



Inferenzstatistik

BSc Psychologie SoSe 2021

Prof. Dr. Dirk Ostwald

(3) Multivariate Verteilungen

Literaturhinweise

Die Lektüre von Georgii (2009, Abschnitte 3.1 - 3.4) wird empfohlen. Entsprechende englischsprachige Darstellungen finden sich in DeGroot and Schervish (2012, Sections 3.4 - 3.7) und Wasserman (2004, Sections 2.9 - 2.10).

Multivariate Verteilungen

- Zufallsvektoren
- Multivariate Verteilungen
- Marginalverteilungen
- Bedingte Verteilungen
- Unabhängige Zufallsvariablen
- Selbstkontrollfragen

Multivariate Verteilungen

- **Zufallsvektoren**
- Multivariate Verteilungen
- Marginalverteilungen
- Bedingte Verteilungen
- Unabhängige Zufallsvariablen
- Selbstkontrollfragen

Definition (Zufallsvektor)

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

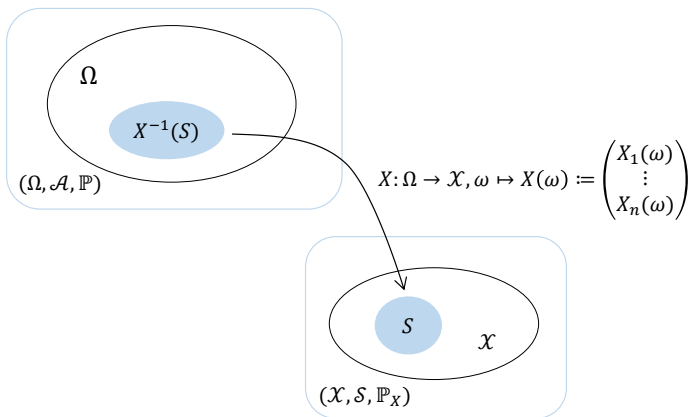
$$X : \Omega \rightarrow \mathcal{X}, \omega \mapsto X(\omega) := \begin{pmatrix} X_1(\omega) \\ \vdots \\ X_n(\omega) \end{pmatrix} \quad (1)$$

mit der *Messbarkeitseigenschaft*

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

Bemerkungen

- X ist messbar, wenn die Komponentenfunktionen $X_1, \dots, X_n$ messbar sind.
- Die Komponentenfunktionen eines Zufallsvektors sind also Zufallsvariablen.
- Ein n -dimensionaler Zufallsvektor ist die Konkatenation von n Zufallsvariablen.
- Für einen Zufallsvektor schreiben wir auch häufig $X := (X_1, \dots, X_n)$.
- Für $n := 1$ ist ein Zufallsvektor eine Zufallsvariable.



$$\mathbb{P}(X^{-1}(S)) = \mathbb{P}(\{\omega \in \Omega | X(\omega) \in S\}) =: \mathbb{P}_X(S)$$

Multivariate Verteilungen

- Zufallsvektoren
- **Multivariate Verteilungen**
- Marginalverteilungen
- Bedingte Verteilungen
- Unabhängige Zufallsvariablen
- Selbstkontrollfragen

Definition (Multivariate Verteilung)

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

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

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

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

die *multivariate Verteilung des Zufallsvektor* X .

Bemerkungen

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

$$\begin{aligned} \mathbb{P}_X(X \in S) &:= \mathbb{P}(\{X \in S\}) = \mathbb{P}(\{\omega \in \Omega | X(\omega) \in S\}) \\ \mathbb{P}_X(X = x) &:= \mathbb{P}(\{X = x\}) = \mathbb{P}(\{\omega \in \Omega | X(\omega) = x\}) \\ \mathbb{P}_X(X \leq x) &:= \mathbb{P}(\{X \leq x\}) = \mathbb{P}(\{\omega \in \Omega | X(\omega) \leq x\}) \end{aligned} \quad (5)$$

$$\mathbb{P}_X(x_1 \leq X \leq x_2) := \mathbb{P}(\{x_1 \leq X \leq x_2\}) = \mathbb{P}(\{\omega \in \Omega | x_1 \leq X(\omega) \leq x_2\})$$

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

Definition (Multivariate kumulative Verteilungsfunktionen)

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

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

multivariate kumulative Verteilungsfunktion von X .

