

## Notationen im Spezialfall $n = 2$

- Im Folgenden wird (auch für weitere Anwendungen) regelmäßig der Spezialfall  $n = 2$  betrachtet.
- Zur Vereinfachung der Darstellung (insbesondere zur Vermeidung doppelter Indizes) sei der betrachtete Zufallsvektor  $\mathbf{X}$  dann mit  $\mathbf{X} = (X, Y)'$  oder  $\mathbf{X} = (X, Y)$  statt  $\mathbf{X} = (X_1, X_2)'$  bezeichnet.
- Stetige 2-dimensionale Zufallsvektoren  $\mathbf{X} = (X, Y)$  werden in der Regel durch die Angabe einer gemeinsamen Dichtefunktion

$$f_{X,Y} : \mathbb{R}^2 \rightarrow \mathbb{R}; (x, y) \mapsto f_{X,Y}(x, y)$$

spezifiziert.

- Ist  $\mathbf{X} = (X, Y)$  ein zweidimensionaler diskreter Zufallsvektor mit „wenigen“ Trägerpunkten, stellt man die gemeinsame Verteilung — analog zu den Kontingenztafeln der deskriptiven Statistik — gerne in Tabellenform dar.

## Beispiel: (Gemeinsame) Wahrscheinlichkeitsfunktion

bei zweidimensionaler diskreter Zufallsvariable

- Ist  $\mathbf{X} = (X, Y)$  zweidimensionale diskrete Zufallsvariable mit endlichem Träger,  $A := T(X) = \{x_1, \dots, x_k\}$  der Träger von  $X$  und  $B := T(Y) = \{y_1, \dots, y_l\}$  der Träger von  $Y$ , so werden die Werte der Wahrscheinlichkeitsfunktion  $p_{\mathbf{X}}$  auch mit

$$p_{ij} := p_{(X,Y)}(x_i, y_j) \quad \text{für } i \in \{1, \dots, k\} \text{ und } j \in \{1, \dots, l\}$$

bezeichnet und wie folgt tabellarisch dargestellt:

$X \setminus Y$	$y_1$	$y_2$	$\cdots$	$y_l$
$x_1$	$p_{11}$	$p_{12}$	$\cdots$	$p_{1l}$
$x_2$	$p_{21}$	$p_{22}$	$\cdots$	$p_{2l}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
$x_k$	$p_{k1}$	$p_{k2}$	$\cdots$	$p_{kl}$

## Randverteilungen I

- Wie in der deskriptiven Statistik lassen sich die Verteilungen der einzelnen Zufallsvariablen eines  $n$ -dimensionalen Zufallsvektors auch aus der gemeinsamen Verteilung gewinnen.
- Analog zu den „Randhäufigkeiten“ erhält man so die **Randverteilungen** der einzelnen Komponenten des Zufallsvektors.
- Ist  $\mathbf{X}$  diskreter  $n$ -dimensionaler Zufallsvektor mit Wahrscheinlichkeitsfunktion  $p_{\mathbf{X}}$ , so erhält man für  $j \in \{1, \dots, n\}$  die Wahrscheinlichkeitsfunktion  $p_{X_j}$  zur  $j$ -ten Komponente  $X_j$  durch:

$$p_{X_j}(x) = \sum_{\substack{\mathbf{x}_i = (x_{i,1}, \dots, x_{i,n}) \\ x_{i,j} = x}} p_{\mathbf{X}}(\mathbf{x}_i)$$

## Randverteilungen II

- Ist  $\mathbf{X}$  stetiger  $n$ -dimensionaler Zufallsvektor mit gemeinsamer Dichtefunktion  $f_{\mathbf{X}}$ , so erhält man für  $j \in \{1, \dots, n\}$  eine Dichtefunktion  $f_{X_j} : \mathbb{R} \rightarrow \mathbb{R}$  zur  $j$ -ten Komponente  $X_j$  durch:

$$f_{X_j}(x) = \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{(n-1)\text{-mal}} f_{\mathbf{X}}(x_1, \dots, x_{j-1}, x, x_{j+1}, \dots, x_n) dx_n \cdots dx_{j+1} dx_{j-1} \cdots dx_1$$

- Für  $\mathbf{X} = (X, Y)$  erhält man also eine Randdichtefunktion  $f_X$  zu  $X$  durch

