Sei A ein Ereignis und $B_1, \dots, B_n$ eine disjunkte Zerlegung von Ω. Dann gilt:

$$\mathbb{P}(B_i \mid A) = \frac{\mathbb{P}(A \mid B_i)\,\mathbb{P}(B_i)}{\sum_{j=1}^{n} \mathbb{P}(A \mid B_j)\,\mathbb{P}(B_j)}$$

$(\Omega, \mathbb{P})$ ein Wahrscheinlichkeitsraum. Eine Zufallsvariable ist eine Abbildung $X: \Omega \to \mathbb{R}$.

Wertebereich von X: $X(\Omega) := \{x \in \mathbb{R} \mid \text{es ex. } \omega \in \Omega \text{ mit } X(\omega) = x\}$

Sei X eine diskrete Zufallsvariable. Dann gilt:

1. $\sum_{x \in X(\Omega)} P_X(x) = 1$ 2. $F_X(x) = \sum_{y \leq x} P_X(y)$

Sei X, Y diskrete ZV mit existierten Erwartungswert $\sum_{x \in X(\Omega)} |x|\,\mathbb{P}(X=x) < \infty$

① $\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$

② $\mathbb{E}[aX] = a\,\mathbb{E}[X]$

③ Falls $X(\omega) \geq 0$ für alle $\omega \in \Omega$, so ist $\mathbb{E}[X] \geq 0$

④ Falls X, Y unabhängig, so gilt $\mathbb{E}[XY] = \mathbb{E}[X] \cdot \mathbb{E}[Y]$

Sei X diskrete ZV. $\mathbb{V}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$

Varianz existiert genau dann, wenn $\mathbb{E}[X^2]$ existiert.

$\delta(X) = \sqrt{\mathbb{V}[X]}$

① $\mathbb{V}[aX+b] = a^2\,\mathbb{V}[X]$

② Falls X, Y unabhängig: $\mathbb{V}[X+Y] = \mathbb{V}[X] + \mathbb{V}[Y]$

③ Falls X konstant: $\mathbb{V}[X] = 0$

Chebyshev-Ungleichung:

$$\mathbb{P}(|X - \mathbb{E}[X]| \geq \varepsilon) \leq \frac{\mathbb{V}[X]}{\varepsilon^2} \xrightarrow[\varepsilon \to \infty]{} 0$$

Markov-Ungleichung:

Sei X ZV. $f: \mathbb{R} \to [0, \infty)$ monoton wachsende Fkt.

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}[f(X)]}{f(a)}$$

Covarianz: $\mathrm{Cov}(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X] \cdot \mathbb{E}[Y]$

>0 positiv korreliert

$=0$ unkorreliert

① $\mathrm{Cov}(X,Y) = \mathrm{Cov}(Y,X)$

② $\mathrm{Cov}(X,X) = \mathbb{V}[X]$

③ $\mathrm{Cov}(aX+bY, Z) = a\,\mathrm{Cov}(X,Z) + b\,\mathrm{Cov}(Y,Z)$

④ Falls X, Y unabhängig: $\mathrm{Cov}(X,Y) = 0$

Unabhängig $\overset{\Rightarrow}{\not\Leftarrow}$ unkorreliert

$$\mathbb{V}[X+Y] = \mathbb{V}[X] + \mathbb{V}[Y] + 2\,\mathrm{Cov}(X,Y)$$

Korrelation: $\mathrm{Corr}(X,Y) = \rho(X,Y) = \dfrac{\mathrm{Cov}(X,Y)}{\sqrt{\mathbb{V}[X]\,\mathbb{V}[Y]}}$

$-1 \leq \mathrm{Corr}(X,Y) \leq 1$

Diskrete Verteilung:

Bernoulli-Verteilung: $X \sim \mathrm{Bernoulli}(p)$

$\mathbb{E}[X] = p \qquad \mathbb{V}[X] = p(1-p)$

Binomialverteilung: $X \sim \mathrm{Bin}(n,p)$

$\mathbb{P}(X=k) = \binom{n}{k} p^k (1-p)^{n-k} = \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}$

