

# Inhaltsverzeichnis

(Ausschnitt)

## 10 Mehrdimensionale Zufallsvariablen

- Borelsche sigma-Algebra
- Diskrete Zufallsvektoren
- Randverteilungen
- (Stochastische) Unabhängigkeit
- Bedingte Verteilungen
- Momente zweidimensionaler Zufallsvektoren
- Momente höherdimensionaler Zufallsvektoren

# Mehrdimensionale Zufallsvariablen/Zufallsvektoren I

- Im Folgenden: *Simultane Betrachtung mehrerer* (endlich vieler) Zufallsvariablen über *demselben* Wahrscheinlichkeitsraum  $(\Omega, \mathcal{F}, P)$ .
- Ist  $n \in \mathbb{N}$  die Anzahl der betrachteten Zufallsvariablen, so fasst man die  $n$  Zufallsvariablen  $X_1, \dots, X_n$  in einem  $n$ -dimensionalen Vektor  $\mathbf{X} = (X_1, \dots, X_n)'$  zusammen.
- Damit ist  $\mathbb{R}^n$  der Wertebereich der Abbildung  $\mathbf{X} : \Omega \rightarrow \mathbb{R}^n$ , als  $\sigma$ -Algebra über  $\mathbb{R}^n$  wählt man die  $n$ -dimensionale Borelsche  $\sigma$ -Algebra  $\mathcal{B}^n$ , in der alle karthesischen Produkte von  $n$  Elementen aus  $\mathcal{B}$  enthalten sind.
- Insbesondere enthält  $\mathcal{B}^n$  alle endlichen und abzählbar unendlichen Teilmengen von  $\mathbb{R}^n$  sowie alle karthesischen Produkte von  $n$  Intervallen aus  $\mathbb{R}$ .
- Damit lassen sich die meisten bekannten Konzepte eindimensionaler Zufallsvariablen leicht übertragen.
- Ähnlich zur Situation bei mehrdimensionalen Merkmalen in der deskriptiven Statistik werden viele Darstellungen im Fall  $n > 2$  allerdings schwierig.

# Mehrdimensionale Zufallsvariablen/Zufallsvektoren II

## Definition 10.1 (Zufallsvektor, Mehrdimensionale Zufallsvariable)

Seien  $(\Omega, \mathcal{F}, P)$  ein Wahrscheinlichkeitsraum,  $n \in \mathbb{N}$ ,

$$\mathbf{X} : \Omega \rightarrow \mathbb{R}^n : \mathbf{X}(\omega) = (X_1(\omega), \dots, X_n(\omega))'$$

eine  $\mathcal{F} - \mathcal{B}^n$ -messbare Abbildung. Dann heißen  $\mathbf{X}$   **$n$ -dimensionale Zufallsvariable** bzw.  **$n$ -dimensionaler Zufallsvektor** über  $(\Omega, \mathcal{F}, P)$  und die gemäß Definition 8.3 gebildete Bildwahrscheinlichkeit

$$P_{\mathbf{X}} : \mathcal{B}^n \rightarrow \mathbb{R}; \mathbf{B} \mapsto P(\mathbf{X}^{-1}(\mathbf{B}))$$

(gemeinsame) **Wahrscheinlichkeitsverteilung** oder kürzer **(gemeinsame) Verteilung** von  $\mathbf{X}$ .  $(\mathbb{R}^n, \mathcal{B}^n, P_{\mathbf{X}})$  ist damit ebenfalls ein Wahrscheinlichkeitsraum. Liegt nach Durchführung des Zufallsexperiments  $(\Omega, \mathcal{F}, P)$  das Ergebnis  $\omega \in \Omega$  vor, so heißt der zugehörige Wert  $\mathbf{x} = \mathbf{X}(\omega)$  die **Realisierung** oder **Realisation** von  $\mathbf{X}$ .

# Mehrdimensionale Zufallsvariablen/Zufallsvektoren III

- Wie im eindimensionalen Fall sind Kurzschreibweisen (zum Beispiel) der Form

$$P\{X_1 \leq x_1, \dots, X_n \leq x_n\} := P_{\mathbf{X}}((-\infty, x_1] \times \dots \times (-\infty, x_n])$$

für  $x_1, \dots, x_n \in \mathbb{R}$  geläufig.

- Auch hier legen die Wahrscheinlichkeiten solcher Ereignisse die Verteilung des  $n$ -dimensionalen Zufallsvektors bereits eindeutig fest, und man definiert analog zum eindimensionalen Fall die gemeinsame Verteilungsfunktion

$$F_{\mathbf{X}} : \mathbb{R}^n \rightarrow \mathbb{R}; F_{\mathbf{X}}(x_1, \dots, x_n) := P_{\mathbf{X}}((-\infty, x_1] \times \dots \times (-\infty, x_n]) .$$

- Gemeinsame Verteilungsfunktionen mehrdimensionaler Zufallsvariablen sind allerdings für den praktischen Einsatz im Vergleich zur eindimensionalen Variante relativ unbedeutend und werden daher hier nicht weiter besprochen.