Bemerkungen

- M.v. kumulative Verteilungsfunktionen können zur Definition m.v. Verteilungen genutzt werden.
- Häufiger ist die Definition m.v. Verteilungen durch multivariate WMFen oder WDFen.

Definition (Diskreter Zufallsvektor, multivariate WMF)

Es sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum und $X : \Omega \rightarrow \mathcal{X}$ ein Zufallsvektor. X heißt *diskreter Zufallsvektor*, wenn der Ergebnisraum $\mathcal{X}$ endlich viele oder höchstens abzählbar viele Elemente $x_i, i = 1, 2, \dots$ enthält. Die *multivariate Wahrscheinlichkeitsmassefunktion (WMF)* eines diskreten Zufallsvektors X wird mit p_X bezeichnet und ist definiert durch

$$p_X : \mathcal{X} \rightarrow [0, 1], x \mapsto p_X(x) := \mathbb{P}_X(X = x). \quad (7)$$

Bemerkungen

- Der Begriff der multivariaten WMF ist analog zum Begriff der WMF.
- Wie univariate WMFen sind multivariate WMFen nicht-negativ und normiert.

Beispiel (Multivariate Wahrscheinlichkeitsmassefunktion)

Wir betrachten einen zweidimensionalen Zufallsvektor $X := (X_1, X_2)$ 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

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

ist dann durch nachfolgende Tabelle definiert.

$p_X(x_1, x_2)$	$x_2 = 1$	$x_2 = 2$	$x_2 = 3$	$x_2 = 4$
$x_1 = 1$	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 $\sum_{x_1=1}^3 \sum_{x_2=1}^4 p_X(x_1, x_2) = 1$.

Definition (Kontinuierlicher Zufallsvektor, multivariate WDF)

Es sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum. Ein Zufallsvektor der Form $X : \Omega \rightarrow \mathbb{R}^n$ heißt *kontinuierlicher Zufallsvektor*. Die *multivariate Wahrscheinlichkeitsdichtefunktion (WDF)* eines kontinuierlichen Zufallsvektors X ist eine Funktion

$$p_X : \mathbb{R}^n \rightarrow \mathbb{R}_{\geq 0}, x \mapsto p_X(x), \quad (9)$$

mit den Eigenschaften

$$(1) \int_{\mathbb{R}^n} p_X(x) dx = 1$$

$$(2) \mathbb{P}_X(x_1 \leq X \leq x_2) = \int_{x_{11}}^{x_{21}} \cdots \int_{x_{1n}}^{x_{2n}} p_X(s_1, \dots, s_n) ds_1 \cdots ds_n$$

Bemerkungen

- Der Begriff der multivariaten WDF ist analog zum Begriff der WDF.
- Wie univariate WDFen sind multivariate WDFen nicht-negativ und normiert.
- Wie für kontinuierliche Zufallsvariablen gilt für kontinuierliche Zufallsvektoren

$$\mathbb{P}_X(X = x) = \mathbb{P}_X(x \leq X \leq x) = \int_{x_1}^{x_1} \cdots \int_{x_n}^{x_n} p_X(s_1, \dots, s_n) ds_1 \cdots ds_n \quad (10)$$

Multivariate Verteilungen

- Zufallsvektoren
- Multivariate Verteilungen
- **Marginalverteilungen**
- Bedingte Verteilungen
- Unabhängige Zufallsvariablen
- Selbstkontrollfragen

Definition (Univariate Marginalverteilung)