$\mathbb{E}[X] = np \qquad \mathbb{V}[X] = np(1-p)$

Geometrische Verteilung: $X \sim \mathrm{Geom}(p)$

$\mathbb{P}(X=k) = (1-p)^{k-1} p \qquad k \in \mathbb{N}$

$\mathbb{E}[X] = \frac{1}{p} \qquad \mathbb{V}[X] = \frac{1-p}{p^2}$

Poissonverteilung: $X \sim \mathrm{Poi}(\lambda)$

$\mathbb{P}(X=k) = \frac{\lambda^k}{k!} e^{-\lambda} \qquad k \in \mathbb{N}_0$

$\mathbb{E}[X] = \lambda \qquad \mathbb{V}[X] = \lambda$

Zipf-Verteilung:

Sei $a > 1$, $X \sim \mathrm{Zipf}(a)$, falls $X(\Omega) = \mathbb{N}_0$

und $\mathbb{P}(X=k) = \frac{k^{-a}}{\zeta(a)}$ wobei $\zeta(a) = \sum_{k=1}^{\infty} k^{-a}$

$\mathbb{E}[X] = \frac{\zeta(a-1)}{\zeta(a)} \quad a > 2 \qquad \mathbb{V}[X] = \frac{\zeta(2-a)\zeta(a) - \zeta(1-a)^2}{\zeta(a)^2} \quad a > 3$

Sei $X \sim \mathrm{Poi}(\lambda)$, $Y \sim \mathrm{Poi}(\mu) \rightsquigarrow X+Y \sim \mathrm{Poi}(\lambda+\mu)$

Poisson-Grenzwertsatz: mit $\lim_{n\to\infty} np_n = \lambda$

$\mathrm{Bin}(n,p) \xrightarrow{n > 30,\ p \leq 0{,}05} \mathrm{Poi}(\lambda) \quad (\lambda = np)$

Stetige Verteilung

Gleichverteilung auf $[a,b]$. $f_X(x) = \begin{cases} \frac{1}{b-a} & x \in [a,b] \\ 0 & \end{cases}$

$F_{(x)} = \begin{cases} 0 \\ \frac{x-a}{b-a} \\ 1 \end{cases} \qquad \mathbb{E}[X] = \frac{1}{2}(a+b) \qquad \mathbb{V}[X] = \frac{(b-a)^2}{12}$

Exponentialverteilung

$X \sim E(\lambda) \qquad f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \end{cases}$

$\mathbb{E}[X] = \frac{1}{\lambda} \qquad \mathbb{V}[X] = \frac{1}{\lambda^2}$

Normalverteilung

$X \sim \mathcal{N}(\mu, \delta^2) \qquad f_X(x) = \frac{1}{\sqrt{2\pi}\,\delta} e^{-\frac{(x-\mu)^2}{2\delta^2}}$

$\frac{X-\mu}{\delta} \sim \mathcal{N}(0,1) \qquad Z = X+Y \sim \mathcal{N}(\mu_1+\mu_2, \delta_1^2+\delta_2^2)$

Satz von großen Zahlen sei $X_1, X_2, \dots$ unab. iden. ZV mit $\mathbb{E}[X_i] = \mu$, $\mathbb{V}[X_i] = \delta^2$. Dann gilt: $\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} X_i = \mu$

f.a. $\varepsilon > 0$: $\lim_{n\to\infty} \mathbb{P}\left(\left| \frac{1}{n} \sum_{i=1}^{n} X_i - \mu \right| > \varepsilon \right) = 0$

Zentraler Grenzwertsatz

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{X_i - \mu}{\delta} \leq x \right) = \Phi(x)$$

$\pi = (\pi_a)_{a\in S}$ heißt invariante Verteilung falls

a) $(P^T - I)\pi = 0$

b) $\sum_{i\in I} \pi_i = 1$

c) $\forall i \in I,\ \pi_i \in [0,1]$

Eine Markov-Ketten heißt irreduzibel, falls es f.a. $a, b \in S$ einen Pfad von a nach b gibt.

