von JD., Seite 1 von 2 schen Abweichung vom Mittelwert BeschreibendeStatistik 1.4.3 Stichprobenstandardabweichungebnisse eines Experiments 1.1 Begriffe R:sd(x) $s = \sqrt{s}$ Streuungsmaß mit gleicher Einheit wie beobachteten Daten $x_i.\bar{x}$ minimiert 1.1.1 Beschreibende/Deskriptive die "quadratische Verlustfunktionöder Statistik die Varianz gibt das Minimum der Feh-Beobachtete Daten werden durch geeiglerquadrate an. nete statistische Kennzahlen charakteri-1.5 p-Quantile siert und durch geeignete Grafiken an-R:quantile(x, p). Teilt die **sortierten** Da- **Schnitt** $E \cap F$: Ereignis E und Ereignis F

schaulich gemacht. ten x_i ca. im Verhältnis p: (1-p) d.h. $\hat{F}(x_p) \approx p$; 1. Quartil = 0.25-Quantil; Me-1.1.2 Schließende/Induktive Stadian = 0.5-Quantil; 3. Quartil = 0.75-Quartil; 1.6 Interquartilsabstand I Aus beobachtete Daten werden Schlüsse $I = x_{0.75} - x_{0.25}$. Ist ein weiterer Streu- 2.2 De Morgan'schen Regeln gezogen und diese im Rahmen vorgegeungsparameter. bener Modelle der Wahrscheinlichkeits-1.7 Chebyshev theorie bewertet. $\frac{N(S_k)}{n} > 1 - \frac{1}{k^2}$, für alle $k \ge 1 \overline{x}$ der 1.1.3 Grundgesamtheit Durchschnitt, s > 0 die Stichproben-

jekt der Grundgesamtheit diskret(<30 Ausprägungen), stetig(≥30 Ausprägungen), univariat(p=1), mulivariat(p>1) 1.2 Lagemaße 1.2.1 Modalwerte x_{mod} Am häufigsten auftretende Ausprägun-

 Ω : Grundgesamtheit ω :Element oder Ob-

Hilfszettel zur Klausur

gen (insbesondere bei qualitativen Merkmalen) 1.2.2 Mittelwert

Schwerpunkt ten. Empfindlich gegemüber Ausreißern. $\overline{X} = \frac{1}{n} \sum_{i=1}^{n} x_i$

R:mean(x)

1.3 Median R:median(x)Liegt in der Mitt der sortierten Daten x_i . Unempfindlich gegenüber Ausreißern.

 $\frac{x_{n+1}}{2}$, falls n ungerade $\left(\frac{1}{2}(x_{\frac{n}{2}} + x_{\frac{n}{2}+1})\right)$, falls n gerade

1.4 Streuungsmaße

1.4.1 Spannweite

 $\max x_i$ - $\min x_i$

1.4.2 Stichprbenverians s^2 R:var(x)Verschiebungssatz:

 $\frac{1}{n-1}(\sum_{i=1}^{n}(x_iy_i-n\overline{xy}))$

1.8 Korrelation

R:cor(x, x); $r = \frac{s_{xy}}{s_x x_y}$; Näherungsweise lin. Zusammenhang zw. x und y, falls $|r| \approx 1$.

Standardabweichung von Beobachtungs-

werten $x_1,...,x_n$. Sei $S_k = \{i, 1 \le i \le n : |x_i - \overline{x}| < k \cdot s\}$; Für eine beliebige Zahl

 $k \ge 1$ liegen mehr als $100 \cdot (1 - \frac{1}{L^2})$ Pro-

zent der Daten im Intervall von $\bar{x} - ks$ bis

 $\overline{x} + ks$. **Speziell:**Für k = 2 liegen mehr als

k=3 liegen mehr als 89% der Daten im

rung: $\overline{S}_k = \{i | |x_i - \overline{x}| \ge k \cdot s\}; \frac{N(S_k)}{n} \le \frac{1}{k^2};$

Die Ungleichheit lifert nur eine sehr gro-

be Abschätzung, ist aber unabhängig

von der Verteilung der Daten. Empiri-

sche Regeln 68% der Daten im Bereich

um $\overline{x} \pm s$. 95% um $\overline{x} \pm 2s$. 99.7% um $\overline{x} \pm 3s$.

