



# Wahrscheinlichkeitstheorie und Frequentistische Inferenz

BSc Psychologie WiSe 2023/24

Prof. Dr. Dirk Ostwald

## (4) Zufallsvektoren

---

Definition

Multivariate Verteilungen

Marginalverteilungen

Bedingte Verteilungen

Unabhängige Zufallsvariablen

Selbstkontrollfragen

---

## Definition

Multivariate Verteilungen

Marginalverteilungen

Bedingte Verteilungen

Unabhängige Zufallsvariablen

Selbstkontrollfragen

## Definition (Zufallsvektor)

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

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

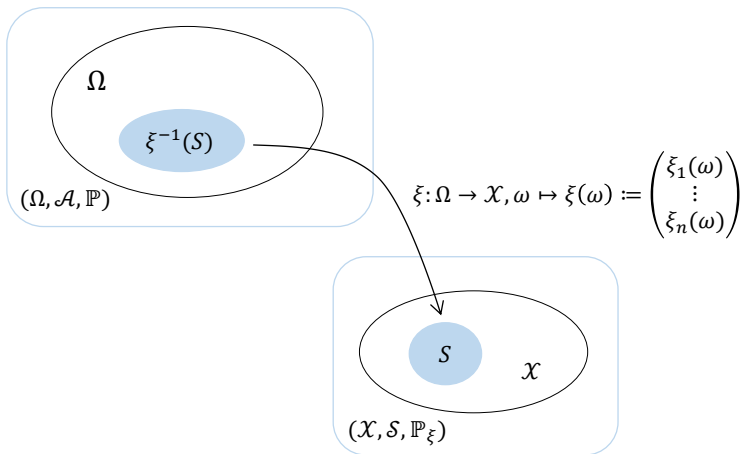
mit der *Messbarkeitseigenschaft*

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

### Bemerkungen

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

## Definition



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

---

Definition

## **Multivariate Verteilungen**

Marginalverteilungen

Bedingte Verteilungen

Unabhängige Zufallsvariablen

Selbstkontrollfragen

## Definition (Multivariate Verteilung)

$(\Omega, \mathcal{A}, \mathbb{P})$  sei ein Wahrscheinlichkeitsraum,  $(\mathcal{X}, \mathcal{S})$  sei ein  $n$ -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  $x, y \in \mathbb{R}^n$ , dass  $x_i \leq y_i$  für alle  $i = 1, \dots, n$ .



## 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

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

existiert, für die gilt

- (1)  $\sum_{x \in \mathcal{X}} p(x) = 1$  und
- (2)  $\mathbb{P}_\xi(\xi = x) = p(x)$  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. Dann entspricht der Ergebnisraum von  $\xi$  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

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

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

$$\sum_{x_1=1}^3 \sum_{x_2=1}^4 p_\xi(x_1, x_2) = 1. \quad (9)$$

## Definition (Kontinuierlicher Zufallsvektor, Multivariate WDF)

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

$$p_\xi : \mathbb{R}^n \rightarrow \mathbb{R}_{\geq 0}, x \mapsto p_\xi(x), \quad (10)$$

existiert, für die gilt

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

(2)  $\mathbb{P}_\xi(x_1 \leq \xi \leq x_2) = \int_{x_{1_1}}^{x_{2_1}} \dots \int_{x_{1_n}}^{x_{2_n}} p_\xi(s_1, \dots, s_n) ds_1 \dots ds_n.$

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

$$\mathbb{P}_\xi(\xi = x) = \mathbb{P}_\xi(x \leq \xi \leq x) = \int_{x_1}^{x_1} \dots \int_{x_n}^{x_n} p_\xi(s_1, \dots, s_n) ds_1 \dots ds_n = 0 \quad (11)$$

- Das Standardbeispiel ist die multivariate Normalverteilung (SoSe 2024).