... heißt aperiodisch, falls f.a. $a \in S$ gilt

$$ggT\{n \in \mathbb{N} : P_{aa}^{(n)} > 0\} = 1$$

- es gibt Pfade zwischen allen Zuständen und die Markov-Ketten ist irreduzibel
- Weil $P_{ii} > 0$ für jedes $i \in \{1,2,3\}$ gibt, erhalten wir $ggT\{n \in \mathbb{N} \mid P_{ii}^{(n)} > 0\} = 1$. Somit ist die Markov-Kette aperiodisch
- π^T ist die eindeutige invariante Verteilung der Markov-Ketten, da irreduzible Markov-Ketten auf endlichen Zustandsräumen eine eindeutige invariante Verteilung besitzen

	Mit Reihenfolge	Ohne Reihenfolge
Mit Zurück	n^k	$\binom{k+n-1}{k}$
ohne Zurück	$\frac{n!}{(n-k)!}$	$\binom{n}{k}$

empirische Kenngrößen

Sei $(x_1, \dots, x_n) \in \mathbb{R}^n$ ein Vektor von Messwert.

empirische Mittel: $\bar{\mu}_n(x_1, \dots, x_n) = \frac{1}{n} \sum_{i=1}^{n} x_i$

empirische Varianz: $\bar{\delta}_n^2(x_1, \dots, x_n) = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{\mu}_n)^2$

Der Schätzer Θ_n heißt erwartungstreu, falls $\mathbb{E}[\Theta_n(x_1, \dots, x_n)] = \Theta$

konsistent $\lim_{n\to\infty} \Theta_n(x_1, \dots, x_n) = \Theta$

effizient $\lim_{n\to\infty} \mathbb{V}[\Theta_n(x_1, \dots, x_n)] = 0$

Maximum-Likelihood-Schätzung

Die Likelihood Fkt: $L((x_1, \dots, x_n); \theta) = \mathbb{P}_\theta(X_1 = x_1, \dots, X_n = x_n) = \prod_{i=1}^{n} P_\theta(x_i)$

Log-Likelihood Fkt: $\ell((x_1, \dots, x_n); \theta) = \log L((x_1, \dots, x_n); \theta)$

BSP: $X_1, \dots, X_{10}$ mit $X_i \sim Ber(p)$ $p \in (0,1)$ Sei $(x_1, \dots, x_{10}) \in \{0,1\}^n$ eine Stichprobe mit $\sum_{i=1}^{10} x_i = 7$

① $L((x_1, \dots, x_{10}); p) = \prod_{i=1}^{10} \mathbb{P}(X_i = x_i) = \prod_{i=1}^{10} p^{x_i} (1-p)^{1-x_i}$

② $\ell(x_1, \dots, x_{10}, p) = \sum_{i=1}^{10} \left(x_i \log p + (1-x_i) \log(1-p) \right)$

$$\frac{d\ell}{dp} = \sum_{i=1}^{10} \left(\frac{x_i}{p} - \frac{1-x_i}{1-p} \right) \stackrel{!}{=} 0 \Rightarrow \sum_{i=1}^{10} \left(x_i(1-p) - p(1-x_i) \right) = \sum_{i=1}^{10} (x_i - p) = \sum_{i=1}^{10} x_i - 10p = 0$$

$$\Rightarrow \bar{p}_{10} = \frac{\sum_{i=1}^{10} x_i}{10}$$

$$\frac{d\ell}{d^2p} = \sum_{i=1}^{10} \left(-\frac{x_i}{p^2} + \frac{1-x_i}{(p-1)^2} \right) < 0 \text{ für } p \in (0,1)$$

$\Rightarrow \bar{p}_n$ Maximumstelle