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

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

die Kovarianz zwischen X und Y. Ist Cov(X, Y) = 0 so sagen wir, dass X und Y unkorreliert sind.

Wir schreiben dann: $X \perp Y$.

Es gilt im Allgemeinen

$$Cov(X, Y) = E(XY) - E(X)E(Y).$$

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

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

 $\bullet \ Var(X) = Cov(X,X)$

Korrelation

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

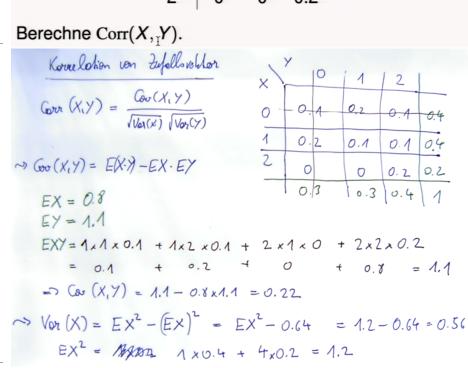
$$Corr(X, Y) := \frac{Cov(X, Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}} \le 1.$$

• Beispiel

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

$X \setminus Y$	0	1	2
0	0.1	0.2	0.1
1	0.2	0.1	0.1
2	0	0	0.2

Berechne Corr(X, $_{_{1}}Y$).



Kovarianzmatrix

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

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

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

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

Seien $X_{_{\mathbb{T}}}Y,X_{i}\in L^{2}$ und $a,b,c,d\in\mathbb{R}.$ Dann gilt

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

$$Cov(aX + b, cY + d) = ac Cov(X, Y).$$

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

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

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