---

Definition

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,  $\xi : \Omega \rightarrow \mathcal{X}$  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

$$\mathbb{P}_{\xi_i} : \mathcal{S}_i \rightarrow [0, 1], S \mapsto \mathbb{P}_\xi(\mathcal{X}_1 \times \cdots \times \mathcal{X}_{i-1} \times S \times \mathcal{X}_{i+1} \times \cdots \times \mathcal{X}_n) \text{ für } S \in \mathcal{S}_i \quad (12)$$

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

$$p_{\xi_i} : \mathcal{X}_i \rightarrow [0, 1], x_i \mapsto p_{\xi_i}(x_i) := \sum_{x_1} \cdots \sum_{x_{i-1}} \sum_{x_{i+1}} \cdots \sum_{x_n} p_\xi(x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_n). \quad (13)$$

(2)  $\xi = (\xi_1, \dots, \xi_n)$  sei ein  $n$ -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

$$p_{\xi_i} : \mathbb{R} \rightarrow \mathbb{R}_{\geq 0},$$
$$x_i \mapsto p_{\xi_i}(x_i) := \int_{x_1} \cdots \int_{x_{i-1}} \int_{x_{i+1}} \cdots \int_{x_n} p_\xi(x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_n) dx_1 \cdots dx_{i-1} dx_{i+1} \cdots dx_n. \quad (14)$$

### Bemerkungen

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



# Marginalverteilungen

## 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 notwendigerweise gilt, dass

$$1 = \sum_{x_1=1}^3 \sum_{x_2=1}^4 p_{\xi}(x_1, x_2) = \sum_{x_1=1}^3 p_{\xi_1}(x_1) \quad \text{und} \quad 1 = \sum_{x_2=1}^4 \sum_{x_1=1}^3 p_{\xi}(x_1, x_2) = \sum_{x_2=1}^4 p_{\xi_2}(x_2) \quad (15)$$

dass aus der Normiertheit von  $p_{\xi}$  die Normiertheit von  $p_{\xi_1}$  und  $p_{\xi_2}$  also direkt folgt.

## Beispiel (Marginale Wahrscheinlichkeitsmassfunktionen)

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

$$\frac{0 + 22 + 12}{100} = \frac{0}{100} + \frac{22}{100} + \frac{12}{100} = 0.00 + 0.22 + 0.12 = 0.34 \quad (16)$$

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

---

Definition

Multivariate Verteilungen

Marginalverteilungen

**Bedingte Verteilungen**

Unabhängige Zufallsvariablen

Selbstkontrollfragen

# Bedingte Verteilungen

## 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

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}. \quad (17)$$

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

$$A := \{\xi_1 \in S_1\} \text{ und } B := \{\xi_2 \in S_2\} \quad (18)$$

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

$$\mathbb{P}(\{\xi_1 \in S_1\}|\{\xi_2 \in S_2\}) = \frac{\mathbb{P}(\{\xi_1 \in S_1\} \cap \{\xi_2 \in S_2\})}{\mathbb{P}(\{\xi_2 \in S_2\})}. \quad (19)$$

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  $p_{\xi_1}$  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

$$p_{\xi_1|\xi_2=x_2} : \mathcal{X}_1 \rightarrow [0, 1], x_1 \mapsto p_{\xi_1|\xi_2=x_2}(x_1|x_2) := \frac{p_{\xi_1, \xi_2}(x_1, x_2)}{p_{\xi_2}(x_2)} \quad (20)$$

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

$$p_{\xi_2|\xi_1=x_1} : \mathcal{X}_2 \rightarrow [0, 1], x_2 \mapsto p_{\xi_2|\xi_1=x_1}(x_2|x_1) := \frac{p_{\xi_1, \xi_2}(x_1, x_2)}{p_{\xi_1}(x_1)} \quad (21)$$

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

$$p_{\xi_1|\xi_2}(x_1|x_2) = \frac{p_{\xi_1, \xi_2}(x_1, x_2)}{p_{\xi_2}(x_2)} = \frac{\mathbb{P}(\{\xi_1 = x_1\} \cap \{\xi_2 = x_2\})}{\mathbb{P}(\{\xi_2 = x_2\})}. \quad (22)$$

- Bedingte Verteilungen sind (lediglich) normalisierte gemeinsame Verteilungen.

## Beispiel (Bedingte 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 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  $\mathbb{R}^2$ , 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

$$p_{\xi_1|\xi_2=x_2} : \mathbb{R} \rightarrow \mathbb{R}_{\geq 0}, x_1 \mapsto p_{\xi_1|\xi_2=x_2}(x_1|x_2) := \frac{p_{\xi_1, \xi_2}(x_1, x_2)}{p_{\xi_2}(x_2)} \quad (23)$$

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

$$p_{\xi_2|\xi_1=x_1} : \mathbb{R} \rightarrow \mathbb{R}_{\geq 0}, x_2 \mapsto p_{\xi_2|\xi_1=x_1}(x_2|x_1) := \frac{p_{\xi_1, \xi_2}(x_1, x_2)}{p_{\xi_1}(x_1)} \quad (24)$$

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  $\mathbb{P}(\xi = x) = 0$ , aber nicht notwendig auch  $p_\xi(x) = 0$ .

