1.10 p-Quantile Hilfszettel zur Klausur von JD., Seite 1 von 4 BeschreibendeStatistik

## Beobachtete Daten werden durch geeignete statistische Kennzahlen charakteri-

siert und durch geeignete Grafiken anschaulich gemacht. 1.2 Schließende/Induktive Statistik Aus beobachtete Daten werden Schlüsse gezogen und diese im Rahmen vorgege-

#### bener Modelle der Wahrscheinlichkeitstheorie bewertet. 1.3 Grundgesamtheit

 $\Omega$ : Grundgesamtheit  $\omega$ :Element oder Objekt der Grundgesamtheit diskret(<30 Ausprägungen), stetig(≥30 Ausprägungen), univariat(p=1), mulivariat(p>1); Diskrete Merkmale haben eine abzählbare Anzahl möglicher Ausprägungen. Stetige Merkmale habne eine nicht abzähl-

### Ausprägungen. Lagemaße **1.4** Modalwerte $x_{mod}$ Am häufigsten auftretende Ausprägungen (insbesondere bei qualitativen Merk-

bare (=überabzählbar) Anzahl möglicher

### 1.5 Mittelwert, quantitativ R:mean(x)Schwerpunkt

ten.**Empfindlich**gegenüber Ausreißern.

# $\overline{X} = \frac{1}{n} \sum_{i=1}^{n} x_i$ 1.6 Median, quantitativ R:median(x)

malen)

Liegt in der Mitt der sortierten Daten  $x_i$ . Unempfindlich gegenüber Ausreißern.

 $x_{0.5} = \begin{cases} x_{\frac{n+1}{2}}, \text{ falls n ungerade} \\ \frac{1}{2}(x_{\frac{n}{2}} + x_{\frac{n}{2}+1}), \text{ falls n gerade} \end{cases}$ Streuungsmaße

# 1.7 Spannweite $\max x_i$ - $\min x_i$

1.8 Stichprobenvarianz s<sup>2</sup> R:var(x)

Verschiebungssatz:  $s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \overline{x}^2) = \frac{1}{n-1} (\sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} x_i^2)$  $n\bar{x}^2$ ) Gemittelte Summe der quadrati-

schen Abweichung vom Mittelwert 1.9 Stichpr.standardabw. R:sd(x)

lerquadrate an.

 $s = \sqrt{s}$  Streuungsmaß mit gleicher Einheit wie beobachteten Daten  $x_i.\overline{x}$  minimiert die "quadratische Verlustfunktionöder die Varianz gibt das Minimum der Feh-

#### R:quantile(x, p). Teilt die **sortierten** Daten $x_i$ ca. im Verhältnis p: (1-p) d.h. $\hat{F}(x_p) \approx p$ ; $\hat{F} = \text{kummul. rel. Häufigkeit}$ ; 1.1 Beschreibende/Deskriptive Statis-1. Quartil = 0.25-Quantil; Median = 0.5-

Quantil; 3. Quartil = 0.75-Quartil;

 $x_p \begin{cases} x_{floor(np)+1}, np \in \mathbb{N} \\ \frac{1}{2}(x_{np} + x_{np+1}, np \notin \mathbb{N}) \end{cases}$ 1.11 Interquartilsabstand I

 $I = x_{0.75} - x_{0.25}$ . Ist ein weiterer Streu-

## ungsparameter. 1.12 Chebyshev $\frac{N(S_k)}{n} > 1 - \frac{1}{k^2}$ , für alle $k \ge 1 \overline{x}$ der $\overline{E_1 \cup E_2} = \overline{E_1} \cap \overline{E_2}$

Durchschnitt, s > 0 die Stichproben-  $\overline{E_1 \cap E_2} = \overline{E_1} \cup \overline{E_2}$ Standardabweichung von Beobachtungswerten  $x_1,...,x_n$ . 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  $100 \cdot (1 - \frac{1}{k^2})$  Prozent der Daten im Intervall von  $\bar{x} - ks$  bis  $\overline{x} + ks$ . **Speziell:**Für k = 2 liegen mehr als

75% der Daten im 2s-Bereich um  $\bar{x}$ . Für k=3 liegen mehr als 89% der Daten im 3s-Bereich um  $\bar{x}$ . Komplement Formulierung:  $\overline{S}_k = \{i | |x_i - \overline{x}| \ge k \cdot s\}; \frac{N(S_k)}{n} \le \frac{1}{k^2};$ 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  $\overline{x} \pm 3s$ . 1.13 Korrelation Grafische Zusammenhang zwischen multivariaten Daten x und y durch ein Streudiagramm. Kennzahlen zur Untersuchung des Zusammenhangs: 1.14 Empirische Kovarianz R:cov(x,y);  $s_{xy} = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \overline{x})(y_i - \overline{y}) =$ 

# $\frac{1}{n-1}(\sum_{i=1}^{n}(x_iy_i)-n\overline{xy}); S_{xy}>0 \text{ steigend};$ $\ddot{S}_{xv} < 0$ fallend;

1.15 Empir. Korrelk.koeff. r

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

Bemerkung: -Der Korrelationskoeffizient kann nur einen statistischen Zusammenhang beschreiben, keinen Kausalen; -Den Korrelationskoeffizient immer im Zusammenhang mit den Streudiagramm sehen (Anscombe-Quartett).