$(\Omega, \mathcal{A}, \mathbb{P})$ sei ein Wahrscheinlichkeitsraum, $(\mathcal{X}, \mathcal{S})$ sei ein n -dimensionaler Messraum, $X : \Omega \rightarrow \mathcal{X}$ sei ein Zufallsvektor, $\mathbb{P}_X$ sei die Verteilung von X , $\mathcal{X}_i \subset \mathcal{X}$ sei der Ergebnisraum der i ten Komponente X_i von X , und $\mathcal{S}_i$ sei eine σ -Algebra auf $\mathcal{X}_i$. Dann heit die durch

$$\mathbb{P}_{X_i} : \mathcal{S}_i \rightarrow [0, 1], S \mapsto \mathbb{P}_X (\mathcal{X}_1 \times \cdots \times \mathcal{X}_{i-1} \times S \times \mathcal{X}_{i+1} \times \cdots \times \mathcal{X}_n) \text{ fr } S \in \mathcal{S}_i \quad (11)$$

definierte Verteilung die *ite univariate Marginalverteilung von X* .

Bemerkungen

- Univariate Marginalverteilungen sind die Verteilungen der Komponenten eines Zufallsvektors.
- Die Komponenten eines Zufallsvektors sind Zufallsvariablen, univariate Marginalverteilungen sind also Verteilungen von Zufallsvariablen.
- Die Festlegung der multivariaten Verteilung von X legt auch die Verteilungen der X_i fest.

Theorem (Marginale Wahrscheinlichkeitsmasse/dichtefunktionen)

(1) $X = (X_1, \dots, X_n)$ sei ein n -dimensionaler diskreter Zufallsvektor mit Wahrscheinlichkeitsmassefunktion p_X und Komponentenergebnisräumen $\mathcal{X}_1, \dots, \mathcal{X}_n$. Dann ergibt sich die Wahrscheinlichkeitsmassefunktion der i ten Komponente X_i von X als

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

(2) $X = (X_1, \dots, X_n)$ sei ein n -dimensionaler kontinuierlicher Zufallsvektor mit Wahrscheinlichkeitsdichtefunktion p_X und Komponentenergebnisraum $\mathbb{R}$. Dann ergibt sich die Wahrscheinlichkeitsdichtefunktion der i ten Komponente X_i von X als

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

Bemerkungen

- Wir verzichten auf einen Beweis
- Die WMFen der univariaten Marginalverteilungen eines diskreten Zufallsvektors werden durch Summation bestimmt, die WDFen der univariaten Marginalverteilungen eines kontinuierlichen Zufallsvektors werden durch Integration bestimmt.

Beispiel (Marginale Wahrscheinlichkeitsmassefunktionen)

Wir betrachten erneut den zweidimensionalen Zufallsvektor $X := (X_1, X_2)$ 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 marginalen WMFen p_{X_1} und p_{X_2}

$p_X(x_1, x_2)$	$x_2 = 1$	$x_2 = 2$	$x_2 = 3$	$x_2 = 4$	$p_{X_1}(x_1)$
$x_1 = 1$	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
$p_{X_2}(x_2)$	0.2	0.3	0.3	0.2	

Man beachte, dass $\sum_{x_1=1}^3 p_{X_1}(x_1) = 1$ und $\sum_{x_2=1}^4 p_{X_2}(x_2) = 1$ gilt.



Multivariate Verteilungen

- Zufallsvektoren
- Multivariate Verteilungen
- Marginalverteilungen
- **Bedingte Verteilungen**
- Unabhängige Zufallsvariablen
- Selbstkontrollfragen

Vorbemerkungen

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 X_1, X_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 X_1 gegeben X_2 mithilfe der Ereignisse $A := \{X_1 \in S_1\}$ und $B := \{X_2 \in S_2\}$ definiert.

So ergibt sich zum Beispiel die bedingte Wahrscheinlichkeit, dass $X_1 \in S_1$ gegeben dass $X_2 \in S_2$ unter der Annahme, dass $\mathbb{P}(\{X_2 \in S_2\}) > 0$, zu

$$\mathbb{P}(\{X_1 \in S_1\}|\{X_2 \in S_2\}) = \frac{\mathbb{P}(\{X_1 \in S_1\} \cap \{X_2 \in S_2\})}{\mathbb{P}(\{X_2 \in S_2\})}. \quad (15)$$

In der Folge betrachten wir zunächst durch die WMFen/WDFen zweidimensionaler Zufallsvektoren definierte bedingte Verteilungen.