---

Definition

Multivariate Verteilungen

Marginalverteilungen

Bedingte Verteilungen

**Unabhängige Zufallsvariablen**

Selbstkontrollfragen



## Definition (Unabhängige Zufallsvariablen)

$(\Omega, \mathcal{A}, \mathbb{P})$  sei ein Wahrscheinlichkeitsraum und  $\xi := (\xi_1, \xi_2)$  ein zweidimensionaler Zufallsvektor. Die Zufallsvariablen  $\xi_1, \xi_2$  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

$$\mathbb{P}_\xi(\xi_1 \in S_1, \xi_2 \in S_2) = \mathbb{P}_{\xi_1}(\xi_1 \in S_1)\mathbb{P}_{\xi_2}(\xi_2 \in S_2). \quad (25)$$

### Bemerkungen

- Die Definition besagt, dass die Ereignisse  $\{\xi_1 \in S_1\}$  und  $\{\xi_2 \in S_2\}$  unabhängig sind.
- Es gilt also auch, dass  $\mathbb{P}(\{\xi_1 \in S_1\}|\{\xi_2 \in S_2\}) = \mathbb{P}(\{\xi_1 \in S_1\})$ .
- Wissen um das Ereignis  $\{\xi_2 \in S_2\}$  verändert die Wahrscheinlichkeit von  $\{\xi_1 \in S_1\}$  nicht.
- Einen formaleren Zugang bietet das Konzept der *Produktwahrscheinlichkeitsräume*.

## Theorem (Unabhängigkeit und Faktorisierung der WMF/WDF)

(1)  $\xi := (\xi_1, \xi_2)$  sei ein diskreter Zufallsvektor mit Ergebnisraum  $\mathcal{X}_1 \times \mathcal{X}_2$ , WMF  $p_\xi$  und marginalen WMFen  $p_{\xi_1}, p_{\xi_2}$ . Dann gilt

$\xi_1$  und  $\xi_2$  sind unabhängige Zufallsvariablen  $\Leftrightarrow$

$$p_\xi(x_1, x_2) = p_{\xi_1}(x_1)p_{\xi_2}(x_2) \text{ für alle } (x_1, x_2) \in \mathcal{X}_1 \times \mathcal{X}_2. \quad (26)$$

(2)  $\xi := (\xi_1, \xi_2)$  sei ein kontinuierlicher Zufallsvektor mit Ergebnisraum  $\mathbb{R}^2$ , WDF  $p_\xi$  und marginalen WDFen  $p_{\xi_1}, p_{\xi_2}$ . Dann gilt

$\xi_1$  und  $\xi_2$  sind unabhängige Zufallsvariablen  $\Leftrightarrow$

$$p_\xi(x_1, x_2) = p_{\xi_1}(x_1)p_{\xi_2}(x_2) \text{ für alle } (x_1, x_2) \in \mathbb{R}^2. \quad (27)$$

### Bemerkungen

- Wir verzichten auf einen Beweis.
- Die Produkteigenschaft  $p_\xi(x_1, x_2) = p_{\xi_1}(x_1)p_{\xi_2}(x_2)$  heißt auch *Faktorisierung*.
- Unabhängigkeit zweier ZVen entspricht der Faktorisierung ihrer gemeinsamen WMF/WDF.

# Unabhängige Zufallsvariablen

## Beispiel (Unabhängige diskrete Zufallsvariablen)

Wir betrachten erneut den zweidimensionalen Zufallsvektor  $\xi := (\xi_1, \xi_2)$ , der Werte in  $\{1, 2, 3\} \times \{1, 2, 3, 4\}$  annimmt, und dessen gemeinsame und marginale WMFen die untenstehende Form haben

$p_\xi(x_1, x_2)$	$x_2 = 1$	$x_2 = 2$	$x_2 = 3$	$x_2 = 4$	$p_{\xi_1}(x_1)$
$x_1 = 1$	0.10	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
$p_{\xi_2}(x_2)$	0.20	0.30	0.30	0.20	

Da hier gilt, dass

$$p_\xi(1, 1) = 0.10 \neq 0.08 = 0.40 \cdot 0.20 = p_{\xi_1}(1)p_{\xi_2}(1) \quad (28)$$

sind die Zufallsvariablen  $\xi_1$  und  $\xi_2$  nicht unabhängig.