## y = mx + t mit $m = r \cdot \frac{s_y}{c}$ und $t = \overline{y} - m \cdot \overline{x}$ ; Für den Bereich $|\pm 0.7|$ bis $\pm 1 \Rightarrow$ linea-

rer Zusammenhang. 2 Wahrscheinlichkeitsrechnung

**Ergebnisraum**  $\Omega$ : Menge aller möglichen Ergebnisse eines Experiments

**Ereignis** $E \subseteq \Omega$ : beliebige Teilmenge des Ergebnisraums  $\Omega$  heißt sicheres Ereignis, Ø heißt unmögliches Ereignis

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

ment von  $\Omega$ 

2.4 Satz 2.1

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

Ereignis  $E_i$ tritt ein.

nicht ein (Komplement von E)

**Vereinigung**  $E \cup F$ : Ereignis E oder Ereignis F treten ein.  $\bigcup_{i=1}^{n} E_i$ : mindestens ein **Schnitt**  $E \cap F$ : Ereignis E und Ereignis F  $\bigcap_{i=1}^n E_i$  alle Ereignisse  $E_i$  treten ein. **Gegenereignis**  $\overline{E} = \Omega / E$ : Ereignis E tritt

**Disjunkte Ereignisse**E und F:  $E \cap F = \emptyset$ 1 - P(F|E)

2.2 De Morgan'schen Regeln 2.3 Wahrscheinlichkeit  $0 \le P(E) \le 1$ ;  $P(\Omega) = 1$ ;  $P(\bigcup_{i=1}^{\infty}) = \sum_{i=1}^{\infty} P(E_i)$ , falls  $E_i \cap E_j = \emptyset$ 

# 2.5 Laplace-Experiment Zufallsexperimente mit n gleich wahr-

(Übungsaufgabe!!! Ergänzen)

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

#### scheinlichen Elementarereignissen. nicht ändert, d.h. falls Dann berechnet sich die Wahrscheinlich- P(E|F) = P(E) or $P(E \cap F) = P(E) \cdot P(F)$ keit P(E) für $E \subseteq \Omega$ aus:

# $P(E) = \frac{\text{Anzahl der für E günstigen Ereignisse}}{\text{Anzahl der für E günstigen Ereignisse}}$ Anzahl der möglichen Ereignisse $\frac{\text{Mächtigkeit von E}}{\text{Mächtigkeit von }\Omega} = \frac{|E|}{n}$

2.6 Bedingte Wahrscheinlichkeit

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

### 2.7 Satz 2.2 $P(E \cap F) = P(E|F) \cdot P(F)$ $P(E \cap F) = P(F|E) \cdot P(E)$

## 2.8 Satz der totalen Wahrscheinlichkeit 1.16 Regressionsgerade y Sei $\Omega = \bigcup_{i=1}^n E_i$ mit $E_i \cap E_j = \emptyset$ für $i \neq j$

Zerlegung bzw. eine Partition von  $\Omega$ . So- $P(F) = \sum_{i=1}^{n} P(F \cap E_i) = \sum_{i=1}^{n} P(F|E_i)$ 

> Summe der Äste des Wahrscheinlichkeitsbaums zu allen Schnitten  $F \cap E_i$ ∈ R. heißt Realisation der ZV X.

d.h. die Ereignisse bilde eine disjunkte

2.9 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(FAE) P(FAE) P(F)

2.10 Formel von Bayes Hilfreich, wenn man man  $P(F|E_i)$  kennt,

# aber nicht $P(E_k|F)$ Satz 2.4 $P(E_k|F)$ = $P(F|E_k) \cdot P(E_k)$ $P(F|E_i) \cdot P(E_i)$

len Wahrscheinlichkeit.

