HAW LANDSHUT — SoSe 2023

FAKULTÄT FÜR INFORMATIK — Eisenreich

Formelsammlung zur Klausur Statistik

Verteilung	Wahrscheinlichkeit	E(X)	Var(X)/ Cov(X)
Bernoulli	$\mathrm{Ber}_p(k) = \begin{cases} p & k=1 \\ 1-p & k=0 \end{cases}$	p	$p(1-p)$
Binomial	$b_{n,p}(k) = \binom{n}{k} p^k (1-p)^{n-k}$	np	$np(1-p)$
geometrische	für $X =$ "Anz. Versuche **vor** 1. Treffer": $P(X=k) = p \cdot (1-p)^k$	$\frac{1-p}{p}$	$\frac{1-p}{p^2}$
	für $X =$ "Anz. Versuche **bis incl** 1. Treffer": $P(X=k) = p \cdot (1-p)^{k-1}$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
kategoriale	$\mathrm{Cat}_p(k) = \begin{cases} p_1 & k=1 \\ \vdots & \\ p_K & k=K \end{cases}$		
Multinomial	$P(X_1 = i_1, \ldots, X_s = i_s) = \frac{n!}{i_1! i_2! \cdots i_s!} p_1^{i_1} p_2^{i_2} \cdots p_s^{i_s}$	$n(p_1, \ldots, p_K)$	s.u.
hypergeometrisch	$h_{N,S,n}(k) = \frac{\binom{S}{k}\binom{N-S}{n-k}}{\binom{N}{n}}$	np	$np(1-p)\frac{N-n}{N-1}$
Poisson	$P(X=k) = \frac{\lambda^k}{k!} e^{-\lambda}$	λ	λ

- Näherung $b_{n,p}$ durch Poisson für $\lambda = n \cdot p \leq 10$ und $n \geq 1500 \cdot p$
- Tschebyscheff-Ungleichung: $\forall \varepsilon > 0$ gilt: $P(|X - E(X)| \geq \varepsilon) \leq \frac{1}{\varepsilon^2} Var(X)$
- $Cov(X,Y) = E(XY) - E(X)E(Y) = \frac{1}{2}(Var(X+Y) - Var(X) - Var(Y))$
- zur Multinomialv.: $Cov(X) = \begin{pmatrix} np_1(1-p_1) & -np_1p_2 & \ldots & -np_1p_K \\ -np_2p_1 & np_2(1-p_2) & \ldots & -np_2p_K \\ \vdots & \vdots & \ddots & \vdots \\ -np_Kp_1 & -np_Kp_2 & \ldots & np_K(1-p_K) \end{pmatrix}$

Verteilung	Dichte	Verteilungsfkt.	E(X)	Var(X)/ Cov(X)
Gleichv.	$\begin{cases} 0 & x < a \text{ oder } x > b \\ \frac{1}{b-a} & a \leq x \leq b \end{cases}.$	$\begin{cases} 0 & x < a \\ 1 & x > b \\ \frac{x-a}{b-a} & a \leq x \leq b \end{cases}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Normalv.	$\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$	$\Phi\left(\frac{x-\mu}{\sigma}\right)$, s.u.	μ	σ^2
Exponentialv.	$\begin{cases} \lambda e^{-\lambda x}, & x > 0, \\ 0 & \text{sonst} \end{cases}$	$\begin{cases} 1 - e^{-\lambda x}, & x > 0, \\ 0 & \text{sonst} \end{cases}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Chi-Quadrat-V.	z.B. $g_1(x) = \frac{1}{\sqrt{2\pi x}} \cdot e^{-\frac{x}{2}}$	s.Tabelle	n	$2n$

- zur Normalv.:
$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(\frac{-t^2}{2}\right) dt$$

- Ist X eine $N(\mu, \sigma^2)$-verteilte Zufallsvariable, so gilt für die Abweichungen vom Erwartungswert:

  $P(|X-\mu| \leq \sigma) \approx 0.6826, P(|X-\mu| \leq 2\sigma) \approx 0.9546, P(|X-\mu| \leq 3\sigma) \approx 0.9973$

- Sei X eine $b_{n,p}$-verteilte Zufallsvariable. Dann ist für $np > 5$ und $n(1-p) > 5$ die folgende Rechnung möglich:

  a.) $P(k_1 \leq X \leq k_2) = \Phi\left(\frac{k_2 - np + 0.5}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{k_1 - np - 0.5}{\sqrt{np(1-p)}}\right), 0 \leq k_1 \leq k_2 \leq n$

  b.) $P(X \leq k) = \Phi\left(\frac{k - np + 0.5}{\sqrt{np(1-p)}}\right), P(X < k) = \Phi\left(\frac{k - np - 0.5}{\sqrt{np(1-p)}}\right), 0 \leq k \leq n$