Grafische Zusammenhang zwischen mul-

tivariaten Daten y und y durch ein Streudiagramm. Kennzahlen zur Unter-

R:cov(x, y); $s_{xy} = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \overline{x})(y_i - \overline{y}) =$

1.8.2 Empirische Korrellationsko-

suchung des Zusammenhangs:

1.8.1 Empirische Kovarians

75% der Daten im 2s-Bereich um \bar{x} . Für

3s-Bereich um \overline{x} . **Komplement Formulie**-

1.8.3 Regressionsgerade y $s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \overline{x}^2) = \frac{1}{n-1} (\sum_{i=1}^{n} x_i^2 - y = mx + t \text{ mit } m = r \cdot \frac{s_y}{s_x} \text{ und } t = \overline{y} - m \cdot \overline{x}$

 $n\bar{x}^2$) Gemittelte Summe der quadrati- 2 Wahrscheinlichkeitsrechnung 2.1 Begriffe **Ergebnisraum** Ω : Menge aller möglichen

Elementarereignis $\omega \in \Omega$: einzelnes Ele-

Ereignis $E \subseteq \Omega$: beliebige Teilmenge des

Ergebnisraums Ω heißt sicheres Ereignis,

Vereinigung $E \cup F$: Ereignis E oder Ereig-

nis F treten ein. $\bigcup_{i=1}^{n} E_i$: mindestens ein

 $\bigcap_{i=1}^n E_i$ alle Ereignisse E_i treten ein. **Ge**-

genereignis $\overline{E} = \Omega / E$: Ereignis E tritt

Disjunkte EreignisseE und F: $E \cap F = \emptyset$

 $P(\bigcup_{i=1}^{\infty}) = \sum_{i=1}^{\infty} P(E_i)$, falls $E_i \cap E_j = \emptyset$

 $P(E \cup F) = P(E) + P(F) - P(E \cap F)$

(Übungsaufgabe!!! Ergänzen)

2.4 Laplace-Experiment

keit P(E) für $E \subseteq \Omega$ aus:

Ø heißt unmögliches Ereignis

nicht ein (Komplement von E)

ment von Ω

Ereignis E_i tritt ein.

 $E_1 \cup E_2 = E_1 \cap E_2$

 $\overline{E_1 \cap E_2} = \overline{E}_1 \cup \overline{E}_2$

2.3.1 Satz 2.1

 $P(\overline{E}) = 1 - P(E)$

2.3 Wahrscheinlichkeit

 $0 \le P(E) \le 1$; $P(\Omega) = 1$;

2.5.2 Permutationen Anzahl einer n-elementigen Menge n**bare Elemente**: $n! = n \cdot (n-1) t ext b f ... 2 \cdot 1$

maliges Ziehen ohne Zurücklgen mit Beachtung der Reihenfolge: n unterscheidk Klassen mit je n_i nicht unterscheidbaren Elementen $n = sum_k^{i=1} n_i$: $\frac{n!}{n_1! \cdot n_2! \cdot n_k!}$

 \overline{E}) = $P(E) - P(\overline{F} \cap E)$; $P(\overline{F}|E) = 1 - P(F|E)$

E E P(TAE) P(TAE) P(T) $P(F|E) = P(F) \cdot P(E|F)$ **Tafel** $= P(F) - P(F \cap F)$

2.6.3 Vierfeldertafel $P(F) = P(F \cap E) + P(F \cap \overline{E})$ $P(\overline{F}) = P(\overline{F} \cap E) + P(\overline{F} \cap E)$

 $P(E \cap F) = P(E|F) \cdot P(F)$ $P(E \cap F) = P(F|E) \cdot P(E)$

2.5.3 Anzahl k-elementigen Teil-

ohne Zurücklegen = $k \le n$.

