

Disto Klausurzettel

- Stichprobenmittel: $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$
- Stichprobenmedian: $m = \frac{1}{2} \left(x_{\lfloor \frac{n}{2} \rfloor} + x_{\lceil \frac{n}{2} \rceil} \right)$
- Stichprobenvarianz: $s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$
- p-Quantil:
 - $np \notin \mathbb{N} : q_p = x_{\lceil np \rceil}$
 - $np \in \mathbb{N} : q_p = \frac{1}{2} (x_{np} + x_{np+1})$
- Interquartilsabstand: $\text{IQR} = q_{0.75} - q_{0.25}$
- Box-Plot:
 - Rechteck von $q_{0.25}$ bis $q_{0.75}$
 - Stab von $q_{0.75}$ bis $\max\{x_i : x_i < q_{0.75} + 1.5 \cdot \text{IQR}\}$
 - Stab von $q_{0.25}$ bis $\min\{x_i : x_i > q_{0.25} - 1.5 \cdot \text{IQR}\}$
 - Ausreißer markieren
- empirische Verteilungsfunktion: $F(x) = \frac{\#\{x_i : x_i \leq x\}}{n}$
- Laplace-Raum: $(P, \Omega) : P(A) = \frac{\#A}{\#\Omega}$
- Kombinatorik:
 - mit Z, mit R: n^k
 - ohne Z, mit R: $\frac{n!}{(n-k)!}$
 - ohne Z, ohne R: $\binom{n}{k} = \frac{n!}{k!(n-k)!}$
 - mit Z, ohne R: $\binom{k+n-1}{k} = \frac{(k+n-1)!}{k!(n-1)!}$
- fixpunktfreie Permutationen: $n! \cdot \sum_{j=0}^{n-1} \frac{(-1)^j}{j!}$
- Wkheitsfunktion: $p^X(x) = P(X = x), P^X(B) = P(X \in B)$
- Verteilungsfunktion: $F^X(x) = P(X \leq x)$
- Erwartungswert: $EX = \sum_{x \in X(\Omega)} x \cdot P(X = x)$
 - Der Erwartungswert ist linear und monoton
 - X_i unabhängig: $E(\prod X_i) = \prod EX_i$
 - Transformationsformel: $E(g \circ X) = \sum_{x \in X(\Omega)} g(x) \cdot P(X = x)$
- Unabhängigkeit: $P(A \cap B) = P(A) \cdot P(B)$
- Produktwkeitsraum: $(\Omega, P) : \Omega = \times_{i=1}^n \Omega_i, P(A) = \sum_{(a_1, \dots, a_n) \in A} P_1(a_1) \cdot \dots \cdot P_n(a_n)$
- Unabhängigkeit von Zufallsvariablen: $P(X_i = x_i) = \prod P(X_i = x_i)$
- bedingte Wkeit: Angenommen B, dann ist die Wkeit von A: $P(A|B) = \frac{P(A \cap B)}{P(B)}$
- Formel von Bayes: $\Omega = \sum_{i=1}^n B_i : P(B_i|A) = \frac{P(A|B_i) \cdot P(B_i)}{\sum_{j=1}^n P(A|B_j) \cdot P(B_j)}$
- bedingter Erwartungswert: $E(X|A) = \frac{1}{P(A)} \sum_{\omega \in A} X(\omega) \cdot P(\omega)$
- Varianz: $VX = E(X^2) - (EX)^2$
 - $V(aX + b) = a^2 VX$
 - X_i unabhängig: $V\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n VX_i$
- Kovarianz: $\text{Cov}(X, Y) = E(XY) - EX \cdot EY$
 - $V\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n VX_i + \sum_{i \neq j} \text{Cov}(X_i, X_j)$
- Markov-Ungleichung: $P(X \geq a) \leq \frac{EX}{a}$
- Chebychev-Ungleichung: $P(|X - EX| \geq a) \leq \frac{VX}{a^2}$
- SGGZ: X_i unabhängig, EX_i gleich, VX_i gleich: $\lim_{n \rightarrow \infty} P\left(\left|\frac{1}{n} \sum_{i=1}^n X_i - EX_1\right| \geq \varepsilon\right) = 0$