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dy$$

sowie eine Randdichtefunktion  $f_Y$  zu  $Y$  durch

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx .$$

## Fortsetzung Beispiel (zweidimensional, diskret)

Ergänzung der gemeinsamen Wahrscheinlichkeitstabelle um Randverteilungen

- Ist  $A = T(X) = \{x_1, \dots, x_k\}$  der Träger von  $X$  und  $B = T(Y) = \{y_1, \dots, y_l\}$  der Träger von  $Y$ , so erhält man für  $i \in \{1, \dots, k\}$  als Zeilensummen

$$p_{i\cdot} := p_X(x_i) = \sum_{j=1}^l p_{(X,Y)}(x_i, y_j) = \sum_{j=1}^l p_{ij}$$

sowie für  $j \in \{1, \dots, l\}$  als Spaltensummen

$$p_{\cdot j} := p_Y(y_j) = \sum_{i=1}^k p_{(X,Y)}(x_i, y_j) = \sum_{i=1}^k p_{ij}$$

und damit insgesamt die folgende ergänzte Tabelle:

$X \setminus Y$	$y_1$	$y_2$	$\dots$	$y_l$	$p_{i\cdot}$
$x_1$	$p_{11}$	$p_{12}$	$\dots$	$p_{1l}$	$p_{1\cdot}$
$x_2$	$p_{21}$	$p_{22}$	$\dots$	$p_{2l}$	$p_{2\cdot}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
$x_k$	$p_{k1}$	$p_{k2}$	$\dots$	$p_{kl}$	$p_{k\cdot}$
$p_{\cdot j}$	$p_{\cdot 1}$	$p_{\cdot 2}$	$\dots$	$p_{\cdot l}$	1

## (Stochastische) Unabhängigkeit von Zufallsvariablen I

- Die Komponenten  $X_1, \dots, X_n$  eines  $n$ -dimensionalen Zufallsvektors werden genau dann **stochastisch unabhängig** genannt, wenn alle Ereignisse der Form

$$\{X_1 \in B_1\}, \dots, \{X_n \in B_n\} \quad \text{für } B_1, \dots, B_n \in \mathcal{B}$$

stochastisch unabhängig sind:

### Definition 10.4

Seien  $n \in \mathbb{N}$  und  $X_1, \dots, X_n$  Zufallsvariablen über demselben Wahrscheinlichkeitsraum  $(\Omega, \mathcal{F}, P)$ . Dann heißen die Zufallsvariablen  $X_1, \dots, X_n$  **(stochastisch) unabhängig**, wenn für alle  $B_1, \dots, B_n \in \mathcal{B}$  gilt:

$$P\{X_1 \in B_1, \dots, X_n \in B_n\} = \prod_{i=1}^n P\{X_i \in B_i\} = P\{X_1 \in B_1\} \cdot \dots \cdot P\{X_n \in B_n\}$$

## (Stochastische) Unabhängigkeit von Zufallsvariablen II

- Man kann weiter zeigen, dass  $n$  diskrete Zufallsvariablen  $X_1, \dots, X_n$  über einem gemeinsamen Wahrscheinlichkeitsraum  $(\Omega, \mathcal{F}, P)$  genau dann stochastisch unabhängig sind, wenn für den (in diesem Fall ebenfalls diskreten) Zufallsvektor  $\mathbf{X} = (X_1, \dots, X_n)'$  bzw. die zugehörigen Wahrscheinlichkeitsfunktionen für alle  $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$  gilt:

$$p_{\mathbf{X}}(\mathbf{x}) = \prod_{i=1}^n p_{X_i}(x_i) = p_{X_1}(x_1) \cdot \dots \cdot p_{X_n}(x_n)$$

- Insbesondere sind  $X$  und  $Y$  im Fall  $n = 2$  mit  $\mathbf{X} = (X, Y)$  bei endlichen Trägern  $A = T(X) = \{x_1, \dots, x_k\}$  von  $X$  und  $B = T(Y) = \{y_1, \dots, y_l\}$  von  $Y$  unabhängig, falls  $p_{ij} = p_{i \cdot} \cdot p_{\cdot j}$  gilt für alle  $i \in \{1, \dots, k\}$  und  $j \in \{1, \dots, l\}$ .

