

Kovarianz

- beschreibt Abhängigkeit zwischen zwei [[Zufallsvariable]]
- Berechnung über [[Erwartungswert]]

Sei (X, Y) ein Zufallsvektor mit Erwartungswert (μ, ν) und die Marginalen seien in L^2 . Dann nennen wir

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

die **Kovarianz zwischen X und Y** . Ist $\text{Cov}(X, Y) = 0$ so sagen wir, dass X und Y **unkorreliert** sind.

Wir schreiben dann: $X \perp Y$.

- Es gilt im Allgemeinen

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

- Sind X und Y diskret, dann gilt

$$E(XY) = \sum_k \sum_\ell x_k y_\ell P(X = x_k, Y = y_\ell).$$

- Haben (X, Y) eine Dichte $f(x, y)$, dann gilt

$$E(XY) = \int \int xyf(x, y) dx dy.$$

- $\text{Var}(X) = \text{Cov}(X, X)$

Korrelation

- Maß der Abhängigkeit zwischen Zufallsvariablen X und Y
- Korrelation zwischen X und Y

$$\text{Corr}(X, Y) := \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)}\sqrt{\text{Var}(Y)}} \leq 1.$$

- Beispiel

Betrachte den Zufallsvektor (X, Y) . Sowohl X als auch Y nimmt die Werte 0, 1, 2 an. Hier ist die PMF:

$x \backslash y$	0	1	2
0	0.1	0.2	0.1
1	0.2	0.1	0.1
2	0	0	0.2

Berechne $\text{Corr}(X, Y)$.

Korrelation von Zufallsvektor

$$\text{Corr}(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)} \sqrt{\text{Var}(Y)}}$$

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

$$E(X) = 0.8$$

$$E(Y) = 1.1$$

$$E(XY) = 1 \times 1 \times 0.1 + 1 \times 2 \times 0.1 + 2 \times 1 \times 0 + 2 \times 2 \times 0.2$$

$$= 0.1 + 0.2 + 0 + 0.8 = 1.1$$

$$\leadsto \text{Cov}(X, Y) = 1.1 - 0.8 \times 1.1 = 0.22$$

$$\leadsto \text{Var}(X) = E(X^2) - (E(X))^2 = E(X^2) - 0.64 = 1.2 - 0.64 = 0.56$$

$$E(X^2) = 1 \times 0.4 + 4 \times 0.2 = 1.2$$

Kovarianzmatrix

Betrachte einen Zufallsvektor mit Komponenten $X = (X_1, \dots, X_n)$ und sei

$$\text{Var}(X) = \Sigma = \begin{pmatrix} \text{Cov}(X_1, X_1) & \text{Cov}(X_1, X_2) & \dots & \text{Cov}(X_1, X_n) \\ \text{Cov}(X_2, X_1) & \text{Cov}(X_2, X_2) & \dots & \text{Cov}(X_2, X_n) \\ \vdots & \vdots & \ddots & \vdots \\ \text{Cov}(X_n, X_1) & \text{Cov}(X_n, X_2) & \dots & \text{Cov}(X_n, X_n) \end{pmatrix}.$$

Man nennt Σ die **Kovarianzmatrix** von X .

-
- Eigenschaften
 - Σ ist nicht-negativ definit [$a' \Sigma a \geq 0$ für alle $a \in \mathbb{R}^n$];
 - Σ ist symmetrisch.
- Herleitungen

Sei X ein Zufallsvektor in \mathbb{R}^n und $A \in \mathbb{R}^{m \times n}$ und $b \in \mathbb{R}^m$. Sei $\Sigma = \text{Var}(X)$. Dann gilt

$$\text{Var}(AX + b) = A\Sigma A'.$$

Seien $X, Y, X_i \in L^2$ und $a, b, c, d \in \mathbb{R}$. Dann gilt

(a) $aX + b$ und $cY + d$ sind in L^2 und

$$\text{Cov}(aX + b, cY + d) = ac \text{Cov}(X, Y).$$

Im Besonderen gilt $\text{Var}(aX + b) = a^2 \text{Var}(X)$.

(b) $\text{Var}(X_1 + \dots + X_n) = \sum_{i=1}^n \text{Var}(X_i) + \sum_{i \neq j} \text{Cov}(X_i, X_j)$.
Sind die X_i unkorreliert, dann

$$\text{Var}(X_1 + \dots + X_n) = \sum_{i=1}^n \text{Var}(X_i).$$