# Diskrete Zufallsvektoren I

- Ist analog zum eindimensionalen Fall

$$\mathbf{X}(\Omega) := \{\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n \mid \mathbf{x} = \mathbf{X}(\omega) \text{ für (mindestens) ein } \omega \in \Omega\}$$

endlich oder abzählbar unendlich bzw. existiert (wiederum etwas allgemeiner) eine endliche oder abzählbar unendliche Menge  $\mathbf{B} \subseteq \mathbb{R}^n$  mit  $P(\{\mathbf{X} \in \mathbf{B}\}) = 1$ , so nennt man auch solche Zufallsvektoren „diskret“.

- Mit Hilfe einer (mehrdimensionalen) Wahrscheinlichkeitsfunktion  $p_{\mathbf{X}}$  mit  $p_{\mathbf{X}}(\mathbf{x}) := P_{\mathbf{X}}(\{\mathbf{x}\})$  für  $\mathbf{x} \in \mathbb{R}^n$  können Wahrscheinlichkeiten  $P\{\mathbf{X} \in \mathbf{A}\}$  für Ereignisse  $\mathbf{A} \in \mathcal{B}^n$  wiederum durch Aufsummieren der Punktwahrscheinlichkeiten aller Trägerpunkte  $\mathbf{x}_i$  mit  $\mathbf{x}_i \in \mathbf{A}$  berechnet werden, also durch:

$$P\{\mathbf{X} \in \mathbf{A}\} = \sum_{\mathbf{x}_i \in \mathbf{A} \cap T(\mathbf{X})} p_{\mathbf{X}}(\mathbf{x}_i) \quad \text{für alle } \mathbf{A} \in \mathcal{B}^n$$

# Diskrete Zufallsvektoren II

## Definition 10.2 (Diskreter Zufallsvektor)

Seien  $(\Omega, \mathcal{F}, P)$  ein Wahrscheinlichkeitsraum,  $n \in \mathbb{N}$ ,  $\mathbf{X}$  ein  $n$ -dimensionaler Zufallsvektor über  $(\Omega, \mathcal{F}, P)$  und  $\mathbf{B} \subseteq \mathbb{R}^n$  endlich oder abzählbar unendlich mit  $P(\{\mathbf{X} \in \mathbf{B}\}) = 1$ . Dann nennt man

- ▶  $\mathbf{X}$  einen **diskreten Zufallsvektor**,
- ▶  $p_{\mathbf{X}} : \mathbb{R}^n \rightarrow [0, 1]$ ;  $p_{\mathbf{X}}(\mathbf{x}) := P_{\mathbf{X}}(\{\mathbf{x}\})$  die (**gemeinsame**) **Wahrscheinlichkeitsfunktion** von  $\mathbf{X}$ ,
- ▶  $T(\mathbf{X}) := \{\mathbf{x} \in \mathbb{R}^n \mid p_{\mathbf{X}}(\mathbf{x}) > 0\}$  den **Träger** von  $\mathbf{X}$  sowie alle Elemente  $\mathbf{x} \in T(\mathbf{X})$  **Trägerpunkte** von  $\mathbf{X}$  und deren zugehörige Wahrscheinlichkeitsfunktionswerte  $p_{\mathbf{X}}(\mathbf{x})$  **Punktwahrscheinlichkeiten**.

## 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.
- Diskrete 2-dimensionale Zufallsvektoren  $\mathbf{X} = (X, Y)$  werden in der Regel durch die Angabe der gemeinsamen Wahrscheinlichkeitsfunktion

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

spezifiziert.

- Ist  $\mathbf{X} = (X, Y)$  dabei 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:

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



# 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:

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

## 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:

$X \setminus Y$	$y_1$	$y_2$	$\cdots$	$y_l$	$p_{i\cdot}$
$x_1$	$p_{11}$	$p_{12}$	$\cdots$	$p_{1l}$	$p_{1\cdot}$
$x_2$	$p_{21}$	$p_{22}$	$\cdots$	$p_{2l}$	$p_{2\cdot}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
$x_k$	$p_{k1}$	$p_{k2}$	$\cdots$	$p_{kl}$	$p_{k\cdot}$
$p_{\cdot j}$	$p_{\cdot 1}$	$p_{\cdot 2}$	$\cdots$	$p_{\cdot l}$	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 10.3

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  $j \in \{1, \dots, l\}$ .

# Bedingte Verteilungen 2-dim. Zufallsvektoren I

## Definition 10.4 (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  $X = x$ .**

# Bemerkungen I

- Mit den üblichen Bezeichnungen und Abkürzungen für zweidimensionale Zufallsvektoren mit endlichem Träger erhält man mit Definition 10.4 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.