∘*E*, *F* unabhängig

Bemerkung

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

2.11 Stochastische Unabhängigkeit

Übung Die Ereignisse E und F heißen (stochastisch) unabhängig, wenn die  $\circ F(x) = (P(X \le x) = \sum_{x_i \le x} p(x_i)$ Information über das Eintreten des einen o F(x) ist eine rechtseitig stetige **Treppen**-Ereignisses die Wahrscheinlichkeit für

das Eintreten des anderen Ereignisses nicht ändert, d.h. falls 
$$P(E|F) = P(E)$$
 or  $P(E \cap F) = P(E) \cdot P(F)$   $= \frac{P(E \cap F)}{2}$ 

gig sind, dann sind auch:  $\circ E, \overline{F}; \circ \overline{E}, F;$ o Stochastische Unabhängigkeit bedeutet nicht notwendigerweise eine

chung mit Venn Diagramm shoih. unabhangiy P(E)= \$ < P(EIT)

 $\circ A, B \neq \emptyset \text{ und } A \cap B = \emptyset$  $P(A \cap B) \stackrel{?}{=} P(A) \cdot P(B)$  $\emptyset \neq P(A) \cdot P(B)$  da P(A) > 0 und P(B) > 0=> A, B stochastisch abhängig 3 Zufallsvariable Abbildung des abstrakte Ergebnisraums

funktion mit Sprüngen bei der Realisation von  $x_i$ . 3.3 Stetige ZVs Stetige ZV X ist die Wahrscheinlichkeitsdichte f  $f: \mathbb{R} \to [0, \infty[$  definiert durch  $P(a < X < b) = \int_a^b f(x) dx$ Es gilt Falls die Ereignisse E, F unabhän- $\circ F(x) = P(X \le x) = \int_{-\infty}^{x} f(t)dt$  und  $X \le b$ ) wegen P(X = a) = 0kausale Abhängigkeit; • Veranschauli-3.4 Verteilungsfunktion

 $\circ$  F(x) ist stetig &  $P(a < X \le b) = P(a \le a)$  $\hat{J_{\text{Untergrenze}}}$  Es wird normal mit - Integriert.

∘ Diskrete ZV:  $X(\Omega) = x_1,...,x_2 (n \in \mathbb{N});$ 

∘ Stetige ZV:  $X(\Omega) \subseteq \mathbb{R}$ ; "z.B. Körpergrö-

Die Wahrscheinlichkeit P(B) für ein Er-

eignis B in R wird zurückgeführt auf die

Wahrscheinlichkeit der entsprechenden

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

Verteilungsfunktion F:  $\mathbb{R} \rightarrow [0,1]$  einer

Für eine diskrete ZV X mit  $X(\Omega) =$ 

 $x_1,...,x_n$  ( n endlich oder abzählbar un-

endlich) ist die Wahrscheinlichkeitsfunk-

 $p(x) = \begin{cases} P(X = x_i), \text{falls } x_i \in X(\Omega) \\ 0, sonst \end{cases}$ 

z.B. X = "Augensumme beim Würfeln '

3.1 Verteilungsfunktion-allg.

 $\circ \lim_{x \to -\infty} F(X) = 0 \lim_{x \to \infty} F(x) = 1$ 

 $\circ P(a < X \le b) = F(b) - F(a)$ 

ße eines Menschen"

ZV X definiert durch:

o monoton wachsend

 $\circ P(X > x) = 1 - F(x)$ 

tion definiert durch:

3.2 Diskrete ZVs

 $F(x) = P(X \le x)$ 

 $0 \le F(x) \le 1$ 

3.5 Zusammenfassung 3.6 Diskrete ZV  $\circ$  Wahrscheinlichkeitsverteilung p(x):

 $\sum_{i=1}^{n} p(x_i) = 1$ ;  $x_i$  ist Realisation der ZV. o Verteilungsfunktion F(x) ist rechtssei-

Treppenfunktion. Sprunghöhen:P(X =

 $x_i) = F(x_i) - \lim_{x \to x_i} F(x) \neq 0$  $\circ P(a < X \le b) = F(b) - F(a) \neq P(a \le X \le b)$ 

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

 $\circ$  Verteilungsfunktion F(x) ist stetig mit  $F'(x) = f(x); P(X = x_i) = 0$  $\Omega$  auf  $\mathbb{R}$ . Eine Abbildung  $X:\Omega\to\mathbb{R}$ ,  $\circ P(a < X \le b) = F(b) - F(a) = P(a \le X \le b)$  $\omega \mapsto X(\omega) = \text{heißt Zufalls variable (ZV). x}$ 

