Mehrdimensionale Zufallsvariable

Wenn $X_1, X_2, \ldots, X_k$ Zufallsvariable sind, heißt $\mathbf{X} = (X_1, X_2, \ldots, X_k)$ *mehrdimensionale Zufallsvariable*:

$$\mathbf{X} : \Omega \to \mathbb{R}^k$$

Wenn $\omega \in \Omega$ eintritt, heißt der Vektor $\mathbf{X}(\omega) = (X_1(\omega), X_2(\omega), \ldots, X_k(\omega))$ *Realisierung* von $\mathbf{X}$.

Die *gemeinsame Verteilungsfunktion* der Zufallsvariablen $X_1, X_2, \ldots, X_k$ ist die Funktion $F_{\mathbf{X}} : \mathbb{R}^k \to [0, 1]$ mit

$$F_{\mathbf{X}}(x_1, x_2, \ldots, x_k) = P(X_1 \leq x_1,\ X_2 \leq x_2,\ \ldots,\ X_k \leq x_k)$$

für alle $x_1, x_2, \ldots, x_k \in \mathbb{R}$.

Die Verteilungen der Zufallsvariablen $X_1, X_2, \ldots, X_k$ heißen *Randverteilungen*, und die zugehörigen Verteilungsfunktionen $F_{X_1}, F_{X_2}, \ldots, F_{X_k}$ (d.h. $F_{X_i} : \mathbb{R} \to [0, 1]$ mit $F_{X_i}(x) = P(X_i \leq x)$) heißen *Randverteilungsfunktionen*.

Zweidimensionale diskrete Zufallsvariable

Eine zweidimensionale Zufallsvariable $\mathbf{X} = (X_1, X_2) : \Omega \to \mathcal{X}_1 \times \mathcal{X}_2$ heißt *diskret*, wenn $\mathcal{X}_1 \times \mathcal{X}_2 \subset \mathbb{R}^2$ endlich oder abzählbar unendlich ist.

$\mathbf{X} = (X_1, X_2)$ ist genau dann diskret, wenn X_1 und X_2 diskret sind.

gemeinsame Wahrscheinlichkeitsverteilung: $f_{\mathbf{X}} : \mathcal{X}_1 \times \mathcal{X}_2 \to [0,1]$ mit
$f_{\mathbf{X}}(x_1, x_2) = P(X_1 = x_1,\, X_2 = x_2)$ für alle $(x_1, x_2) \in \mathcal{X}_1 \times \mathcal{X}_2$

gemeinsame Verteilungsfunktion: $F_{\mathbf{X}} : \mathbb{R}^2 \to [0,1]$ mit
$F_{\mathbf{X}}(x_1, x_2) = \sum_{x_1' \in \mathcal{X}_1 : x_1' \leq x_1} \sum_{x_2' \in \mathcal{X}_2 : x_2' \leq x_2} f_{\mathbf{X}}(x_1', x_2')$

Randwahrscheinlichkeitsverteilungen: $f_{X_1} : \mathcal{X}_1 \to [0,1]$ und $f_{X_2} : \mathcal{X}_2 \to [0,1]$ mit
$f_{X_1}(x_1) = P(X_1 = x_1) = \sum_{x_2 \in \mathcal{X}_2} f_{\mathbf{X}}(x_1, x_2)$ für alle $x_1 \in \mathcal{X}_1$ und
$f_{X_2}(x_2) = P(X_2 = x_2) = \sum_{x_1 \in \mathcal{X}_1} f_{\mathbf{X}}(x_1, x_2)$ für alle $x_2 \in \mathcal{X}_2$

Erwartungswert: $E\left(g(X_1, X_2)\right) = \sum_{x_1 \in \mathcal{X}_1} \sum_{x_2 \in \mathcal{X}_2} g(x_1, x_2)\, f_{\mathbf{X}}(x_1, x_2)$ für alle Funktionen $g : \mathbb{R}^2 \to \mathbb{R}$

Bedingte Verteilungen

Sei $\mathbf{X} = (X_1, X_2)$ eine diskrete zweidimensionale Zufallsvariable.

Die *bedingte Wahrscheinlichkeitsverteilung* von X_1 gegeben $X_2 = x_2$ (mit $P(X_2 = x_2) = f_{X_2}(x_2) > 0$) ist die Funktion $f_{X_1 \mid X_2}(\,\cdot\mid x_2) : \mathcal{X}_1 \to [0,1]$ mit

$$f_{X_1 \mid X_2}(x_1 \mid x_2) = P(X_1 = x_1 \mid X_2 = x_2) = \frac{f_{\mathbf{X}}(x_1, x_2)}{f_{X_2}(x_2)}$$

für alle $x_1 \in \mathcal{X}_1$.

