Vorlesung 4 Wahrscheinlichkeitstheorie cont'd

Wiederholung: Bayes' Theorem sagt:

$$P(A|B) = P(B|A) \frac{P(A)}{P(B)}$$

Definition 4.1 Seien A, B Ereignisse. Wir nennen A und unabhängig, falls:

$$P(A|B) = P(A).$$

Seien X, Y Zufallsvariablen, dann nennen wir sie unabhängig, falls

$$P(X \in A \text{ und } Y \in A) = P(X \in A)\, P(Y \in A)$$

für alle Ereignisse A. Wir nennen $(X_n)_{n \in \mathbb{N}}$ unabhängig und identisch verteilt (iid), falls sie paarweise unabhängig sind und alle die gleiche Zufallsverteilung.

Sei nun $X \in \mathbb{R}$ eine ZV. Falls der Wertbereich $X \in \{x_1, \ldots, x_n\}$ endlich (oder diskret) ist, dann ist X vollständig durch Angabe von $\underset{\substack{\| \\ P(X \in \{x_i\})}}{P(X = x_i)}$, $i = 1, \ldots, n$, charakterisiert.

z.B: $X \in \{0,1\}$, dann ist X durch Angabe von $P(X=0) =: p \in [0,1]$ und $P(X=1) = 1-p$ vollständig charakterisiert.

Falls nun aber der Wertbereich von X eine kontinuierliche Teilmenge von $\mathbb{R}$ enthält, gilt obiges nicht mehr. Falls der Wertebereich von X das Intervall $[a,b]$ enthält, dann muss $P(X = x) = 0$ für alle $x \in [a,b]$.

In vielen Fällen haben kontinuierliche ZV $X \in \mathbb{R}$ aber eine Dichte, mit der wir die Verteilung von X beschreiben können.

Definition 4.2

Sei $X \in \mathbb{R}$ eine kontinuierliche ZV. Eine integrierbare Funktion $f: \mathbb{R} \to \mathbb{R}_{\geq 0}$ heißt Dichte von X, falls für alle Ereignisse A gilt:

$$P(X \in A) = \int_A f(x)\,dx, \quad A \subset \mathbb{R}.$$

(Es gilt dann $\int_{\mathbb{R}} f(x)\,dx = 1$).

Falls $X \in \mathbb{R}^n$, $X = (X_1, \dots, X_n)$, eine kontinuierliche ZV ist, nennen wir eine Funktion $f: \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ eine Dichte, falls

$$P(X \in A) = \int_A f(x)\,dx, \quad A \subset \mathbb{R}^n.$$

Wir nennen f auch die gemeinsame Dichte der X_i.

Definition 4.3

Sei $(X, Y) \in \mathbb{R}^n \times \mathbb{R}^m$ eine ZV mit Dichte $P_{(X,Y)}$. Die bedingte Dichte von Y gegeben $X = x$, $x \in \mathbb{R}^n$, ist:

$$P_{Y|X=x}(y) = \frac{P_{(X,Y)}(x, y)}{P_X(x)},$$

wobei P_X die Dichte von X ist.

Wir nennen $Y|X=x$ die ZV mit der Dichte $P_{Y|X=x}$.

Es gilt: $P_X(x) = \int_{\mathbb{R}^m} P_{(X,Y)}(x, y)\,dy$, weil:

$$P(X \in A) = P(X \in A \text{ und } Y \in \mathbb{R}^m)$$

Im Folgenden schreiben wir $P_X(x)$ für die Dichte von $X \in \mathbb{R}$, so dass $$P(X\in A) = \int_A p_X(x)\,dx.$$

Theorem 4.4 (Bayes' Theorem für Dichten)

Sei $(X,Y) \in \mathbb{R}^n \times \mathbb{R}^m$ eine ZV mit Dichte $p_{(X,Y)}$, und $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$. Dann:

$$p_{Y|X=x}(y) = p_{X|Y=y}(x)\,\frac{p_Y(y)}{p_X(x)}$$

Beweis

$$p_{Y|X=x}(y) = \frac{p_{(X,Y)}(x,y)}{p_X(x)}$$ und

$$p_{X|Y=y}(x) = \frac{p_{(X,Y)}(x,y)}{p_Y(y)}$$ ☐.