Zurücklegen: $\binom{n}{k} = \frac{n!}{(n-k)!k!}$

Zurücklegen: $\frac{n!}{(n-k)!}$

Zurücklegen $\binom{n+k-1}{k}$

2.6.1 Satz 2.2

rücklegen: nk

mit Zurücklegen = k > n möglich.

mit Beachtung der Reihenfolge, ohne

ohne Beachtung der Reihenfolge, ohne

mit Beachtung der Reihenfolge, mit Zu-

ohne Beachtung der Reihenfolge, mit

2.6 Bedingte Wahrscheinlichkeit

 $P(E|F) = P_F(E) = \frac{|E \cap F|}{|F|} = \frac{P(E \cap F))}{P(F)}$

mengen einer n-elementigen

Menge k-maliges Ziehen aus

einer n-elementigen Menge

2.6.2 Satz der totalen Wahrschein-Sei $\Omega = \bigcup_{i=1}^n E_i$ mit $E_i \cap E_j = \emptyset$ für $i \neq j$

Summe der Äste des Wahrscheinlichkeitsbaums zu allen Schnitten $F \cap E_i$

Zufallsexperimente mit n gleich wahr-Elementarereignissen. d.h. die Ereignisse bilde eine disjunkte Dann berechnet sich die Wahrscheinlich-Zerlegung bzw. eine Partition von Ω . Somit gilt: $P(F) = \sum_{i=1}^{n} P(F \cap E_i) = \sum_{i=1}^{n} P(F|E_i)$

 $P(E) = \frac{AnzahlderfrEgnstigenEreignisse}{AnzahldermglichenEreignisse}$ $\frac{\textit{MchtigkeitvonE}}{\textit{Mchtigkeitvon}\Omega} = \frac{|E|}{\Omega} \textbf{text}$

2.5 Kombinatorik 2.5.1 Allgmeines Zählprinzip

Anzahl der Möglihckeiten für ein kstufiges Zufallsexperiment mit n_i Varianten im i-ten Schritt: $n_1 \cdot n_2 \cdot ... \cdot n_k$

Satz 2.2 $P(E \cap F)P(E)$

nicht ändert, d.h. falls $P(E|F) = P(E)oderP(E \cap F) = P(E) \cdot P(F)$ $= \frac{P(E \cap F)}{P(F)}$

Es gilt Falls die Ereignisse E, F unabhängig sind, dann sind auch:

 \overline{E} , \overline{F} unabhängig **Bemerkung** Stochastische Unabhängigkeit bedeutet nicht notwendigerweise ei-

2.6.4 Formel von Bayes

 $P(F|E_k) \cdot P(E_k)$

 $P(F|E_i) \cdot P(E_i)$

len Wahrscheinlichkeit.

Hilfreich, wenn man man $P(F|E_i)$ kennt,

aber nicht $P(E_k|F)$ Satz 2.4 $P(E_k|F) =$

Nur Nenner!P(F) aus dem Satz der tota-

2.6.5 Stochastische Unabhängig-

Uebung Die Ereignisse E und F heißen

(stochastisch) unabhängig, wenn die

Information über das Eintreten des einen

Ereignisses die Wahrscheinlichkeit für

das Eintreten des anderen Ereignisses

ne kausale Abhängigkeit Veranschaulichung mit Venn Dia-

 $P(E) = \frac{4}{2} = P(E(F))$ gramm stock unabhanging P(E)= 1 < P(E| F)

• $A, B \neq \emptyset$ und $A \cap B = \emptyset$ $P(A \cap B) \stackrel{?}{=} P(A) \cdot P(B)$ $\emptyset \neq P(A) \cdot P(B)$ da P(A) > 0 und P(B) > 0=> A, B stochastisch abhängig

3 Zufallsvariable Abbildung des **abstrakte** Ergebnisraums Ω auf \mathbb{R} . Eine Abbildung $X: \Omega \to \mathbb{R}$, $\omega \mapsto X(\omega)$ = heißt Zufallsvariable (ZV). x

∈ R. heißt Realisation der ZV X. • Diskrete ZV: $X(\Omega) = x_1, ..., x_2 (n \in$

 \mathbb{N}); z.B. X = "Augensumme beim"Würfeln

• Stetige ZV: $X(\Omega) \subseteq \mathbb{R}$; "z.B. Körpergröße eines Menschen"

3.1 Verteilungsfunktion-allg.

Die Wahrscheinlichkeit P(B) für ein Ereignis B in R wird zurückgefürht auf die Währscheinlichkeit der entsprechenden

Ereignisse in Ω . Für jedes $X \in \mathbb{R}$ ist die

Verteilungsfunktion F: $\mathbb{R} \to [0,1]$ einer ZV X definiert durch: $F(x) = P(X \le x)$

