbb{R}^{n\times n}$. Dann ist die bedingte Verteilung von ξ gegeben υ eine m-dimensionale Normalverteilung mit bedingter WDF

$$p_{\xi|\upsilon}(\cdot|y) : \mathbb{R}^m \to \mathbb{R}_{>0}, x \mapsto p_{\xi|\upsilon}(x|y) := N(x; \mu_{\xi|\upsilon}, \Sigma_{\xi|\upsilon}) \tag{65}$$

mit

$$\mu_{\xi|\upsilon} = \mu_\xi + \Sigma_{\xi\upsilon}\Sigma_{\upsilon\upsilon}^{-1}(y - \mu_\upsilon) \in \mathbb{R}^m \tag{66}$$

und

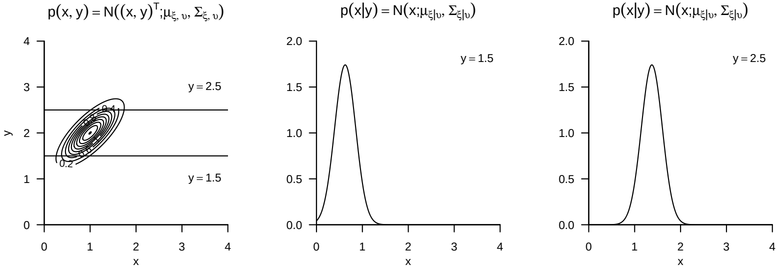
$$\Sigma_{\xi|\upsilon} = \Sigma_{\xi\xi} - \Sigma_{\xi\upsilon}\Sigma_{\upsilon\upsilon}^{-1}\Sigma_{\upsilon\xi} \in \mathbb{R}^{m\times m}. \tag{67}$$

Bemerkungen

- Die Parameter einer bedingten (multivariaten) Normalverteilung ergeben sich aus den Parametern einer gemeinsamen multivariaten Normalverteilung. Im Zusammenspiel mit den vorherigen Theoremen können die Parameter bedingter und marginale Normalverteilungen aus den Parametern der komplementären bedingten und marginalen Normalverteilungen bestimmt werden.

Bedingte Normalverteilungen

$$m := 2, n := 1, \mu := \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \Sigma := \begin{pmatrix} 0.12 & 0.09 \\ 0.09 & 0.12 \end{pmatrix}$$



Selbstkontrollfragen

1. Geben Sie Definition eines Zufallsvektors wieder.
2. Geben Sie Definition der multivariaten Verteilung eines Zufallsvektors wieder.
3. Geben Sie Definition einer multivariaten WMF wieder.
4. Geben Sie Definition einer multivariaten WDF wieder.
5. Geben Sie die Definition des Erwartungswerts eines Zufallsvektors wieder.
6. Geben Sie die Definition der Kovarianzmatrix eines Zufallsvektors wieder.
7. Was repräsentieren die Diagonalelemente der Kovarianzmatrix eines Zufallsvektors?
8. Was repräsentieren die Nichtdiagonalelemente der Kovarianzmatrix eines Zufallsvektors?
9. Geben Sie die Definition der Korrelationsmatrix eines Zufallsvektors wieder.
10. Geben Sie die Definition der univariaten Marginalverteilung eines Zufallsvektors wieder.
11. Wie berechnet man die WMF der $i$ten Komponente eines diskreten Zufallsvektors?
12. Wie berechnet man die WDF der $i$ten Komponente eines kontinuierlichen Zufallsvektors?
13. Geben Sie Definition der bedingten WMF und der diskreten bedingten Verteilung wieder.
14. Geben Sie Definition der bedingten WDF und der kontinuierlichen bedingten Verteilung wieder.
15. Geben Sie die Definition der WDF eines multivariaten normalverteilten Zufallsvektors wieder.
16. Erläutern Sie die Komponenten der WDF eines multivariaten normalverteilten Zufallsvektors.
17. Welche Werte haben der Erwartungswert und die Kovarianzmatrix eines normalverteilten Zufallsvektors?