Definition (Bedingte WMF, diskrete bedingte Verteilung)

$X := (X_1, X_2)$ sei ein diskreter Zufallsvektor mit Ergebnisraum $\mathcal{X} := \mathcal{X}_1 \times \mathcal{X}_2$, WMF $p_X = p_{X_1, X_2}$ und marginalen WMFen p_{X_1} und p_{X_2} . Die bedingte WMF von X_1 gegeben $X_2 = x_2$ ist dann für $p_{X_2}(x_2) > 0$ definiert als

$$p_{X_1|X_2=x_2} : \mathcal{X}_1 \rightarrow [0, 1], x_1 \mapsto p_{X_1|X_2=x_2}(x_1|x_2) := \frac{p_{X_1, X_2}(x_1, x_2)}{p_{X_2}(x_2)} \quad (16)$$

Analog ist für $p_{X_1}(x_1) > 0$ die bedingte WMF von X_2 gegeben $X_1 = x_1$ definiert als

$$p_{X_2|X_1=x_1} : \mathcal{X}_2 \rightarrow [0, 1], x_2 \mapsto p_{X_2|X_1=x_1}(x_2|x_1) := \frac{p_{X_1, X_2}(x_1, x_2)}{p_{X_1}(x_1)} \quad (17)$$

Die bedingten Verteilungen mit WMFen $p_{X_1|X_2=x_2}$ und $p_{X_2|X_1=x_1}$ heißen dann die *diskreten bedingten Verteilungen* von X_1 gegeben $X_2 = x_2$ und X_2 gegeben $X_1 = x_1$, respektive.

Bemerkungen

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

$$p_{X_1|X_2}(x_1|x_2) = \frac{p_{X_1, X_2}(x_1, x_2)}{p_{X_2}(x_2)} = \frac{\mathbb{P}(\{X_1 = x_1\} \cap \{X_2 = x_2\})}{\mathbb{P}(\{X_2 = x_2\})}. \quad (18)$$

- **Bedingte Verteilungen sind lediglich normalisierte gemeinsame Verteilungen!**

Beispiel (Bedingte Wahrscheinlichkeitsmassenfunktionen)

Wir betrachten erneut den zweidimensionalen Zufallsvektor $X := (X_1, X_2)$ 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 $p_{X_2|X_1=x_1}$

$p_{X_1 X_2}(x_1, x_2)$	$x_2 = 1$	$x_2 = 2$	$x_2 = 3$	$x_2 = 4$
$p_{X_2 X_1=1}(x_2 x_1 = 1)$	$\frac{0.1}{0.4} = 0.25$	$\frac{0.0}{0.4} = 0.00$	$\frac{0.2}{0.4} = 0.50$	$\frac{0.1}{0.4} = 0.25$
$p_{X_2 X_1=2}(x_2 x_1 = 2)$	$\frac{0.1}{0.3} = 0.3\bar{3}$	$\frac{0.2}{0.3} = 0.6\bar{6}$	$\frac{0.0}{0.3} = 0.00$	$\frac{0.0}{0.3} = 0.00$
$p_{X_2 X_1=3}(x_2 x_1 = 3)$	$\frac{0.0}{0.3} = 0.00$	$\frac{0.1}{0.3} = 0.3\bar{3}$	$\frac{0.1}{0.3} = 0.3\bar{3}$	$\frac{0.1}{0.3} = 0.3\bar{3}$

Bemerkungen

- Man beachte, dass $\sum_{x_2=1}^4 p_{X_2|X_1=x_1}(x_2|x_1) = 1$ für alle $x_1 \in \mathcal{X}_1$.
- Man beachte die qualitative Ähnlichkeit der WMFen $p_{X_1, X_2}(x_1, x_2)$ und $p_{X_2|X_1}(x_2|x_1)$.
- **Bedingte Verteilungen sind lediglich normalisierte gemeinsame Verteilungen!**

Definition (Bedingte WDF, kontinuierliche bedingte Verteilungen)