- 0 < F(x) < 1• $\lim_{x \to -\infty} F(X) = 0 \lim_{x \to \infty} F(x) = 1$
- · monoton wachsend • P(X > x) = 1 - F(x)

Hilfszettel zur Klausur von JD., Seite 2 von 2

- $P(a < X \le b) = F(b) F(a)$
- 3.2 Diskrete ZVs Für eine diskrete ZV X mit $X(\Omega) =$ $x_1,...,x_n$ (n endlich oder abzählbar unendlich) ist die Wahrscheinlichkeitsfunktion definiert durch:
- $p(x) = \begin{cases} P(X = x_i), \text{falls } x_i \in X(\Omega) \\ 0, sonst \end{cases}$ Es gilt:

•
$$F(x) = (P(X \le x) = \sum_{x_i \le x} p(x_i)$$

- F(x) ist eine rechtseitig stetige Treppenfunktion mit Sprüngen
- bei der Realisation von x_i . 3.3 Stegite ZVs Stetige ZV X ist die Wahrscheinlichkeitsdichte f $f: \mathbb{R} \to [0, \infty]$ definiert durch
- $P(a < X < b) = \int_{a}^{b} f(x) dx$

 $X \leq b$

- $F(x) = P(X \le x) = \int_{-\infty}^{x} f(t)dt$ und F'(x) = f(x)
- F(x) ist stetig & $P(a < X \le b) =$ $P(a \le X \le b)$ wegen P(X = a) = 0

3.4 Verteilungsfunktion

- Untergrenze Es wird normal mit Inte-
- 3.5 Zusammenfassung
- 3.5.1 Diskrete ZV Wahrscheinlichkeitsverteilung $p(x) \sum_{i=1}^{n} p(x_i) = 1x_i$ ist Realisation
 - der ZV. • Verteilungsfunktion F(x) ist rechtsseitig stetige Treppenfunktion. **Sprunghöhen:** $P(X = x_i) = F(x_i) \lim \neq 0$
 - $P(a < X \le b) = F(b) F(a) \ne P(a \le b)$

 $Var[X] = E[X^2] - (E[X])^2$ Beim Minuend wird beim Erwartungswert nur das einfach stehende x quadriert nicht f(x)!!!

 $Z = \frac{X - \mu}{5} = \frac{x}{5} - \frac{\mu(konstant)}{5}$ • Verteilungsfunktion F(x) ist stetig 3.9 Kovarianz mit F'(x) = f(x); $P(X = x_i) = 0$ Eigenschaften: • $P(a < X \le b) = F(b) - F(a) = P(a \le b)$

Cov[X,Y]=0

3.10 Satz 3.3

3.11.1 E[X]

den gilt:

- $X \le b$) = $F(a \le X < b)$ = P(a < X < b)Der Erwartungswert E[X] = einer ZV
- X ist der **Schwerpunkt** ihrer Verteilung or der durchschnittliche zu erwartende Wert der ZV. • diskrete ZV: $E[X] = \sum_{i=1}^{n} x_i \cdot p(x_i)$

• Dichtefunktion fx $\int_{-\infty}^{\infty} f(x)dx = 1$

ZV ist konstant. E[X] verhält sich linear. Eigenschaften von E[X]: • E[b] = b

• stetige ZV: $E[X] = \int_{-\infty}^{\infty} x \cdot f(x) dx$

• $E[X_i + ... + X_n] = \sum_{i=1}^n E[X_i]$

• E[aX + b] = aE[X] + b

• $\sum_{i=1}^{n} x_i$

3.5.2 Stetige ZV

3.6 Erwartungswert

3.6.1 Satz 3.1

- Sei Y = g(X) eine Funktion der ZV X. Dann gilt: • für diskrete ZV:E[g(X)] =
 - $\sum_{i=1}^{n} g(x) \cdot p(x_i)$ $\sum_{i=1}^{n} E[X_i];$ • für stetige ZV: E[g(X)] = $\int_{-\infty}^{\inf fty} g(x) \cdot f(x) dx.$ Das vertauschen von E und g nur bei linearen Funktionen möglich. \Rightarrow g(E[X])
- 3.7 Varianz Die Varianz einer ZV X mit µ ist ein qua-
- dratisches Streungsmaß. $\sigma^2 = Var[X] =$ $E[(X-)^2]$ falls $\underset{=}{\text{stetig}} \int_{-\infty}^{\infty} (x-\mu)^2 \cdot f(x)$
- g(X)Die Standardabweichung $\sigma = \sqrt{Var[X]}$ hat im Gegensatz zur Varianz die gleiche
 - Var[b] = 0
 - $Var[aX + b] = a^2 Var[X]$

Dimension von die ZV X.