- Varianz einer Stichprobe $x = (x_1, \ldots, x_n)$ mit Mittelwert $\overline{x}$: $s^2 := \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \overline{x})^2$

- Konfidenzintervall Bernoulli-Experiment: $k, n-k > 30$, $\gamma = 1 - \alpha$:

$$p_{u,o} = \frac{k}{n} \pm \frac{c}{n} \sqrt{\frac{k(n-k)}{n}}, \quad \text{mit } c = \Phi^{-1}\left(\frac{1+\gamma}{2}\right), B = 2 \cdot \frac{c}{n} \sqrt{\frac{k(n-k)}{n}}$$

- Bernoulli-Experiment, Konfidenzniveau γ, Konfidenzintervallbreite B: Stichprobengröße $n \geq \frac{c^2}{B^2}$, mit $c = \Phi^{-1}\left(\frac{1+\gamma}{2}\right)$

- n-faches Bernoulliexperiment mit k Treffern, Nullhypothese $P(A) = p_0$, Irrtumswahrscheinlichkeit α:

$$\delta = c \cdot \sqrt{p_0(1-p_0)/n}, \quad c := \Phi^{-1}\left(1 - \frac{\alpha}{2}\right).$$

- n-faches Bernoulliexperiment mit k Treffern, Nullhypothese $P(A) \leq p_0$, Irrtumswahrscheinlichkeit α:

$$\delta = c \cdot \sqrt{\frac{p_0(1-p_0)}{n}}, \quad c := \Phi^{-1}(1-\alpha).$$

- X_k sei b_{n,p_k}-verteilt, dann ist die Pearson'sche Testfunktion:

$$\chi^2 := \sum_{k=1}^{m} \frac{(X_k - np_k)^2}{np_k}$$

- unbegrent lange Warteschlange mit Ankunftsrate λ und Bedienrate μ mit $\lambda < \mu$, $X =$ Anzahl der Kunden im System im stationären Zustand. Dann:

$$P(X=k) = \left(1 - \frac{\lambda}{\mu}\right)\left(\frac{\lambda}{\mu}\right)^k, k \in \mathbb{N}_0, E(X) = \frac{\lambda}{\mu - \lambda}$$

  - $\frac{\lambda}{\mu-\lambda} - \frac{\lambda}{\mu}$: Erwartungswert für die Länge der Warteschlange,
  - $\frac{1}{\mu-\lambda}$: mittlere Aufenthaltsdauer im System,
  - $\frac{1}{\mu-\lambda} - \frac{1}{\mu}$: mittlere Wartezeit in der Schlange.

- In einem Poisson-Prozess $(X_t)_{t\in\mathbb{R}^+}$ ist die Zufallsvariable X_t für alle t Poisson-verteilt mit Parameter λt. Die Wartezeit zwischen zwei Ereignissen ist exponentialverteilt mit Erwartungswert $\frac{1}{\lambda}$

- Der empirische Median (Zentralwert) ist definiert als

$$x_{1/2} := \begin{cases} x_{\left(\frac{n+1}{2}\right)}, & \text{falls } n \text{ eine ungerade Zahl ist} \\ \frac{1}{2}\left(x_{\left(\frac{n}{2}\right)} + x_{\left(\frac{n}{2}+1\right)}\right), & \text{falls } n \text{ eine gerade Zahl ist.} \end{cases}$$

- (empirisches) p-Quantil von $x_1, \ldots, x_n$:

$$x_p := \begin{cases} x_{(\lfloor np+1 \rfloor)}, & \text{falls } np \notin \mathbb{N}, \\ \frac{1}{2}\left(x_{(np)} + x_{(np+1)}\right), & \text{falls } np \in \mathbb{N}. \end{cases}$$

-
  - $|\mathrm{Per}_k^n(mW)| = n^k$,
  - $|\mathrm{Per}_k^n(oW)| = n(n-1)\cdots(n-k+1), \quad k \leq n$
  - $|\mathrm{Kom}_k^n(mW)| = \binom{n+k-1}{k}$,
  - $|\mathrm{Kom}_k^n(oW)| = \binom{n}{k}, \quad n \leq k.$