Als nächstes führen wir wichtige Kennzahlen von ZV ein:

Definition 4.5

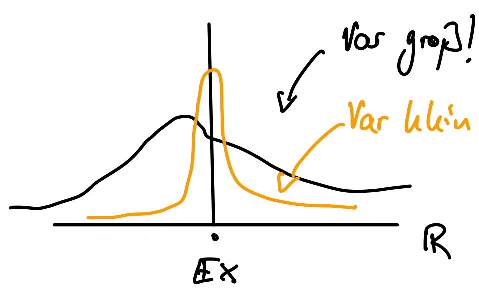
Sei $X \in \mathbb{R}$ eine ZV.

1) Sei $X \in \{x_1, x_2, \ldots\}$ diskret. Der Erwartungswert von X ist

$$\mathbb{E}X := \sum_{i=1}^{\infty} x_i\, P(X=x_i).$$

2) Sei $X \in \mathbb{R}$ kontinuierlich mit Dichte p_X. Der Erwartungswert von X ist

$$\mathbb{E}X := \int_{\mathbb{R}} x \cdot p_X(x)\,dx$$



3) In beiden Fällen ist die Varianz

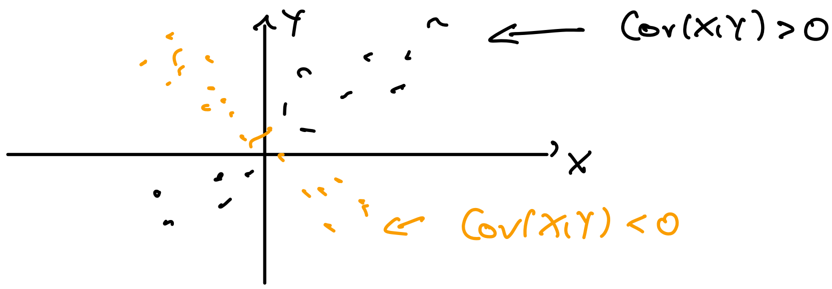
$$\mathrm{Var}(X) = \mathbb{E}(X - \mathbb{E}X)^2.$$

Die Standardabweichung ist $s(X) = \sqrt{Var(X)}$

4) Seien X,Y zwei Zufallsvariablen. Dann ist ihre Kovarianz

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

Beispiel:



Insbesondere: $Var(X) = Cov(X,X)$.

<u>Lemma</u> 4.7

Seien X,Y zwei Zufallsvariablen, $\mathbb{E}X, \mathbb{E}Y < \infty$. Dann gilt für alle $a,b \in \mathbb{R}$,

$$\mathbb{E}\; aX+bY = a\,\mathbb{E}X + b\,\mathbb{E}Y.$$

<u>Beweis</u>

In der Übung.

Lemma 4.7 impliziert:

$$Var(X) = \mathbb{E}\,(X-\mathbb{E}X)^2 = \mathbb{E}\,(X^2 - X\mathbb{E}X + (\mathbb{E}X)^2) = \mathbb{E}\,X^2 - (\mathbb{E}X)^2$$

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

<u>Lemma 4.8</u>

Sei $X \in \mathbb{R}^n$ eine ZV und $g: \mathbb{R}^n \to \mathbb{R}$ (g ist <u>messbar</u>)

1) Falls $X \in \{x_1, x_2, \ldots\}$ diskret ist, gilt,

$$\mathbb{E}\; g(X) = \sum_{i=1}^{\infty} g(x_i)\, P(X=x_i).$$

2) Falls $X \in \mathbb{R}$ kontinuierlich ist mit Dichte P_X, gilt:

$$\mathbb{E}\, g(X) = \int_{\mathbb{R}^n} g(x)\, P_X(x)\, dx.$$

falls $\int_{\mathbb{R}^n} |g(x)|\, P_X(x)\, dx < \infty$.