- SGGZ von Bernoulli: X_i unabhängig, $X_i \sim \text{Bin}(1, p)$: $P\left(\left|\frac{1}{n} \sum_{i=1}^n X_i - p\right| \geq \varepsilon\right) \leq \frac{p(1-p)}{n \cdot \varepsilon^2} \leq \frac{1}{4 \cdot n \cdot \varepsilon^2}$
- Poisson-Approximation der Binomialverteilung: $\lim_{n \rightarrow \infty} b_{n, p_n}(k) = p_{np_n}(k)$ für festes k
 - Fehlerabschätzung von Stein-Chen: $X \sim \text{Bin}(n, p) : |P(X \in A) - \sum_{k \in A} p_{np}(k)| \leq (1 - e^{-np}) \cdot p$
- Normalapproximation der Binomialverteilung: $S_n \sim \text{Bin}(n, p), np(1-p) \geq 9 : P(i \leq S_n \leq j) \approx \Phi\left(\frac{j-np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{i-np}{\sqrt{np(1-p)}}\right)$
- ZGS von Lindeberg-Levy: $S_n = \sum_{i=1}^n X_i, X_i$ unabhängig, P^{X_i} gleich: $\lim_{n \rightarrow \infty} P\left(a \leq \frac{S_n - ES_n}{\sqrt{VS_n}} \leq b\right) = \Phi(b) - \Phi(a)$
 - $ES_n = n \cdot EX_1, VS_n = n \cdot VX_1$
 - Fehlerabschätzung von Berry-Essen: $\gamma(X_1) = E(|X_1 - EX_1|^3) : |P\left(\frac{S_n - ES_n}{\sqrt{VS_n}} \leq b\right) - \Phi(b)| \leq \frac{0.8 \cdot \gamma(X_1)}{V(X_1)^{\frac{3}{2}} \cdot \sqrt{n}}$
- Dichtefunktion: $\int_{-\infty}^{\infty} f(x) dx = 1 : P(a \leq X \leq b) = \int_a^b f(x) dx$
- Verteilungsfunktion: $F(t) = P(X \leq t) = \int_{-\infty}^t f(x) dx$
- Erwartungswert: $EX = \int_{-\infty}^{\infty} x \cdot f(x) dx$
 - Transformationsformel: $E(g \circ X) = \int_{-\infty}^{\infty} g(x) \cdot f(x) dx$
- Schätzung der Erfolgswkeit: $\hat{p} = \frac{\sum X_i}{n}$
- erforderliche Stichprobenumfang mit Sicherheitsniveau a , Fehlertoleranz ε : $P_p(-\varepsilon \leq \hat{p} - p \leq \varepsilon) \geq a$ wenn $n = \frac{c^2}{4\varepsilon^2}$ mit c als Lösung zu $\Phi(c) = \frac{a+1}{2}$
- Konfidenzintervall mit Niveau a , Stichprobenumfang n : Für jeden wahren Parameter p enthält $I(\hat{p})$ mit Wkeit $\geq a$ den Parameter p . $I(\hat{p}) = \left[\hat{p} - \frac{c}{2\sqrt{n}}, \hat{p} + \frac{c}{2\sqrt{n}}\right]$
- Chi-Quadrat: $N_k = \#\{i : X_i = k\} : p\text{-Wert} = P\left(X \geq \sum_{k \in \Omega} \frac{(N_k - n \cdot p_k)^2}{n \cdot p_k}\right)$
- Smirnov: Sei $\mathbb{P}$ Dichtefunktion, F Verteilungsfunktion des statistischen Tests, F_n die der Annahme. $p\text{-Wert} = P\left(X \geq \sup_{x \in \mathbb{P}} |F_n(x) - F(x)| \cdot \sqrt{n}\right)$
- Midsquare-Generator: y_0 sei 4-stellig. $y_{n+1} = \lfloor \frac{y_n^2}{100} \rfloor \bmod 10^4$
- Linearer Kongruenzgenerator: $z_{n+1} = a \cdot z_n + r \bmod m$
- Periode: $d(i) = \gcd(\{n \geq 1 \mid P^{n(i,i)} > 0\})$
- Übergangsmatrix ist irreduzibel wenn es einen Rundgang gibt, aperiodisch wenn $d = 1$
- Pfadwkeit: $P(X_k = i_k) = \pi_0(i_0) \prod_{k=0}^{m-1} p_{i_k, i_{k+1}}$
- m -Schritt Übergangswkeit: $P(X_{n+m} = j \mid X_n = i) = P^{m(i,j)}$
- Verteilung zur Zeit: $\pi_n = \pi_0 \cdot P^n$
- invariante Verteilung: $\Pi \cdot P = \Pi$
- Konvergenzsatz: $\lim_{n \rightarrow \infty} P^{n(i,j)} = \Pi(j), \lim_{n \rightarrow \infty} P(X_n = j) = \Pi(j), \lim_{n \rightarrow \infty} \pi_n = \Pi$
- Rückkehrzeit: $T_i = \inf\{n \geq 1 \mid X_n = i\}$
 - $ET_i = \frac{1}{\Pi(i)}$
- Starkes Gesetz: $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} f(x_k) = \sum_{i \in S} f(i) \cdot \Pi(i)$