Die *bedingte Wahrscheinlichkeitsverteilung* von X_2 gegeben $X_1 = x_1$ (mit $P(X_1 = x_1) = f_{X_1}(x_1) > 0$) ist die Funktion $f_{X_2 \mid X_1}(\,\cdot\mid x_1) : \mathcal{X}_2 \to [0,1]$ mit

$$f_{X_2 \mid X_1}(x_2 \mid x_1) = P(X_2 = x_2 \mid X_1 = x_1) = \frac{f_{\mathbf{X}}(x_1, x_2)}{f_{X_1}(x_1)}$$

für alle $x_2 \in \mathcal{X}_2$.

Die folgende Aussagen sind äquivalent:

- X_1 und X_2 sind unabhängig
- $f_{\mathbf{X}}(x_1, x_2) = f_{X_1}(x_1)\, f_{X_2}(x_2)$ für alle $(x_1, x_2) \in \mathcal{X}_1 \times \mathcal{X}_2$
- $f_{X_1 \mid X_2}(\,\cdot\mid x_2) = f_{X_1}$ für alle $x_2 \in \mathcal{X}_2$ mit $f_{X_2}(x_2) > 0$
- $f_{X_2 \mid X_1}(\,\cdot\mid x_1) = f_{X_2}$ für alle $x_1 \in \mathcal{X}_1$ mit $f_{X_1}(x_1) > 0$

Zweidimensionale stetige Zufallsvariable

Eine zweidimensionale Zufallsvariable $\mathbf{X} = (X_1, X_2) : \Omega \to \mathbb{R}^2$ heißt *stetig*, wenn es eine Funktion $f_\mathbf{X} : \mathbb{R}^2 \to [0, +\infty)$ gibt, so dass

$$F_\mathbf{X}(x_1, x_2) = P(X_1 \leq x_1,\, X_2 \leq x_2) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} f_\mathbf{X}(y_1, y_2)\, dy_2\, dy_1$$

für alle $x_1, x_2 \in \mathbb{R}$; eine solche Funktion $f_\mathbf{X}$ heißt *gemeinsame Dichte* von X_1 und X_2.

Wenn $\mathbf{X} = (X_1, X_2)$ stetig ist, sind X_1 und X_2 stetig.

Wahrscheinlichkeit eines Rechteckes: $P(a_1 \leq X_1 \leq b_1,\, a_2 \leq X_2 \leq b_2) = F_\mathbf{X}(b_1, b_2) - F_\mathbf{X}(b_1, a_2) - F_\mathbf{X}(a_1, b_2) + F_\mathbf{X}(a_1, a_2) = \int_{a_1}^{b_1} \int_{a_2}^{b_2} f_\mathbf{X}(x_1, x_2)\, dx_2\, dx_1$ für alle $a_1, b_1, a_2, b_2 \in \mathbb{R}$ mit $a_1 \leq b_1$ und $a_2 \leq b_2$

Randdichten: $f_{X_1} : \mathbb{R} \to [0, +\infty)$ und $f_{X_2} : \mathbb{R} \to [0, +\infty)$ mit $f_{X_1}(x_1) = \int_{-\infty}^{+\infty} f_\mathbf{X}(x_1, x_2)\, dx_2$ für alle $x_1 \in \mathbb{R}$ und $f_{X_2}(x_2) = \int_{-\infty}^{+\infty} f_\mathbf{X}(x_1, x_2)\, dx_1$ für alle $x_2 \in \mathbb{R}$

Erwartungswert: $E\left(g(X_1, X_2)\right) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x_1, x_2)\, f_\mathbf{X}(x_1, x_2)\, dx_2\, dx_1$ für alle Funktionen $g : \mathbb{R}^2 \to \mathbb{R}$

Bedingte Verteilungen

Sei $\mathbf{X} = (X_1, X_2)$ eine stetige zweidimensionale Zufallsvariable.

Die *bedingte Dichte* von X_1 gegeben $X_2 = x_2$ (mit $f_{X_2}(x_2) > 0$) ist die Funktion $f_{X_1 \mid X_2}(\cdot \mid x_2) : \mathbb{R} \to [0, +\infty)$ mit

$$f_{X_1 \mid X_2}(x_1 \mid x_2) = \frac{f_{\mathbf{X}}(x_1, x_2)}{f_{X_2}(x_2)}$$

für alle $x_1 \in \mathcal{X}_1$.