## (Stochastische) Unabhängigkeit von Zufallsvariablen III

- Weiterhin sind  $n$  stetige Zufallsvariablen  $X_1, \dots, X_n$  über einem gemeinsamen Wahrscheinlichkeitsraum  $(\Omega, \mathcal{F}, P)$  stochastisch unabhängig, wenn der Zufallsvektor  $\mathbf{X} = (X_1, \dots, X_n)'$  stetig ist und eine gemeinsame Dichte  $f_{\mathbf{X}}$  bzw. Randdichten  $f_{X_1}, \dots, f_{X_n}$  existieren, so dass für alle  $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$  gilt:

$$f_{\mathbf{X}}(\mathbf{x}) = \prod_{i=1}^n f_{X_i}(x_i) = f_{X_1}(x_1) \cdot \dots \cdot f_{X_n}(x_n)$$

- Insbesondere sind stetige Zufallsvariablen  $X$  und  $Y$  im Fall  $n = 2$  mit  $\mathbf{X} = (X, Y)$  genau dann unabhängig, wenn es Dichtefunktionen  $f_X$  von  $X$ ,  $f_Y$  von  $Y$  sowie  $f_{X,Y}$  von  $(X, Y)$  gibt mit

$$f_{X,Y}(x, y) = f_X(x) \cdot f_Y(y) \quad \text{für alle } x, y \in \mathbb{R}.$$

## Bedingte Verteilungen 2-dim. Zufallsvektoren I

### Definition 10.5 (Bedingte Wahrscheinlichkeitsverteilung (diskret))

Seien  $\mathbf{X} = (X, Y)$  ein diskreter zweidimensionaler Zufallsvektor mit Wahrscheinlichkeitsfunktion  $p_{(X,Y)}$ ,  $p_Y$  die Wahrscheinlichkeitsfunktion zur Randverteilung von  $Y$ ,  $y \in T(Y)$ . Dann heißt die durch die Wahrscheinlichkeitsfunktion

$$p_{X|Y=y} : \mathbb{R} \rightarrow [0, 1]; p_{X|Y=y}(x) = \frac{p_{(X,Y)}(x, y)}{p_Y(y)}$$

definierte diskrete Wahrscheinlichkeitsverteilung **bedingte**

**Wahrscheinlichkeitsverteilung von  $X$  unter der Bedingung  $Y = y$ .**

Analog nennt man für  $x \in T(X)$  die durch die Wahrscheinlichkeitsfunktion

$$p_{Y|X=x} : \mathbb{R} \rightarrow [0, 1]; p_{Y|X=x}(y) = \frac{p_{(X,Y)}(x, y)}{p_X(x)}$$

definierte diskrete Wahrscheinlichkeitsverteilung **bedingte**

**Wahrscheinlichkeitsverteilung von  $Y$  unter der Bedingung  $X = x$ .**

## Bedingte Verteilungen 2-dim. Zufallsvektoren II

### Definition 10.6 (Bedingte Wahrscheinlichkeitsverteilung (stetig))

Seien  $\mathbf{X} = (X, Y)$  ein stetiger zweidimensionaler Zufallsvektor,  $f_{(X,Y)}$  eine gemeinsame Dichtefunktion von  $(X, Y)$ ,  $f_Y$  eine Dichtefunktion zur Randverteilung von  $Y$ ,  $y \in \mathbb{R}$  eine Stetigkeitsstelle von  $f_Y$  mit  $f_Y(y) > 0$ . Dann heißt die durch die Dichtefunktion

$$f_{X|Y=y} : \mathbb{R} \rightarrow \mathbb{R}; f_{X|Y=y}(x) = \frac{f_{(X,Y)}(x, y)}{f_Y(y)}$$

definierte stetige Wahrscheinlichkeitsverteilung **bedingte**

**Wahrscheinlichkeitsverteilung von  $X$  unter der Bedingung  $Y = y$ .**

Analog nennt man für Stetigkeitsstellen  $x \in \mathbb{R}$  von  $f_X$  mit  $f_X(x) > 0$  die durch die Dichtefunktion

$$f_{Y|X=x} : \mathbb{R} \rightarrow \mathbb{R}; f_{Y|X=x}(y) = \frac{f_{(X,Y)}(x, y)}{f_X(x)}$$