diskrete Verteilungen

Name	Abkürzung	P	E	V
Gleich		$P(X = x_i) = \frac{1}{n}$	$\frac{1}{n} \sum_{i=1}^n x_i$	
Binomial	$X \sim \text{Bin}(n, p)$	$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$	np	$np(1-p)$
Bernoulli	$X \sim \text{Bin}(1, p)$			
Geometrisch	$X \sim \text{Geo}(p)$	$P(X = k) = (1-p)^{k-1} \cdot p$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
		$P(X > n) = (1-p)^n$		
		$P(X = n + k \mid X > n) = P(X = k)$		
Poisson	$X \sim \text{Poi}(\lambda)$	$P(X = k) = \frac{e^{-\lambda} \cdot \lambda^k}{k!}$	λ	λ
Power-Law	Parameter α, c	$P(X > n) \sim \frac{c}{n^\alpha}$	$\alpha \leq 1 : EX = \infty,$ $\alpha > 1 : E(X - n \mid X > n) \sim \frac{n}{\alpha-1}$	
Zipf	$X \sim \text{Zipf}(\alpha)$, X Power-Law-Verteilt mit α und $c = \frac{\beta}{\alpha}$	$P(X = k) = \frac{\beta}{k^{(1+\alpha)}}$ mit $\beta = \frac{1}{\sum_{k=1}^{\infty} \frac{1}{k^{(1+\alpha)}}}$		
Hypergeometrisch	$X \sim \text{Hyp}(n, r, s)$	$P(X = k) = \frac{\binom{r}{k} \binom{s}{n-k}}{\binom{r+s}{n}}$	$\frac{nr}{r+s}$	

stetige Verteilungen

Name	Abkürzung	Dichte	$P(a \leq X \leq b)$	Verteilung	E	V
Normal	$X \sim N(\mu, \sigma^2)$	$\frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}}$		$\Phi\left(\frac{t-\mu}{\sigma}\right)$	μ	σ^2
Gleich	$X \sim U(r, s)$	$\frac{\mathbb{1}_{[r,s]}(x)}{s-r}$	$\frac{1}{s-r}$	$\min\left(\frac{t-r}{s-r}, 1\right) \cdot \mathbb{1}_{[r,\infty)}(t)$	$\frac{s+r}{2}$	$\frac{(s-r)^2}{12}$
Exponential	$X \sim \text{Exp}(\lambda)$	$\lambda \cdot e^{-\lambda x} \cdot \mathbb{1}_{[0,\infty)}(x)$	$e^{-\lambda a} - e^{-\lambda b}$	$(1 - e^{-\lambda t}) \cdot \mathbb{1}_{[0,\infty)}(t)$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Pareto	$X \sim \text{Par}(\alpha)$	$\frac{\alpha \cdot \mathbb{1}_{[0,\infty)}(x)}{(x+1)^{\alpha+1}}$		$\left(1 - \frac{1}{(t+1)^\alpha}\right) \cdot \mathbb{1}_{[0,\infty)}(t)$	$a \leq 1 : \infty,$ $\alpha > 1 : \frac{1}{\alpha-1}$	