 $(b) = F(a \le X < b) = P(a < X < b)$ 

```
3.8 Erwartungswert
Der Erwartungswert E[X] = \mu einer ZV
X ist der Schwerpunkt ihrer Verteilung
or der durchschnittliche zu erwartende
Wert der ZV.
o diskrete ZV: E[X] = \sum_{i=1}^{n} x_i \cdot p(x_i)
o stetige ZV: E[X] = \int_{-\infty}^{\infty} x \cdot f(x) dx
ZV ist konstant. E[X] verhält sich linear.
Eigenschaften von E[X]:
\circ E[b] = b
\circ E[aX + b] = aE[X] + b
\circ E[X_i + ... + X_n] = \sum_{i=1}^n E[X_i]
\circ \sum_{i=1}^n x_i
3.9 Satz 3.1
Sei Y = g(X) eine Funktion der ZV X. \mu
Dann gilt:
o für diskrete ZV:E[g(X)] = \sum_{i=1}^{n} g(x).
o für stetige ZV: E[g(X)] = \int_{-\infty}^{\infty} g(x).
f(x)dx. Das vertauschen von E und g
nur bei linearen Funktionen möglich. \Rightarrow
g(E[X])
3.10 Varianz
Die Varianz einer ZV X mit u ist ein qua-
dratisches Streungsmaß. \sigma^2 = Var[X] =
E[(X-)^2] \stackrel{\text{falls x stetig}}{=} \int_{-\infty}^{\infty} (x-\mu)^2 \cdot f(x)
Die Standardabweichung \sigma = \sqrt{Var[X]}
hat im Gegensatz zur Varianz die gleiche
Dimension von die ZV X.
\circ Var[b] = 0
\circ Var[aX+b] = a^2Var[X]
3.11 Satz 3.2
Var[X] = E[X^2] - (E[X])^2 Beim Minuend
wird beim Erwartungswert nur das ein-
fach stehende x quadriert nicht f(x)!!!
3.12 Z-Transformation, Standardisie-
Sei X eine ZV mit \mu und \sigma. Dann ist
Z = \frac{X - \mu}{\sigma} = \frac{x}{\sigma} - \frac{\mu(konstant)}{\sigma}
3.13 Kovarianz
Eigenschaften:

\circ Cov[X,Y] = Cov[Y,X] 

\circ Cov[X,X] = Var[X]

\circ Cov[aX,Y] = aCov[X,Y]
Die Kovarianz zweier ZV (X, Y) ist defi-
niert durch Cov[X,Y] = E[(X-E[X])(Y-X)]
E[Y]) Die Kovarianz beschreibt die Ab-
hångigkeit zweier ZV X und Y. Je
stärker diese Korrelieren, desto (be-
tragsmäßig) größer ist die Kovarianz.
Falls X, Y(stochastisch) unabhängig \Rightarrow
Cov[X,Y]=0
```

Hilfszettel zur Klausur

von JD., Seite 2 von 4

## Falls $X_i, X_i$ paarweise unabhängig: $Var[X_1 + \dots + X_n] = \sum_{i=1}^n Var[X_i]$ $Var[X_i] = \sigma^2 \Rightarrow Var[\overline{X}] = Var[\frac{1}{n}(x_1 + ... +$ $|x_n| = \frac{1}{n^2} \sum_{i=1}^n Var[X_i] = \frac{1}{n^2} \cdot n \cdot \sigma^2 = \frac{\sigma^2}{n}$ 3.19 Quantile Sei X eine ZV mit Verteilungsfunktion F(x) und 0 . Dann ist das p-Quantil definiert als der Wert $x_n \in \mathbb{R}$ für den gilt: $F(x_n) \geq p$ . p-Quantil einer stetigen ZV mit streng monoton wachsenden $F(x:)x_p = F^{-1}(p)d$ . h. umkehrbar. 4 Spezielle Verteilung 4.1 Diskrete Verteilung 4.2 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 Erfolgswahrscheinlichkeit; $E[X] = p = \sum x_i \cdot p(x_i) = 1$ $p(1); Var[X] = p(1-p) = E[X^2] - (E[X])^2 =$ $p - p^2 = p(1 - p);$ 4.3 Binominalverteilung Anzahl der Erfolge beim n-maligen Ziehen mit Zurücklegen; Wahrscheinlichkeit $P(x = k) = \binom{n}{k} \cdot p^k$ $(1 - p)^{n-k}, k \in [0, 1, ..., n]$ ; Verteilung $X \sim B_{n,p}$ ; E[X] = np; Var[X] =np(1-p); **R:** dbinom(k,n,p)=P(X=k) ≜Wahrscheinlichkeits-

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

rbinom(k,n,p)\(\hat{p}\)kbinomialverteilte Zu-

≜Verteilungsfunktion;

fallszahlen;

qbinom(q,n,p) $\hat{=}$ q-Quantil;

3.14 Satz 3.3

 $Cov[X, Y] = E[XY] - E[X] \cdot E[Y]$ 

3.15 Varianz einer Summe von ZV

 $Var[X_1 + ... + X_n] = \sum_{i=1}^{n} Var[X_i]$ 

3.16 Overview  $\mu \sigma$ 

Falls  $X_1, X_2$  unabhängig:

 $\frac{1}{n} \sum_{i=1}^{n} E[x_i] = \frac{1}{n} \cdot n \cdot \mu = \mu$ 

 $Var[aX + b] = a^2 Var[X]$ 

3.17 E[X]

 $\sum_{i=1}^n E[X_i]$ 

3.18 Varianz

 $Var[X_i + ... + X_n]$ 

 $\sum_{i=1}^{n} \sum_{j=1}^{n} Cov[X_i, X_j]; Var[X_1 + X_2] =$ 

 $Var[X_1] + Var[X_2] + 2Cov[X_1, X_2] \circ$ Falls  $X_i, X_j$  paarweise unabhängig !!!:

E[aX + b] = aE[X] + b;  $E[X_1 + ... + E_n] =$ 

 $E[X_i] = \mu = E[\overline{X}] = E[\frac{1}{n}(X_1 + ... + X_n)] =$ 

 $X \sim U_{\{x_1,...,x_n\}}; E[X] = \frac{1}{n} \sum_{k=1}^n x_k = \overline{x};$  $Var[X] = \frac{1}{n} \sum_{k=1}^{n} x_k^2 - \overline{x}^2$ ; **R**: sample(1 :N,n)  $\hat{=}$  n Zufallszahlen zwischen 1 und 4.7 Gleichverteilung 4.8 Stetige Gleichverteilung Zufallszahlen aus einem Intervall [a, b]; **Dichte:**  $f(x) = \frac{1}{b-a}$  für  $x \in [a,b]$ ; **Verteilung:**  $X \sim U_{[a,b]}$ ;  $E[X] = \frac{a+b}{2}$ ;  $Var[X] = \frac{(b-a)^2}{12} \mathbf{R} : \frac{d}{d}unif(x, a, b) = f(x);$ puni f(x,a,b) = F(x); runi f(n) = n Zufallszahlen zwischen 0 und 1; runi f(n,a,b)  $\hat{=}$ n Zufallszahlen zwischen a und b; 4.9 Normalverteilung Beschreibt viele reale Situationen,  $dnorm(x,\mu,\sigma) = f(x)$ ;  $pnorm(x,\mu,\sigma) = teilter ZV$ ; Verteilung:  $X \sim \chi_n^2$ ; E[X] =

4.4 Hypergeometrische Verteilung

Anzahl der Erfolge beim n-maligen

Ziehen ohne Zurücklegen aus einer

Menge mit M Elementen, die Erfolg be-

deuten, und N Elementen, die Misserfolg

bedeuten. Gesamtum fang = M + N;

Wahrscheinlichkeit P(X = k) =

 $\binom{\binom{M}{k} \cdot \binom{N}{(n-k)}}{\binom{M+N}{k}}, k \in \{0, 1, ..., min\{n, M\}\};$  Ver-

teilung  $X \sim H_{M,N,n}$ ;  $E[X] = n \frac{M}{M+N}$ ;

 $Var[X] = n\frac{M}{M+N}(1 - \frac{M}{M+N})\frac{M+N-n}{M+N-1};$ 

→ 1 falls n klein im Verhältnis zu

M+N; **R**: dhyper(k, M, N, n) = P(X = k);

Verteilung der seltenen Ereignisse Häu-

figkeit punktförmiger Ereignisse in ei-

nem Kontinuum. Die durchschnittlich

zu erwartende Anzahl der Erfolge  $\lambda$  pro-

Maßeinheit (i. a. Zeiteinheit) sei bekannt.

 $k \in \mathbb{N}_0 \rightarrow diskret$  Wahrscheinlich-

**keit** $P(X = k) = \frac{\lambda^{\kappa}}{k!} e^{-\lambda} \text{ mit } \sum_{k=0}^{\infty} P(X = k)$ 

k) = 1,  $da \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{\lambda}$ ; Verteilung

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

 $e^{-\lambda} \sum_{k=1}^{\infty} \lambda \frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda} \sum_{i=0}^{\infty} \frac{\lambda^{i}}{i!} = \lambda;$ 

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

Alle Werte  $\{x_1,...,x_n\}$ einer ZV X sind gleich wahrscheinlich; **Wahrscheinlich**-

keit  $P(X = x_k) = \frac{1}{n}$ ; Verteilung

 $\frac{M}{M+N}$   $\hat{=}$  Tref ferwahrscheinlichkeit;