3.7.1 Satz 3.2

4.1 Diskrete Verteilung

 $p-p^2=p(1-p);$

≜Wahrscheinlichkeits-

≜Verteilungsfunktion;

lung

qbinom(q,n,p)=q-Quantil;

4.1.1 Bernouilliverteilung

Indikatorvariable mit den Werten 1 bei

Erfolg und 0 bei Misserfolg; Wahrschein-

lichkeit:P(X = 1) = p, P(X = 0) = 1 - p;**Verteilung:** $X \sim B_{1,p}$ p ist Erfolgswahr-

scheinlichkeit; $E[X] = p = \sum x_i \cdot p(x_i) = 1$.

 $p(1); Var[X] = p(1-p) = E[X^2] - (E[X])^2 =$

np(1 - p); **R:** dbinom(k,n,p)=P(X=k)

/Dichtefunktion; pbinom(k,n,p)=F(k)

3.8 Z-Transformation, Standardisie- 4 Spezielle Verteilung

• Cov[X, Y] = Cov[Y, X]

• Cov[X, X] = Var[X]• Cov[aX, Y] = aCov[X, Y]Die Kovarianz zweier ZV (X, Y) definiert durch Cov[X, Y]

Sei X eine ZV mit μ und σ . Dann ist

3.10.1 Varianz einer Summe von

E[(X - E[X])(Y - E[Y]) Die Kovarianz beschreibt die Abhängigkeit zweier ZV X und Y. Je stärker diese Korrelieren, desto Anzahl der Erfolge beim n-1 Anzahl der Erfolge beim n-maligen Ziehen**mit** Zurücklegen; (betragsmäßig) größer ist die Kovarianz. Falls X, Y stochastisch unabhängig \Rightarrow scheinlichkeit $P(x = k) = \binom{n}{k} \cdot p^k$ $(1 - p)^{n-k}, k \in [0, 1, ..., n]$; Verteilung $X \sim B_{n,p}$; E[X] = np; Var[X] = $Cov[X, Y] = E[XY] - E[X] \cdot E[Y]$

• $Var[X_i + ... + X_n] = \sum_{i=1}^{n} \sum_{j=1}^{n} Cov[X_i, X_j]; Var[X_1 +$ rbinom(k,n,p)\(\hat{p}\) kbinomialverteilte Zufallszahlen; X_2] = $Var[X_1] + Var[X_2] + 2Cov[X_1, X_2]$ 4.1.3 Hypergeometrische Vertei-

• Falls X_i, X_i paarweise unabhängig !!!: $Var[X_1 + ... + X_n] = \sum_{i=1}^n Var[X_i]$ Anzahl der Erfolge beim n-maligen Ziehen ohne Zurücklegen aus einer

3.11 Overview $\mu \sigma$ $E[aX + b] = AE[X] + b; EX_1 + ... + E_n =$

Falls X_1, X_2 unabhängig: $E[X_i] = \mu => E[\overline{X}] = E[\frac{1}{n}(X_1 + ... + X_n)] =$ $\frac{1}{n}\sum_{i=1}^{n} E[x_i] = \frac{1}{n} \cdot n \cdot \mu = \mu$

3.11.2 Varianz $Var[aX + b] = a^{2}Var[X]$ Falls X_{i} , X_{j} parweise unabhängig:

 $Va[X_1 + ... + X_n] = \sum_{i=1}^n Var[X_i]$ $Var[X_i] = \sigma^2 \Longrightarrow Var[\overline{X}] = Var[\frac{1}{n}(x_1 + ... +$ $[x_n] = \frac{1}{n^2} \sum_{i=1}^n Var[X_i] = \frac{1}{n^2} \cdot n \cdot \sigma^2 = \frac{\sigma^2}{n}$ 3.12 Ouantile Sei X eine ZV mit Verteilungsfunktion F(x) und 0 . Dann ist das p-Quantil definiert als der Wert $x_p \in \mathbb{R}$ für