$X := (X_1, X_2)$ sei ein kontinuierlicher Zufallsvektor mit Ergebnisraum $\mathbb{R}^2$, WDF $p_X = p_{X_1, X_2}$ und marginalen WDFen p_{X_1} und p_{X_2} . Die bedingte WDF von X_1 gegeben $X_2 = x_2$ ist dann für $p_{X_2}(x_2) > 0$ definiert als

$$p_{X_1|X_2=x_2} : \mathbb{R} \rightarrow \mathbb{R}_{\geq 0}, x_1 \mapsto p_{X_1|X_2=x_2}(x_1|x_2) := \frac{p_{X_1, X_2}(x_1, x_2)}{p_{X_2}(x_2)} \quad (19)$$

Analog ist für $p_{X_1}(x_1) > 0$ die bedingte WMF von X_2 gegeben $X_1 = x_1$ definiert als

$$p_{X_2|X_1=x_1} : \mathbb{R} \rightarrow \mathbb{R}_{\geq 0}, x_2 \mapsto p_{X_2|X_1=x_1}(x_2|x_1) := \frac{p_{X_1, X_2}(x_1, x_2)}{p_{X_1}(x_1)} \quad (20)$$

Die Verteilungen mit WDFen $p_{X_1|X_2=x_2}$ und $p_{X_2|X_1=x_1}$ heißen dann die *kontinuierlichen bedingten Verteilungen* von X_1 gegeben $X_2 = x_2$ und X_2 gegeben $X_1 = x_1$, respektive.

Bemerkungen

- Im kontinuierlichen Fall gilt zwar $\mathbb{P}(X = x) = 0$, aber nicht notwendig auch $p_X(x) = 0$!

Multivariate Verteilungen

- Zufallsvektoren
- Multivariate Verteilungen
- Marginalverteilungen
- Bedingte Verteilungen
- **Unabhängige Zufallsvariablen**
- Selbstkontrollfragen

Definition (Unabhängige Zufallsvariablen)

$(\Omega, \mathcal{A}, \mathbb{P})$ sei ein Wahrscheinlichkeitsraum und $X := (X_1, X_2)$ ein zweidimensionaler Zufallsvektor. Die Zufallsvariablen X_1, X_2 mit Ergebnisräumen $\mathcal{X}_1, \mathcal{X}_2$ heißen *unabhängig*, wenn für alle $S_1 \subseteq \mathcal{X}_1$ und $S_2 \subseteq \mathcal{X}_2$ gilt, dass

$$\mathbb{P}_X(X_1 \in S_1, X_2 \in S_2) = \mathbb{P}_{X_1}(X_1 \in S_1)\mathbb{P}_{X_2}(X_2 \in S_2). \quad (21)$$

Bemerkungen

- Die Definition besagt, dass die Ereignisse $\{X_1 \in S_1\}$ und $\{X_2 \in S_2\}$ unabhängig sind.
- Es gilt also auch, dass $\mathbb{P}(\{X_1 \in S_1\}|\{X_2 \in S_2\}) = \mathbb{P}(\{X_1 \in S_1\})$.
- Wissen um das Ereignis $\{X_2 \in S_2\}$ verändert die Wahrscheinlichkeit von $\{X_1 \in S_1\}$ also nicht.
- Einen formaleren Zugang bietet das Konzept der *Produktwahrscheinlichkeitsräume*.

Theorem (Unabhängigkeit und Faktorisierung der WMF/WDF)

(1) $X := (X_1, X_2)$ sei ein diskreter Zufallsvektor mit Ergebnisraum $\mathcal{X}_1 \times \mathcal{X}_2$, WMF p_X und marginalen WMFen p_{X_1}, p_{X_2} . Dann gilt

X_1 und X_2 sind unabhängige Zufallsvariablen $\Leftrightarrow$

$$p_X(x_1, x_2) = p_{X_1}(x_1)p_{X_2}(x_2) \text{ für alle } (x_1, x_2) \in \mathcal{X}_1 \times \mathcal{X}_2. \quad (22)$$