# Unabhängige Zufallsvariablen

## Beispiel (Unabhängige diskrete Zufallsvariablen)

Die gemeinsame Verteilung von  $\xi_1$  und  $\xi_2$  unter der Annahme der Unabhängigkeit von  $\xi_1$  und  $\xi_2$  bei gleichen Marginalverteilungen ergibt sich zu

$p_{\xi}(x_1, x_2)$	$x_2 = 1$	$x_2 = 2$	$x_2 = 3$	$x_2 = 4$	$p_{\xi_1}(x_1)$
$x_1 = 1$	0.08	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
$p_{\xi_2}(x_2)$	0.20	0.30	0.30	0.20	

Weiterhin ergeben sich im Falle der Unabhängigkeit von  $\xi_1$  und  $\xi_2$  zum Beispiel die bedingten Wahrscheinlichkeitsmassefunktion  $p_{\xi_2|\xi_1}$  zu

$p_{\xi_1 \xi_2}(x_1, x_2)$	$x_2 = 1$	$x_2 = 2$	$x_2 = 3$	$x_2 = 4$
$p_{\xi_2 \xi_1=1}(x_2 x_1 = 1)$	$\frac{0.08}{0.40} = 0.2$	$\frac{0.12}{0.40} = 0.3$	$\frac{0.12}{0.40} = 0.3$	$\frac{0.08}{0.40} = 0.2$
$p_{\xi_2 \xi_1=2}(x_2 x_1 = 2)$	$\frac{0.06}{0.30} = 0.2$	$\frac{0.09}{0.30} = 0.3$	$\frac{0.09}{0.30} = 0.3$	$\frac{0.06}{0.30} = 0.2$
$p_{\xi_2 \xi_1=3}(x_2 x_1 = 3)$	$\frac{0.06}{0.30} = 0.2$	$\frac{0.09}{0.30} = 0.3$	$\frac{0.09}{0.30} = 0.3$	$\frac{0.06}{0.30} = 0.2$

Im Falle der Unabhängigkeit von  $\xi_1$  und  $\xi_2$  ändert sich die Verteilung von  $\xi_2$  gegeben (oder im Wissen um) den Wert von  $\xi_1$  also nicht und entspricht jeweils der Marginalverteilung von  $\xi_2$ . Dies entspricht natürlich der Intuition der Unabhängigkeit von Ereignissen im Kontext elementarer Wahrscheinlichkeiten.

## Definition ( $n$ unabhängige Zufallsvariablen)

$\xi := (\xi_1, \dots, \xi_n)$  sei ein  $n$ -dimensionaler Zufallsvektor mit Ergebnisraum  $\mathcal{X} = \times_{i=1}^n \mathcal{X}_i$ . Die  $n$  Zufallsvariablen  $\xi_1, \dots, \xi_n$  heißen *unabhängig*, wenn für alle  $S_i \in \mathcal{X}_i, i = 1, \dots, n$  gilt, dass

$$\mathbb{P}_\xi(\xi_1 \in S_1, \dots, \xi_n \in S_n) = \prod_{i=1}^n \mathbb{P}_{\xi_i}(\xi_i \in S_i). \quad (29)$$

Wenn der Zufallsvektor eine  $n$ -dimensionale WMF oder WDF  $p_\xi$  mit marginalen WMFen oder WDFen  $p_{\xi_i}, i = 1, \dots, n$  besitzt, dann ist die Unabhängigkeit von  $\xi_1, \dots, \xi_n$  gleichbedeutend mit der Faktorisierung der gemeinsamen WMF oder WDF, also mit

$$p_\xi(\xi_1, \dots, \xi_n) = \prod_{i=1}^n p_{\xi_i}(x_i). \quad (30)$$

## Bemerkung

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

## Definition (Unabhängig und identisch verteilte Zufallsvariablen)

$n$  Zufallsvariablen  $\xi_1, \dots, \xi_n$  heißen *unabhängig und identisch verteilt (u.i.v.)*, wenn

- (1)  $\xi_1, \dots, \xi_n$  unabhängige Zufallsvariablen sind, und
- (2) die Marginalverteilungen der  $\xi_i$  übereinstimmen, also gilt, dass

$$\mathbb{P}_{\xi_i} = \mathbb{P}_{\xi_j} \text{ für alle } 1 \leq i, j \leq n. \quad (31)$$