 $F(x_p) \ge p$. p-Quantil einer stetigen ZV mit streng monoton wachsenden $F(x)x_p = F^{-1}(p)d$. h. umkehrbar.

ppois $(k, \lambda) = F(k)$;

4.1.5 Gleichverteilung

Alle Werte $\{x_1,...,x_n\}$ einer ZV X sind gleich wahrscheinlich; **Wahrscheinlich**keit $P(X = x_k) = \frac{1}{n}$; Verteilung

 $Var[X] = \lambda \mathbf{R} : \frac{d}{pois}(k, \lambda) = P(X = k);$

 $X \sim U_{\{x_1,...,x_n\}}; E[X] = \frac{1}{n} \sum_{k=1}^n x_k = \overline{x};$ $Var[X] = \frac{1}{n} \sum_{k=1}^{n} x_k^2 - \overline{x}^2$; **R:** sample(1: N,n) $\hat{=}$ n Zufallszahlen zwischen 1 und

4.2 Gleichverteilung 4.2.1 Stetige Gleichverteilung

Zufallszahlen aus einem Intervall [a, b]; **Dichte:** $f(x) = \frac{1}{h-a}$ für $x \in [a,b]$;

Verteilung: $X \sim U_{[a,b]}$; $E[X] = \frac{a+b}{2}$;

 $Var[X] = \frac{(b-a)^2}{12} \mathbf{R} : \frac{d}{duni} f(x, a, b) = f(x);$ punif(x, a, b) = F(x); runif(n) = n Zufallszahlen zwischen 0 und 1; runi f(n, a, b) $\hat{=}$ n Zufallszahlen zwischen a und b;

4.2.2 Normalverteilung

Beschreibt viele reale Situationen, ist insbesondere Grenzverteilung

unabhängiger Summen;

 $f(x) = \frac{1}{\sigma\sqrt{2\pi}} \left(-\frac{1}{2} \left(\frac{x-\mu}{\sigma}\right)^2\right)$; Verteilung:

 $X \sim N_{\mu,\sigma^2}$; $E[X] = \mu$; $Var[X] = \sigma^2$; **R**:

Menge mit M Elementen, die Erfolg be $dnorm(x, \mu, \sigma) = f(x); pnorm(x, \mu, \sigma) =$ deuten, und N Elementen, die Misserfolg bedeuten. Gesamtum fang = M + N; F(x); qnorm (q, μ, σ) : q - Quantil; **Maxi**-Wahrscheinlichkeit $P(\bar{X} = k) =$ malstelle von f(x) bei $x = \mu$; Wende**stelle** von f(x) bei $x = \mu \pm \sigma$; E[aX + b] =

 $\frac{\binom{M}{k}\cdot\binom{N}{n-k}}{\binom{M+N}{k}}, k \in \{0,1,...,min\{n,M\}\}; \text{ Ver-}$ aE[X] + b; $Var[aX + b] = a^2 Var[X]$; $X \sim N_{\mu,\sigma^2} \Rightarrow aX + b \sim N_{a\mu+b,a^2\sigma^2}$ und teilung $X \sim H_{M,N,n}$; $E[X] = n \frac{M}{M+N}$; $\frac{X-\mu}{\sigma}$ ~ $N_{0,1}$; X_1 ~ N_{μ_1,σ_1^2} und X_2 ~ $\frac{M}{M+N}$ $\hat{=}$ Tref f erwahrscheinlichkeit;

 $Var[X] = n \frac{M}{M+N} (1 - \frac{M}{M+N}) \frac{M+N-n}{M+N-1};$ $\rightarrow 1$ falls n klein im Verhältnis zu $N_{\mu_2,\sigma_2^2} \Rightarrow X_1 + X_2 \sim N_{\mu_1 + \mu_2,\sigma_1^2 + \sigma_2^2};$ X_1, X_2 stochastisch unabhängig M+N; **R**: dhyper(k, M, N, n) = P(X = k);