definierte stetige Wahrscheinlichkeitsverteilung **bedingte**

**Wahrscheinlichkeitsverteilung von  $Y$  unter der Bedingung  $X = x$ .**

## Bemerkungen I

- Mit den üblichen Bezeichnungen und Abkürzungen für zweidimensionale Zufallsvektoren mit endlichem Träger erhält man mit Definition 10.5 analog zu den bedingten Häufigkeiten bei zweidimensionalen Merkmalen

$$p_{X|Y=y_j}(x_i) = \frac{p_{ij}}{p_{\cdot j}} \quad \text{und} \quad p_{Y|X=x_i}(y_j) = \frac{p_{ij}}{p_{i\cdot}}$$

für  $i \in \{1, \dots, k\}$  und  $j \in \{1, \dots, l\}$ .

- Ebenfalls analog zur deskriptiven Statistik sind zwei Zufallsvariablen also genau dann stochastisch unabhängig, wenn die bedingten Verteilungen jeweils mit den zugehörigen Randverteilungen übereinstimmen, also wenn
  - ▶ im diskreten Fall  $p_{X|Y=y}(x) = p_X(x)$  für alle  $x \in \mathbb{R}$  und alle  $y \in T(Y)$  sowie  $p_{Y|X=x}(y) = p_Y(y)$  für alle  $y \in \mathbb{R}$  und alle  $x \in T(X)$  gilt.
  - ▶ im stetigen Fall (bedingte) Dichtefunktionen  $f_X, f_Y, f_{Y|X=x}, f_{X|Y=y}$  existieren mit  $f_{X|Y=y}(x) = f_X(x)$  für alle  $x \in \mathbb{R}$  und alle Stetigkeitsstellen  $y \in \mathbb{R}$  von  $f_Y$  mit  $f_Y(y) > 0$  sowie  $f_{Y|X=x}(y) = f_Y(y)$  für alle  $y \in \mathbb{R}$  und alle Stetigkeitsstellen  $x \in \mathbb{R}$  von  $f_X$  mit  $f_X(x) > 0$ .

## Bemerkungen II

- Durch bedingte Dichtefunktionen  $f_{X|Y=y}$  bzw.  $f_{Y|X=x}$  oder bedingte Wahrscheinlichkeitsfunktionen  $p_{X|Y=y}$  bzw.  $p_{Y|X=x}$  definierte bedingte Verteilungen können *im weiteren Sinne* als Verteilungen eindimensionaler Zufallsvariablen  $X|Y = y$  bzw.  $Y|X = x$  aufgefasst werden.
- Damit lassen sich viele für eindimensionale Zufallsvariablen bekannten Methoden/Konzepte in natürlicher Weise übertragen, insbesondere auf
  - ▶ bedingte Verteilungsfunktionen  $F_{X|Y=y}$  bzw.  $F_{Y|X=x}$ ,
  - ▶ bedingte Wahrscheinlichkeitsmaße  $P_{X|Y=y}$  bzw.  $P_{Y|X=x}$ ,
  - ▶ bedingte Träger  $T(X|Y = y)$  bzw.  $T(Y|X = x)$ ,
  - ▶ bedingte Erwartungswerte  $E(X|Y = y)$  bzw.  $E(Y|X = x)$ ,
  - ▶ bedingte Varianzen  $\text{Var}(X|Y = y)$  bzw.  $\text{Var}(Y|X = x)$ ,
  - ▶ bedingte höhere Momente  $m_k(X|Y = y)$  bzw.  $m_k(Y|X = x)$ ,
  - ▶ bedingte zentrale höhere Momente  $\mu_k(X|Y = y)$  bzw.  $\mu_k(Y|X = x)$ .
- Den bedingten Erwartungswert  $E(Y|X = x)$  erhält man zum Beispiel
  - ▶ im diskreten Fall durch:  $E(Y|X = x) = \sum_{y_i} y_i \cdot p_{Y|X=x}(y_i)$
  - ▶ im stetigen Fall durch:  $E(Y|X = x) = \int_{-\infty}^{+\infty} y \cdot f_{Y|X=x}(y) dy$
- Gültig bleibt auch der VZS:  $\text{Var}(X|Y = y) = E(X^2|Y = y) - [E(X|Y = y)]^2$