(2) $X := (X_1, X_2)$ sei ein kontinuierlicher Zufallsvektor mit Ergebnisraum $\mathbb{R}^2$, WDF p_X und marginalen WMFen p_{X_1}, p_{X_2} . Dann gilt

X_1 und X_2 sind unabhängige Zufallsvariablen $\Leftrightarrow$

$$p_X(x_1, x_2) = p_{X_1}(x_1)p_{X_2}(x_2) \text{ für alle } (x_1, x_2) \in \mathbb{R}^2. \quad (23)$$

Bemerkungen

- Wir verzichten auf einen Beweis.
- Die Produkteigenschaft $p_X(x_1, x_2) = p_{X_1}(x_1)p_{X_2}(x_2)$ heißt auch *Faktorisierung*.
- Unabhängigkeit zweier ZVen entspricht der Faktorisierung ihrer gemeinsamen WMF/WDF.

Beispiel (Unabhängige diskrete Zufallsvariablen)

Wir betrachten erneut den zweidimensionalen Zufallsvektor $X := (X_1, X_2)$, der Werte in $\{1, 2, 3\} \times \{1, 2, 3, 4\}$ annimmt, und dessenn gemeinsame und marginale WMFen die untenstehende Form haben

$p_X(x_1, x_2)$	$x_2 = 1$	$x_2 = 2$	$x_2 = 3$	$x_2 = 4$	$p_{X_1}(x_1)$
$x_1 = 1$	0.10	0.00	0.20	0.10	0.40
$x_1 = 2$	0.10	0.20	0.00	0.00	0.30
$x_1 = 3$	0.00	0.10	0.10	0.10	0.30
$p_{X_2}(x_2)$	0.20	0.30	0.30	0.20	

Da hier gilt, dass

$$p_X(1, 1) = 0.10 \neq 0.08 = 0.40 \cdot 0.20 = p_{X_1}(1)p_{X_2}(1) \quad (24)$$

sind die Zufallsvariablen X_1 und X_2 nicht unabhängig.

Beispiel (Unabhängige diskrete Zufallsvariablen)

Die gemeinsame Verteilung von X_1 und X_2 unter der Annahme der Unabhängigkeit von X_1 und X_2 bei gleichen Marginalverteilungen ergibt sich zu

$p_X(x_1, x_2)$	$x_2 = 1$	$x_2 = 2$	$x_2 = 3$	$x_2 = 4$	$p_{X_1}(x_1)$
$x_1 = 1$	0.08	0.12	0.12	0.08	0.40
$x_1 = 2$	0.06	0.09	0.09	0.06	0.30
$x_1 = 3$	0.06	0.09	0.09	0.06	0.30
$p_{X_2}(x_2)$	0.20	0.30	0.30	0.20	

(NEU) Weiterhin ergeben sich im Falle der Unabhängigkeit von X_1 und X_2 zum Beispiel die bedingten Wahrscheinlichkeitsmassenfunktion $p_{X_2|X_1}$ zu

$p_{X_1 X_2}(x_1, x_2)$	$x_2 = 1$	$x_2 = 2$	$x_2 = 3$	$x_2 = 4$
$p_{X_2 X_1=1}(x_2 x_1 = 1)$	$\frac{0.08}{0.40} = 0.2$	$\frac{0.12}{0.40} = 0.3$	$\frac{0.12}{0.40} = 0.3$	$\frac{0.08}{0.40} = 0.2$
$p_{X_2 X_1=2}(x_2 x_1 = 2)$	$\frac{0.06}{0.30} = 0.2$	$\frac{0.09}{0.30} = 0.3$	$\frac{0.09}{0.30} = 0.3$	$\frac{0.06}{0.30} = 0.2$
$p_{X_2 X_1=3}(x_2 x_1 = 3)$	$\frac{0.06}{0.30} = 0.2$	$\frac{0.09}{0.30} = 0.3$	$\frac{0.09}{0.30} = 0.3$	$\frac{0.06}{0.30} = 0.2$

Im Falle der Unabhängigkeit von X_1 und X_2 ändert sich die Verteilung von X_2 gegeben (oder im Wissen um) den Wert von X_1 also nicht und entspricht jeweils der Marginalverteilung von X_2 . Dies entspricht natürlich der Intuition der Unabhängigkeit von Ereignissen im Kontext elementarer Wahrscheinlichkeiten.