4.1.4 Poisson-Verteilung

4.2.3 Standardnormalverteilung

Dichte: $\varphi(x) = \frac{1}{\sqrt{2}}e^{(-\frac{1}{2}x^2)}$; Verteilung

 $X \sim P_{\lambda}$; $E[X] = \lambda, da \sum_{k=0}^{\infty} k \frac{\lambda^{k}}{k!} e^{-\lambda} =$

 $\mathbf{keit}P(X=k) = \frac{\lambda^k}{k!}e^{-\lambda} \text{ mit } \sum_{k=0}^{\infty} P(X=k)$ k) = 1, $da \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{\lambda}$; Verteilung

 $\phi(x) = \int_{-\infty}^{x} \varphi(t)dt$; Quantile: $\phi(-x) = 1$ $\phi(x) \Rightarrow -x_p = x_{1-p} \text{ z.B. } -x_{0.25} = x_{0.75};$

nem Kontinuum. Die durchschnittlich

phyper(k, M, N, n) = F(k);

zu erwartende Anzahl der Erfolge λ pro Maßeinheit (i. a. Zeiteinheit) sei bekannt. $k \in \mathbb{N}_0 \rightarrow diskret$ Wahrscheinlich-

Verteilung der seltenen Ereignisse Häu-

figkeit punktförmiger Ereignisse in ei-

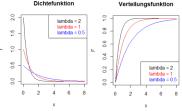
 $e^{-\lambda} \sum_{k=1}^{\infty} \lambda \frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda} \sum_{i=0}^{\infty} \frac{\lambda^{i}}{i!} = \lambda;$ werte: $Z = \frac{x-\mu}{\sigma} \sim N_{0,1}$

Hilfszettel zur Klausur von **JD**., Seite 3 von 2

4.2.4 Exponentialverteilung

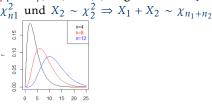
Modellierung von Lebensdauern, Wartezeiten Sei $Y_t \sim P_{\lambda t}$ im Intervall [0,t]von t Zeiteinheiten, dann beschreibt die Exponentialverteilung die Wartezeit X bis zum Eintreten eines Ereignisses; Dichte- und Verteilungsfunktion: $f(x) = \lambda e^{-\lambda x} (x \ge 0)$ und F(x) = 1 $e^{-\lambda x}$; Verteilung: $X \sim Exp_{\lambda}$; E[X] = $\frac{1}{1} \Rightarrow$ Berechnung mit partieller Integra-

tion; $Var[X] = \frac{1}{12}$; **R**: $dexp(x, \lambda) = f(x)$; $pexp(x, \lambda) = F(x)$; **Eigenschaft:** Eine exponentialverteile ZV X ist gedächtnislos, d.h. P(X > s + t)|X > t = P(X > s);



4.2.5 Chiquadrat-Verteilung

 $Z_1,...,Z_n$ seien unabhängige, standardnormalverteilte ZV \Rightarrow X = $Z_1^2 + + Z_n^2$ hat Chiquadratverteilung mit n Freiheitsgraden; Anwendungsmodell: Summen unabhängiger, standardnormalverteilter ZV; **Verteilung:** $X \sim \chi_n^2$; E[X] =n; Var[X] = 2n; **R**: $\frac{d}{d}chisq(x,n) = f(x)$; ppchisq(x,n) = F(x); Eigenschaft: $X_1 \sim$



4.2.6 t-Verteilung

 $Z \sim N_{0,1}$ und $X \sim \chi_n^2 \Rightarrow Y =$ $\frac{Z}{X}$ ist t-verteilt mit n Freiheitsgra-

den; Anwendungsmodell: Schätz- und Testverfahren bei unbekannter Varianz; Verteilung: $Y \sim t_n$; E[Y] = 0 für n > 1; $Var[Y] = \frac{n}{n-2}$ für n > 2; R: dt(y,n) = f(x); pt(y,n) = F(x); Eigenschaf**ten:** Für $n \to \infty$: $t_n \to N_{0,1}$; Achsensym--qnorm(p) =metrie der Dichtefunktion $\Rightarrow -y_p = x_{1-p}$