Wenn die Zufallsvariablen  $\xi_1, \dots, \xi_n$  unabhängig und identisch verteilt sind und die *ite* Marginalverteilung  $\mathbb{P}_\xi := \mathbb{P}_{\xi_i}$  ist, so schreibt man auch

$$\xi_1, \dots, \xi_n \sim \mathbb{P}_\xi. \quad (32)$$

### Bemerkungen

- Man sagt kurz, dass  $\xi_1, \dots, \xi_n$  u.i.v. sind.
- Im Englischen spricht man von *independent and identically distributed (i.i.d)* Zufallsvariablen.
- In der Statistik werden Fehlerterme meist durch u.i.v. Zufallsvariablen modelliert.
- $n$  u.i.v. normalverteilte ZVen werden als  $\xi_1, \dots, \xi_n \sim N(\mu, \sigma^2)$  geschrieben.

---

Definition

Multivariate Verteilungen

Marginalverteilungen

Bedingte Verteilungen

Unabhängige Zufallsvariablen

**Selbstkontrollfragen**

# Selbstkontrollfragen

---

1. Geben Sie die Definition des Begriffs des Zufallsvektors wieder.
2. Geben Sie die Definition des Begriffs der multivariaten Verteilung eines Zufallsvektors wieder.
3. Geben Sie die Definition des Begriffs der multivariaten WMF wieder.
4. Geben Sie die Definition des Begriffs der multivariaten WDF wieder.
5. Geben Sie die Definition des Begriffs der univariaten Marginalverteilung eines Zufallsvektors wieder
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. Geben Sie die Definition des Begriffs der Unabhängigkeit zweier Zufallsvariablen wieder.
9. Wie erkennt man an der gemeinsamen WMF oder WDF eines zweidimensionalen Zufallsvektors, ob die Komponenten des Zufallsvektors unabhängig sind oder nicht?
10. Geben Sie die Definition des Begriffs der Unabhängigkeit von  $n$  Zufallsvariablen wieder.
11. Geben Sie die Definition des Begriffs der unabhängig und identisch verteilten Zufallsvariablen wieder.



# Selbstkontrollfragen - Lösungen

---

1. Siehe Definition (Zufallsvektor).
2. Siehe Definition (Multivariate Verteilung).
3. Siehe Definition (Diskreter Zufallsvektor, Multivariate WMF).
4. Siehe Definition (Kontinuierlicher Zufallsvektor, Multivariate WDF).
5. Siehe Definition (Univariate Marginalverteilung).
6. Siehe Theorem (Marginale Wahrscheinlichkeitsmasse- und -dichtefunktionen). Man erhält die WMF der  $i$ ten Komponente eines diskreten Zufallsvektors durch Summation der Werte der WMF des Zufallsvektors über alle Komponenten des Zufallsvektors *außer* der  $i$ ten Komponente.
7. Siehe Theorem (Marginal Wahrscheinlichkeitsmasse- und dichtefunktionen). Man erhält die WDF der  $i$ ten Komponente eines kontinuierlichen Zufallsvektors durch Integration der Werte der WDF des Zufallsvektors über alle Komponenten des Zufallsvektors *außer* der  $i$ ten Komponente.
8. Siehe Definition (Unabhängige Zufallsvariablen).
9. Man prüft, ob sich alle Werte der gemeinsamen WMF oder WDF  $p_{\xi}(x_1, x_2)$  des Zufallsvektors  $\xi(\xi_1, \xi_2)$  durch Multiplikation der entsprechenden marginalen WMF oder WDF Werte  $p_{\xi_1}(x_1)$  und  $p_{\xi_2}(x_2)$  ergeben. Ist dies für alle Werte  $(x_1, x_2) \in \mathcal{X}_1 \times \mathcal{X}_2$  der Fall, so sind  $\xi_1$  und  $\xi_2$  unabhängig, ist dies nicht für alle Werte  $(x_1, x_2) \in \mathcal{X}_1 \times \mathcal{X}_2$ , so sind  $\xi_1$  und  $\xi_2$  nicht unabhängig.
10. Siehe Definition ( $n$  unabhängige Zufallsvariablen).
11. Siehe Definition (Unabhängig und identisch verteilte Zufallsvariablen).