<u>Beweis von Teil (1)</u>

Sei $Z := g(X)$. Dann gilt nach Definition vom Erwartungswert:

$$\mathbb{E}\, Z = \sum_{i=1}^{\infty} z_i\, P_Z(Z = z_i),$$

wobei $z_i = g(x_i)$.

(denn $X \in \{x_1, x_2, \dots\}$ impliziert $Z \in \{g(x_1), g(x_2), \dots\}$)

D.h.

$$\mathbb{E} Z = \sum_{i=1}^{\infty} g(x_i)\, P_Z(Z = g(x_i))$$

$$= \sum_{i=1}^{\infty} g(x_i)\, P_X(X = x_i).$$

Beachte:

$$\underbrace{Z = g(x_i) \Leftarrow X = x_i}$$

$$\Downarrow$$

$$P(Z = z_i) = \sum_{k:\, g(x_k) = z_i} P(X = x_k).$$

Insbesondere gilt nach Lemma 4.8:

$$\mathrm{Cov}(X, Y) = \mathbb{E} XY - \mathbb{E} X \mathbb{E} Y \qquad \text{mit } g(X, Y) = XY$$

$$= \int_{\mathbb{R}} xy\, P_{(X,Y)}(x,y)\, d(x,y) - \mathbb{E} X \mathbb{E} Y.$$

<u>Beispiel 4.9</u>

Folgende Verteilungen sind wichtig:

1) <u>Bernoulli-Verteilung</u>: $X \in \{0, 1\}$ und $P(X = 0) = p \in [0, 1]$

Wir schreiben $X \sim \mathrm{Ber}(p)$

2) <u>Binomial-Verteilung</u> $X \in \{0, \ldots, n\}$, $P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$, $p \in [0,1]$.
Wir schreiben $X \sim \mathrm{Bin}(n,p)$.
<u>Beachte</u>: $P(X=k) = P(\#\{i \mid Z_i = 0,\ 1 \le i \le n\} = k)$, $Z_1, \ldots, Z_n \overset{iid}{\sim} \mathrm{Ber}(p)$

3) <u>Diskrete Gleichverteilung</u> $X \in \{a_1, \ldots, a_n\}$ und $P(X=a_i) = \frac{1}{n}$
Wir schreiben $X \sim \mathrm{Unif}(\{a_1, \ldots, a_n\})$

4) <u>Kontinuierliche Gleichverteilung</u> $X \in [a,b] \subset \mathbb{R}$, $P(A) = \int_A \frac{1}{b-a}\, dx$, $A \subset [a,b]$.
Wir schreiben $X \sim \mathrm{Unif}([a,b])$.

5) <u>Normalverteilung</u> $X \in \mathbb{R}$ und für $A \subset \mathbb{R}$:

$$P(X \in A) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_A e^{-1/2\sigma^2 (x-\mu)^2}\, dx$$

mit $\sigma^2 > 0$, $\mu \in \mathbb{R}$. Wir schreiben $X \sim N(\mu, \sigma^2)$.
In der Übung: $\mathbb{E} X = \mu$, $\mathrm{Var}(X) = \sigma^2$.

6) <u>Multivariate Normalverteilung</u> $X \in \mathbb{R}^n$ und für $A \subset \mathbb{R}^n$:

$$P(X \in A) = \frac{1}{\sqrt{(2\pi)^n \det(\Sigma)}} \int_A e^{-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)}\, dx,$$

$\Sigma \in \mathbb{R}^{n \times n}$ positiv-definit. und $\mu \in \mathbb{R}^n$. Wir schreiben $X \sim N(\mu, \Sigma)$

Wir schreiben auch:

$$\Phi(x \mid \mu, \Sigma) = \frac{1}{\sqrt{(2\pi)^n \det(\Sigma)}} e^{-1/2 (x-\mu)^T \Sigma^{-1} (x-\mu)}$$

für die Dichte.

Lemma 4.10

Sei $X \sim N(\mu, \Sigma)$, $X = (X_1, \dots, X_n)$. Dann gilt:

$$\mathrm{Cov}(X_i, X_j) = \Sigma_{ij}$$

Deswegen heißt Σ auch Kovarianzmatrix

Beweis

Übung.