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

4.5 Poisson-Verteilung

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

4.6 Gleichverteilung

4.12 Chiquadrat-Verteilung

F(x);  $qnorm(q, \mu, \sigma): q - Quantil$ ; Maxi-n; Var[X] = 2n; R: dchisq(x, n) = f(x);

malstelle von f(x) bei  $x = \mu$ ; Wende- ppchisq(x, n) = F(x); Eigenschaft:  $X_1 \sim$ 

insbesondere Grenzverteilung  $Z_1,...,Z_n$  seien unabhängige, standardunabhängiger Summen; **Dichte:** normalverteilte  $ZV \Rightarrow X = Z_1^2 + + Z_n^2$  $f(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{(-\frac{1}{2}(\frac{x-\mu}{\sigma})^2)}$ ; **Verteilung:** hat Chiquadratverteilung mit n Freiheitsgraden; **Anwendungsmodell:** Sumten Ausreißer, aber schief ist (Exponentialverteilung); n>15: falls die unbekann- $X \sim N_{u,\sigma^2}$ ;  $E[X] = \mu$ ;  $Var[X] = \sigma^2$ ; **R**: men unabhängiger, standardnormalverte Verteilung annähernd symmetrisch

# dungsmodell: Schätz- und Testverfahren bei unbekannter Varianz; Verteilung:

4.13 t-Verteilung

**stelle** von f(x) bei  $x = \mu \pm \sigma$ ;  $E[aX + b] = \chi_{n_1}^2$  und  $X_2 \sim \chi_2^2 \Rightarrow X_1 + X_2 \sim \chi_{n_1 + n_2}$ 

Schätz-

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

 $\frac{X-\mu}{\sigma} \sim N_{0,1}; \ X_1 \sim N_{\mu_1,\sigma_1^2} \ \ \text{und} \ \ X_2 \sim$ 

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

 $\phi(x) = \int_{-\infty}^{x} \varphi(t)dt$ ; Quantile:  $\phi(-x) = 1$ 

 $\phi(x) \Rightarrow -x_p = x_{1-p} \text{ z.B. } -x_{0.25} = x_{0.75};$ 

 $P(\mu - \sigma \le X \le \mu + \sigma) = P(-1 \le Z \le 1) \approx$ 

 $P(\mu - 2\sigma \le X \le \mu + 2\sigma) = P(-2 \le Z \le 2) \approx$ 

 $P(\mu - 3\sigma \le X \le \mu + 3\sigma) = P(-3 \le Z \le 3) \approx$ 

tezeiten 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 Ereignis-

ses; 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 ex-

ponentialverteile ZV X ist gedächtnis-

los, d.h. P(X > s + t)|X > t = P(X > s);

 $N_{\mu_2,\sigma_2^2} \Rightarrow X_1 + X_2 \sim N_{\mu_1 + \mu_2,\sigma_1^2 + \sigma_2^2};$ 

 $X_1, X_2$  stochastisch unabhängig

werte:  $Z = \frac{x - \mu}{\sigma} \sim N_{0,1}$ ;

4.11 Exponentialverteilung

4.10 Standardnormalverteilung

 $Y \sim t_n$ ; E[Y] = 0 für n > 1;  $Var[Y] = \frac{n}{n-2}$ für n > 2; **R**:  $\frac{d}{dt}(y, n) = f(x)$ ; pt(y, n) = F(x);  $qt(y,n) = F^{-1}(x)$ ; Eigenschaften: Für  $n \to \infty$  $\infty$ :  $t_n \rightarrow N_{0,1}$ ; Achsensymmetrie

 $Z \sim N_{0.1}$  und  $X \sim \chi_n^2 \Rightarrow Y = \frac{Z}{X}$  ist t-

verteilt mit n Freiheitsgraden; Anwen-

der Dichtefunktion  $\Rightarrow -y_p = x_{1-p}$ 

# Abbildung Dichtefunktion Modellierung von Lebensdauern, War-

# 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 Erwartungs-

lich dominanter?! als die anderen.Für

ist(Binomialverteilung);  $n \le 15$ : falls die

unbekannte Verteilung annähernd nor-

wert  $\mu$  und Varianz  $\sigma^2$ . Dann gilt für hinreichend große n (>30) und  $\overline{X} = \frac{1}{n} \sum_{i=1}^{n}$ näherungsweise:

 $\sum_{i=1}^{n} X_i \sim N_{n\mu,n\sigma^2} \&$  $\frac{\sum X_i - n\mu}{\sqrt{n} \cdot \sigma} \sim N_{0,1}$ 

 $\sum X_i$  bezieht sich auf Y;  $\sum X_i - n\mu$  bezieht sich auf  $X_i$ ;  $\overline{X} \sim N_{\mu,\frac{\sigma^2}{\mu}} & \frac{\overline{X}-\mu}{\sigma} \sim N_{0,1}$ ; Der Satz gilt sogar allgemeiner, wenn die  $X_i$  abhängig und nicht identisch verteilt sind, vorausgesetzt kein  $X_i$  ist deut-

malverteilt ist;

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. Faustre-

gel: **Je** schiefer die Verteilung der  $X_i$ **desto** größer muss n sein: n>30: falls die unbekannte Verteilung ohne markan-