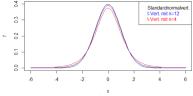


Abbildung Dichtefunktion

5 Zentraler Grenzwertsatz

 $\mu\sigma^2$ bekannt aber nicht die Verteilung

Seien X_i (i = 1,...,n) unabhängige identische verteilte (i.i.d) ZV mit Erwartungswert μ und Varianz σ^2 . Dann gilt für hinreichend große n (>30) und $\overline{X} = \frac{1}{n} \sum_{i=1}^{n}$ näherungsweise:

$$\sum_{i=1}^{n} X_i \sim N_{n\mu,n\sigma^2} \&$$

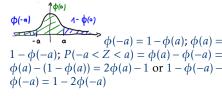
$$\frac{\sum X_i - n\mu}{\sqrt{n \cdot \sigma}} \sim N_{0,1}$$

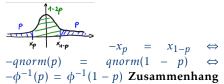
$$\sum X_i \text{ bezieht sich auf Y; } \sum$$

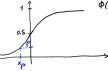
 $\sum X_i$ bezieht sich auf Y; $\sum X_i - n\mu$ bezieht sich auf X_i ; $\overline{X} \sim N_{\mu,\frac{\sigma^2}{n}} & \overline{X} - \mu \sim N_{0,1}$;

Der Satz gilt sogar allgemeiner, wenn die X_i abhängig und nicht identisch verteilt sind, vorausgesetzt kein X_i ist deutlich dominanter?! als die anderen.Für die Voraussetzung des ZGW ist, dass die X_i nicht normalverteilt sein müssen., damit $\sum_{i=1}^{n} X_i$ oder \overline{X} bei **hinreichend** großem n normalverteilt sind. Faustregel: Je schiefer die Verteilung der X_i desto größer muss n sein: n>30: falls die unbekannte Verteilung ohne markanten Ausreißer, aber schief ist (Exponentialverteilung); n>15: falls die unbekannte Verteilung annähernd symmetrisch ist(Binomialverteilung); $n \le 15$: falls die unbekannte Verteilung annähernd nor- 5.4.2 Stichprobenvarianz malverteilt ist;

5.2 ϕ







Aufgabentypen: Seien X_i i.i.d. ZV mit

 μ und σ^2 , aber unbekannter Verteilung. Dann sind $Z_1 = \frac{\sum X_i - n\mu}{\sqrt{n}\sigma}$ und $Z_2 = \frac{\overline{X} - \mu}{\sigma}$

näherungsweise standardnormalverteilt.

- Es lassen sich Wahrscheinlichkeiten für $\sum X_i, X, Z_1$ oder Z_2 berech-
- Es lässt sich n bestimmen, so dass, zu vorgegebener Schranke k und Wahrscheinlichkeit p gilt: $P(Z_i >$ $k \ge p$ or $P(-k \le Z_i \le k) \ge p$

5.4 Stichprobenverteilungen für normalverteilte Grundgesamtheiten

5.4.1 Stichprobenmittel

Die Stichprobenfunktion $\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$ ist eine erwartungstreue Schätzfunktion für Erwartungswert μ , d. h. $E[X] = \mu$

Stichprobenfunktion $S^2 =$ $\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2 = \frac{1}{n-1}(\sum_{i=1}^{n}X_i^2$ $n\overline{X}^2$)ist eine erwartungstreue Schätzfunktion für die Varianz σ^2 , d. h. $E[S^2] = \sigma^2$; $E[\overline{X}] = E[\frac{1}{n}\sum X_i] =$ $\frac{1}{n}E[\sum X_i] = \frac{1}{n}\sum_{i=1}^n E[X_i] = \frac{1}{n}n\mu = \mu;$ $Var[\overline{X}] = Var[\frac{1}{n}\sum X_i] = \frac{1}{n^2}Var[\sum X_i] =$ $\frac{1}{n^2}n\sigma^2 = \frac{\sigma^2}{n}$; Seien $X_i(i = 1,...,n)$ unabhängige normalverteilte ZV mit Erwartungswert μ und Varianz σ^2 . Dann gilt: bei unbekannter Varianz: $\frac{X-\mu}{\sigma}\sqrt{n} \sim$ $N_{0,1}$; $\frac{(n-1)S^2 = \sum (x-\overline{x})^2}{\sigma^2 \Rightarrow \text{Standardisierung}} \sim \chi_{n-1}^2$; **Bei** unbekannter Varinanz: $\frac{X-\mu}{S}\sqrt{n} \sim t_{n-1}$;