Definition (n unabhängige Zufallsvariablen)

$X := (X_1, \dots, X_n)$ sei ein n -dimensionaler Zufallsvektor mit Ergebnisraum $\mathcal{X} = \times_{i=1}^n \mathcal{X}_i$. Die n Zufallsvariablen $X_1, \dots, X_n$ heißen *unabhängig*, wenn für alle $S_i \in \mathcal{X}_i, i = 1, \dots, n$ gilt, dass

$$\mathbb{P}_X(X_1 \in S_1, \dots, X_n \in S_n) = \prod_{i=1}^n \mathbb{P}_{X_i}(X_i \in S_i). \quad (25)$$

Wenn der Zufallsvektor eine n -dimensionale WMF oder WDF p_X mit marginalen WMFen oder WDFen $p_{X_i}, i = 1, \dots, n$ besitzt, dann ist die Unabhängigkeit von $X_1, \dots, X_n$ gleichbedeutend mit der Faktorisierung der gemeinsamen WMF oder WDF, also mit

$$p_X(x_1, \dots, x_n) = \prod_{i=1}^n p_{X_i}(x_i). \quad (26)$$

Bemerkung

- Es handelt sich um eine direkte Generalisierung des zweidimensionalen Falls.

Definition (Unabhängig und identisch verteilte Zufallsvariablen)

n Zufallsvariablen $X_1, \dots, X_n$ heißen *unabhängig und identisch verteilt (u.i.v.)*, wenn

- (1) $X_1, \dots, X_n$ unabhängige Zufallsvariablen sind, und
- (2) die Marginalverteilungen der X_i übereinstimmen, also gilt, dass

$$p_{X_i} = p_{X_j} \text{ für alle } 1 \leq i, j \leq n. \quad (27)$$

Wenn die Zufallsvariablen $X_1, \dots, X_n$ unabhängig und identisch verteilt sind, und die i te Marginalverteilung $\mathbb{P}_X := \mathbb{P}_{X_i}$ ist, so schreibt man auch

$$X_1, \dots, X_n \sim \mathbb{P}_X. \quad (28)$$

Bemerkungen

- Man sagt kurz, dass $X_1, \dots, X_n$ u.i.v. sind.
- Im Englischen spricht man von *independent and identically distributed (i.i.d)* Zufallsvariablen.
- In der Statistik werden Stichproben meist mit u.i.v. Zufallsvariablen modelliert.
- n u.i.v. normalverteilte ZVen werden als $X_1, \dots, X_n \sim N(\mu, \sigma^2)$ geschrieben.

Multivariate Verteilungen

- Zufallsvektoren
- Multivariate Verteilungen
- Marginalverteilungen
- Bedingte Verteilungen
- Unabhängige Zufallsvariablen
- **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 WMF.
4. Definieren Sie den Begriff der multivariaten WDF.
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 den Begriff der Unabhängigkeit zweier Zufallsvariablen.
9. Wie erkennt man an der gemeinsamen WMF oder WDF eines zweidimensionalen Zufallsvektors, ob die Komponenten des Zufallsvektors unabhängig sind oder nicht?
10. Definieren Sie den Begriff der Unabhängigkeit von n Zufallsvariablen.
11. Definieren Sie den Begriff n unabhängig und identisch verteilter Zufallsvariablen.

Literatur

- DeGroot, M. H. and Schervish, M. J. (2012). *Probability and Statistics*. Addison-Wesley, Boston, 4th ed edition.
- Georgii, H.-O. (2009). *Stochastik: Einführung in die Wahrscheinlichkeitstheorie und Statistik*. DeGruyter-Lehrbuch. de Gruyter, Berlin, 4., überarb. und erw. aufl edition.
- Wasserman, L. (2004). *All of Statistics*.