6.2 Punkschätzer  $\phi(a) - (1 - \phi(a)) = 2\phi(a) - 1 \text{ or } 1 - \phi(-a) -$ E[X]: Stichprobenmittel:  $X = \frac{1}{n} \sum_{i=1}^{n} X_i$ ;  $\phi(-a) = 1 - 2\phi(-a)$ Varianz: Stichprobenvarianz:  $s^2$  = **5.3**  $\phi^{-1}$  $\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$ ; Schätzwert für wahren Parameter, aber keine Aussage über Unsicherheit der Schätzung, Geringe Sicherheit für wahren Parameter; 6.3 Intervallschätzer Intervall für wahren Parameter, vorgegebener Sicherheit; Vorqnorm(1gabe (95% or 99%); Dichtefunkti- $-\phi^{-1}(p) = \phi^{-1}(1-p)$  Zusammenhang  $P(-a \le \overline{x} \le a) > 0.95$ ;  $\sigma ist$  unbekann Aufgabentypen: Seien  $X_i$  i.i.d. ZV mit  $\mu$  und  $\sigma^2$ , aber unbekannter Verteilung.  $P(x_{0.025} < \frac{x - \mu}{\sigma} \sqrt{n} < x_{0.975}) \ge 0.95$ Dann sind  $Z_1 = \frac{\sum X_i - n\mu}{\sqrt{n}\sigma}$  und  $Z_2 = \frac{X - \mu}{\sigma}$  $-1.96; N_{0.1}; 1.96;$ näherungsweise standardnormalverteilt. o Es lassen sich Wahrscheinlichkeiten für 6.4  $\mu$ , unbekannt,  $\sigma^2$ , bekannt  $\sum X_i, \overline{X}, Z_1$  oder  $Z_2$  berechnen.  $\sim \text{Es lässt sich } \frac{1}{n} \text{ bestimmen, so dass, } I = ]\overline{X} - \phi^{-1} (1 - \frac{\alpha}{2}) \frac{\sigma}{\sqrt{n}}$ zu vorgegebener Schranke k und Wahrscheinlichkeit p gilt:  $P(Z_i > k) \ge p$  or  $qnorm(1-\frac{\alpha}{2})$  $P(-k \le Z_i \le k) \ge p$ 5.4 Stichprobenvert.normalvert. φ-1/0,95)≈ 1,64S Grundgesamt. φ<sup>-1</sup>(0,975) ≈ 1,96 95% 2,5% 5.5 Stichprobenmittel 99% 0,5% \$^1(0,995)≈ 2,576 Die Stichprobenfunktion  $\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$ ist eine erwartungstreue Schätzfunktion für Erwartungswert  $\mu$ , d. h. $E[\overline{X}] = \mu$ 5.6 Stichprobenvarianz Die Stichprobenfunktion 6.5  $\mu \& \sigma^2$ , unbekannt  $\frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})^2 = \frac{1}{n-1} (\sum_{i=1}^{n} X_i^2 - X_i^2)$  $I = \overline{X} - t_{n-1}^{-1} (1 - \frac{\alpha}{2}) \frac{S}{\sqrt{n}}, \overline{X} + t_{n-1}^{-1} (1 - \frac{\alpha}{2}) \frac{S}{\sqrt{n}}$  $n\overline{X}^2$ )ist eine erwartungstreue Schätz-6.6 Zusammenfassung funktion für die Varianz  $\sigma^2$ , d. h.  $E[S^2] = \sigma^2$ ;  $E[\overline{X}] = E[\frac{1}{n}\sum X_i] =$ Wie verändert sich das  $(1 - \alpha)$ -Konfidenzintervall, n-größer ⇒ I  $\frac{1}{n}E[\sum X_i] = \frac{1}{n}\sum_{i=1}^{n}E[X_i] = \frac{1}{n}n\mu = \mu;$ kürzer;  $1-\alpha$  größer  $\Rightarrow$  I länger; Für **7.4 Klassischer Parametertest**  $\frac{L}{2} = 2\phi^{-1}(1 - \frac{\alpha}{2})\frac{\sigma}{\sqrt{n}}\frac{1}{2} = 2\phi^{-1}(1 - \frac{\alpha}{2})\frac{\sigma}{\sqrt{4n}}$  $Var[\overline{X}] = Var[\frac{1}{n}\sum X_i] = \frac{1}{n^2}Var[\sum X_i] =$  $\frac{1}{n^2}n\sigma^2 = \frac{\sigma^2}{n}$ ; Seien  $X_i(i=1,...,n)$  unab- **6.7** Aufgabentypen hängige normalverteilte ZV mit Erwar- Geg: n, 1- $\alpha$ ; Ges: I s.o. Geg:  $\overline{X}$ ,  $\sigma$ , 1 –  $\alpha$ , L; tungswert  $\mu$  und Varianz  $\sigma^2$ . Dann gilt:  $L = 2\phi^{-1}(1-\frac{\alpha}{2})\frac{\sigma}{\sqrt{n}}$ ; **Ges:** n;  $\sqrt{n} > 2\phi^{-1}(1-\frac{\alpha}{2})\frac{\sigma}{\sqrt{n}}$ bei bekannter Varianz:  $\frac{\overline{X}-\mu}{\sigma}\sqrt{n}\sim N_{0,1}; \quad \frac{\alpha}{2})\frac{\sigma}{L}$  Geg: n, I, L; Ges: 1-  $\alpha$ ;  $1-\frac{\alpha}{2}=$ 