Die *bedingte Dichte* von X_2 gegeben $X_1 = x_1$ (mit $f_{X_1}(x_1) > 0$) ist die Funktion $f_{X_2 \mid X_1}(\cdot \mid x_1) : \mathbb{R} \to [0, +\infty)$ mit

$$f_{X_2 \mid X_1}(x_2 \mid x_1) = \frac{f_{\mathbf{X}}(x_1, x_2)}{f_{X_1}(x_1)}$$

für alle $x_2 \in \mathcal{X}_2$.

Die folgende Aussagen sind äquivalent:

- X_1 und X_2 sind unabhängig
- die Funktion $f : \mathbb{R}^2 \to [0, +\infty)$ mit $f(x_1, x_2) = f_{X_1}(x_1)\, f_{X_2}(x_2)$ ist eine gemeinsame Dichte von X_1 und X_2

Kovarianz

Kovarianz der Zufallsvariablen X und Y mit $\mu = E(X)$ und $\nu = E(Y)$:

$$Cov(X, Y) = E\left((X - \mu)\,(Y - \nu)\right)$$

Die Kovarianz ist ein Zusammenhangsmaß.

Eigenschaften der Kovarianz für Zufallsvariable $X, Y, X_1, X_2, \ldots, X_n, Y_1, Y_2, \ldots, Y_m$:

- $Var(X + Y) = Var(X) + Var(Y) + 2\,Cov(X, Y)$
- X und Y unabhängig $\Rightarrow Cov(X, Y) = 0$
- $X = c$ konstant (d.h. $X(\omega) = c$ für alle $\omega \in \Omega$) $\Rightarrow Cov(X, Y) = 0$
- $Cov(X, Y) = E(X\,Y) - E(X)\,E(Y)$
- $Var(X) = Cov(X, X)$
- *Symmetrie*: $Cov(X, Y) = Cov(Y, X)$
- *Bilinearität*: $Cov\left(\sum_{i=1}^{n} \alpha_i\, X_i,\ \sum_{j=1}^{m} \beta_j\, Y_j\right) = \sum_{i=1}^{n} \sum_{j=1}^{m} \alpha_i\, \beta_j\, Cov(X_i, Y_j)$ für alle $\alpha_1, \alpha_2, \ldots, \alpha_n, \beta_1, \beta_2, \ldots, \beta_m \in \mathbb{R}$
- $Var\left(\sum_{i=1}^{n} \alpha_i\, X_i\right) = \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i\, \alpha_j\, Cov(X_i, X_j) = \left(\sum_{i=1}^{n} \alpha_i^2\, Var(X_i)\right) + 2\left(\sum_{i,j \in \{1,2,\ldots,n\}: i<j} \alpha_i\, \alpha_j\, Cov(X_i, X_j)\right)$ für alle $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{R}$

Korrelation

Korrelationskoeffizient der Zufallsvariablen X und Y:

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)\,Var(Y)}} = Cov\left(\frac{X - E(X)}{\sqrt{Var(X)}}, \frac{Y - E(Y)}{\sqrt{Var(Y)}}\right)$$

Der Korrelationskoeffizient ist ein standardisiertes Zusammenhangsmaß ($\rho(X,Y)$ ist die Korrelation der zu X und Y gehörigen standardisierten Zufallsvariablen).

Die Zufallsvariablen X und Y heißen *unkorreliert*, wenn $\rho(X,Y) = 0$ (d.h. $Cov(X,Y) = 0$).

Eigenschaften des Korrelationskoeffizienten für Zufallsvariable X, Y:

- X und Y unabhängig $\Rightarrow$ X und Y unkorreliert
- X und Y unkorreliert $\Rightarrow Var(X+Y) = Var(X) + Var(Y)$
- *Symmetrie*: $\rho(X,Y) = \rho(Y,X)$
- $\rho(\alpha X + \beta, \gamma Y + \delta) = \rho(X,Y)$ für alle $\alpha, \beta, \gamma, \delta \in \mathbb{R}$ mit $\alpha\gamma > 0$
- $\rho(\alpha X + \beta, \gamma Y + \delta) = -\rho(X,Y)$ für alle $\alpha, \beta, \gamma, \delta \in \mathbb{R}$ mit $\alpha\gamma < 0$
- $\rho(X,Y) \in [-1, 1]$
- $\rho(X,Y) = 1 \Leftrightarrow P(Y = \alpha X + \beta) = 1$ mit $\alpha, \beta \in \mathbb{R}$ und $\alpha > 0$
- $\rho(X,Y) = -1 \Leftrightarrow P(Y = \alpha X + \beta) = 1$ mit $\alpha, \beta \in \mathbb{R}$ und $\alpha < 0$