# Bemerkungen II

- Durch 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 bekannte 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.  $\text{Var}(Y|X = x)$ ,
- Den bedingten Erwartungswert  $E(Y|X = x)$  erhält man im diskreten Fall zum Beispiel durch:  $E(Y|X = x) = \sum_{y_i} y_i \cdot p_{Y|X=x}(y_i)$ .
- Gültig bleibt auch der VZS:  $\text{Var}(X|Y = y) = E(X^2|Y = y) - [E(X|Y = y)]^2$

# Erwartungswert von $G(X, Y)$

## Definition 10.5

Es seien  $(X, Y)$  ein zweidimensionaler diskreter Zufallsvektor und  $G : \mathbb{R}^2 \rightarrow \mathbb{R}$  eine  $\mathcal{B}^2 - \mathcal{B}$ -messbare Abbildung. Sind  $(x_i, y_j)$  die Trägerpunkte sowie  $p_{(X, Y)}$  die Wahrscheinlichkeitsfunktion von  $(X, Y)$  und gilt

$$\sum_{x_i} \sum_{y_j} |G(x_i, y_j)| \cdot p_{(X, Y)}(x_i, y_j) < \infty ,$$

dann existiert der Erwartungswert  $E(G(X, Y))$  und es gilt

$$E(G(X, Y)) = \sum_{x_i} \sum_{y_j} G(x_i, y_j) \cdot p_{(X, Y)}(x_i, y_j).$$



- *Beispiel 1:*  $G(X, Y) = a \cdot X + b \cdot Y + c$  für  $a, b, c \in \mathbb{R}$

Wegen der Linearität von Summenbildung und Integration gilt stets:

$$E(a \cdot X + b \cdot Y + c) = a \cdot E(X) + b \cdot E(Y) + c$$

- *Beispiel 2:*  $G(X, Y) = X \cdot Y$

Hier gilt im allgemeinen **nicht**  $E(X \cdot Y) = E(X) \cdot E(Y)$ !

Sind  $X$  und  $Y$  allerdings **stochastisch unabhängig**, so gilt im diskreten Fall wegen  $p_{(X,Y)}(x, y) = p_X(x) \cdot p_Y(y)$  insgesamt

$$\begin{aligned} E(X \cdot Y) &= \sum_{x_i} \sum_{y_j} x_i \cdot y_j \cdot p_X(x_i) \cdot p_Y(y_j) \\ &= \sum_{x_i} x_i \cdot p_X(x_i) \cdot \sum_{y_j} y_j \cdot p_Y(y_j), \end{aligned}$$

womit man **in diesem Fall** speziell  $E(X \cdot Y) = E(X) \cdot E(Y)$  erhält.

# Abhängigkeitsmaße

- Analog zur deskriptiven Statistik ist man an Maßzahlen für die Abhängigkeit von zwei Zufallsvariablen  $X$  und  $Y$  über demselben Wahrscheinlichkeitsraum interessiert.
- Bei stochastischer Unabhängigkeit von  $X$  und  $Y$  sollten diese Maßzahlen naheliegenderweise den Wert 0 annehmen.
- Wie in deskriptiver Statistik: Maß für **lineare** Abhängigkeit von  $X$  und  $Y$  vordergründig:

## Definition 10.6 (Kovarianz)

Es sei  $(X, Y)$  ein zweidimensionaler Zufallsvektor. Den Erwartungswert

$$\sigma_{XY} := \text{Cov}(X, Y) := E[(X - E(X)) \cdot (Y - E(Y))]$$

nennt man (im Falle der Existenz) **Kovarianz** zwischen  $X$  und  $Y$ .

- Eine dem Varianzzerlegungssatz ähnliche Rechenvorschrift zeigt man auch leicht für die Berechnung der Kovarianz:

# Rechenregeln für Kovarianzen

## Satz 10.1 (Kovarianzzerlegungssatz)

Ist  $(X, Y)$  ein zweidimensionaler Zufallsvektor, so gilt (im Fall der Existenz der beteiligten Erwartungswerte)

$$\text{Cov}(X, Y) = E(X \cdot Y) - E(X) \cdot E(Y)$$

Sind  $X, Y$  und  $Z$  Zufallsvariablen (über demselben Wahrscheinlichkeitsraum  $(\Omega, \mathcal{F}, P)$ ) und  $a, b \in \mathbb{R}$ , so gelten außerdem die folgenden Rechenregeln:

- 1  $\text{Cov}(aX, bY) = ab \text{Cov}(X, Y)$
- 2  $\text{Cov}(X + a, Y + b) = \text{Cov}(X, Y)$   
(Translationsinvarianz)
- 3  $\text{Cov}(X, Y) = \text{Cov}(Y, X)$   
(Symmetrie)
- 4  $\text{Cov}(X + Z, Y) = \text{Cov}(X, Y) + \text{Cov}(Z, Y)$
- 5  $\text{Cov}(X, X) = \text{Var}(X)$
- 6  $X, Y$  stochastisch unabhängig  $\Rightarrow \text{Cov}(X, Y) = 0$

- Die erreichbaren Werte der Größe  $\text{Cov}(X, Y)$  hängen nicht nur von der Stärke der linearen Abhängigkeit ab, sondern (wie insbesondere aus Rechenregel 1 von Folie 269 ersichtlich) auch von der Streuung von  $X$  bzw.  $Y$ .
- Wie in deskriptiver Statistik: Alternatives Abhängigkeitsmaß mit normiertem „Wertebereich“, welches invariant gegenüber Skalierung von  $X$  bzw.  $Y$  ist.
- Hierzu Standardisierung der Kovarianz über Division durch Standardabweichungen von  $X$  und  $Y$  (falls möglich!):

### Definition 10.7

Es sei  $(X, Y)$  ein zweidimensionaler Zufallsvektor mit  $\text{Var}(X) = \sigma_X^2 > 0$  und  $\text{Var}(Y) = \sigma_Y^2 > 0$ . Man nennt

$$\rho_{XY} := \text{Korr}(X, Y) := \frac{\sigma_{XY}}{\sigma_X \cdot \sigma_Y} = \frac{\text{Cov}(X, Y)}{+\sqrt{\text{Var}(X) \cdot \text{Var}(Y)}}$$

den **Korrelationskoeffizienten** (nach Bravais-Pearson) zwischen  $X$  und  $Y$ .

# Rechenregeln für Korrelationskoeffizienten

Sind  $X$  und  $Y$  Zufallsvariablen (über denselben Wahrscheinlichkeitsraum  $(\Omega, \mathcal{F}, P)$ ) mit  $\text{Var}(X) > 0$ ,  $\text{Var}(Y) > 0$  und  $a, b \in \mathbb{R}$ , so gilt:

$$\textcircled{1} \text{Korr}(aX, bY) = \begin{cases} \text{Korr}(X, Y) & \text{falls } a \cdot b > 0 \\ -\text{Korr}(X, Y) & \text{falls } a \cdot b < 0 \end{cases}$$

$$\textcircled{2} \text{Korr}(X + a, Y + b) = \text{Korr}(X, Y)$$

(*Translationsinvarianz*)

$$\textcircled{3} \text{Korr}(X, Y) = \text{Korr}(Y, X)$$

(*Symmetrie*)

$$\textcircled{4} -1 \leq \text{Korr}(X, Y) \leq 1$$

$$\textcircled{5} \text{Korr}(X, X) = 1$$

$$\textcircled{6} \left. \begin{array}{l} \text{Korr}(X, Y) = 1 \\ \text{Korr}(X, Y) = -1 \end{array} \right\} \text{genau dann, wenn } Y = aX + b \text{ mit } \begin{cases} a > 0 \\ a < 0 \end{cases}$$

$$\textcircled{7} X, Y \text{ stochastisch unabhängig} \Rightarrow \text{Korr}(X, Y) = 0$$

Zufallsvariablen  $X, Y$  mit  $\text{Cov}(X, Y) = 0$  (!) heißen **unkorreliert**.

# Zusammenhang Unkorreliertheit und Unabhängigkeit

- Offensichtlich gilt stets (falls die Kovarianz existiert):

$$X, Y \text{ stochastisch unabhängig} \Rightarrow X, Y \text{ unkorreliert}$$

- Die Umkehrung ist allerdings *im allgemeinen* falsch, es gilt **außer in speziellen Ausnahmefällen**:

$$X, Y \text{ unkorreliert} \not\Rightarrow X, Y \text{ stochastisch unabhängig}$$

# Varianzen von Summen zweier Zufallsvariablen

- Durch Verknüpfung verschiedener Rechenregeln aus Folie 269 lässt sich leicht zeigen, dass stets

$$\text{Var}(X + Y) = \text{Var}(X) + 2 \text{Cov}(X, Y) + \text{Var}(Y)$$

bzw. für  $a, b, c \in \mathbb{R}$  allgemeiner stets

$$\text{Var}(aX + bY + c) = a^2 \text{Var}(X) + 2ab \text{Cov}(X, Y) + b^2 \text{Var}(Y)$$

gilt.

- **Nur für unkorrelierte** (also insbesondere auch für stochastisch unabhängige) Zufallsvariablen  $X$  und  $Y$  gilt offensichtlich spezieller

$$\text{Var}(aX + bY + c) = a^2 \text{Var}(X) + b^2 \text{Var}(Y) .$$

Dies kann für mehr als zwei Zufallsvariablen weiter verallgemeinert werden.