 $(n-1)S^2 = \sum (x-\overline{x})^2$ 

6 Konfidenzintervall

6.1 Begriffe

kannter Varianz:  $\frac{X-\mu}{C}\sqrt{n} \sim t_{n-1}$ ;

Irrtumswahrscheinlichkeit =  $\alpha$ ; Konfi

denzniveau =  $1 - \alpha$  = ; Konfidenzintervall

Hilfszettel zur Klausur von JD., Seite 3 von 4

 $1 - \phi(-a)$ ;  $P(-a < Z < a) = \phi(a) - \phi(-a) =$ 

 $\sim \chi_{n-1}^2$ ; Bei unbe-  $\phi(\frac{L\sqrt{n}}{2\sigma})$ 

7 Hypothesentests

wert  $\mu$  gültig ist or nicht.

achtetes Signifikanzniveau

7.2 Null- und Gegenhypothese

7.1 Def

Basierend auf n unabhängig und iden-

Treffen der Testentscheidung, basierend auf einer konkreten Stichprobe 
$$\{x_1, ..., x_n\}$$
; Berechnung der Realisation  $tg = TG(x_1, ..., x_n)$  der Prüfgröße TG; Ablehnungsbereich / Kritischer Bereich C: Werte der Testgröße, die für H1, sprechem & bei Gültigkeit von  $H_0$  mit Wahrscheinlichkeit  $\leq \alpha$  (meist  $0.1, 0.05,$  or  $0.01$ ) auftreten Fehler 1. Art: $\alpha$  ist die Wahrscheinlichkeit, dass  $H_0$  verworfen wird, obwohl sie richtig ist. Annahmebereich: Komplement  $\overline{C}$  des Ablehnungsbereichs.  $H_0$  kann nicht abgeleht werden, falls  $tg \in \overline{C}(P(tg \in \overline{C}) \geq 1 - \alpha)$ . Fehler 2. Art: Die Wahrscheinlichkeit, dass  $H_0$  nichtig  $t_0$  wird abgelehnt.  $t_0$  wird abgelehnt  $t_0$  wird abgelehnt.  $t_0$  wird abgelehnt  $t_0$  wird abgelehnt.  $t_0$  wird abgelehnt, falls  $t_0$   $t_0$ 

 $\overline{C}$ )  $\geq 1 - \alpha \Leftrightarrow TG^* \in [\phi^{-1}(\frac{\alpha}{2}), \phi^{-1}(1 - \frac{\alpha}{2})];$ tisch Verteilte (i.i.d) Zufallsvariablen  $X_1,...,X_n$  (Messungen) soll eine Entschei-Wird dann  $H_0$  verworfen, spricht man dung getroffen werden, ob eine Hypothevon einer signifikanten Schlussfolgerung. se für einen unbekannten Erwartungs-Kann  $H_0$  nicht verworfen werden, dann lässt sich keine Aussage über den Fehler 2. Art treffen & man spricht von einer  $\alpha$  = Signifikanzniveau/ Fehlerwahrschwachen Schlussfolgerung. scheinlichkeit TG = Prüfgröße; TG\* = 7.5 Zweiseitiger Gauß Test standardisierte Prüfgröße; siginifikante  $H_0: \mu = \mu_0$  gegen  $H_1: \mu \neq \mu_0$ ;  $\overline{X} \sim$ Schlussfolgerung =  $H_0$  verworfen  $\rightarrow$  klas- $N_{\mu_0,\sigma_0^2/n} \Rightarrow \frac{\overline{X}-\mu_0}{\sigma_0}\sqrt{n} \sim N_{0,1}; P_{\mu 0}(\overline{X} \in$ sischer Parametertest; schwache Schlussfolgerung =  $H_0$  wird nicht verworfen  $\rightarrow$ klassischer Parametertest. p-Wert = beob- $C) \le \alpha \Leftrightarrow |TG| = \frac{|\overline{X} - \mu_0|}{\sigma_0} \sqrt{n} > \phi^{-1} (1 - \frac{\alpha}{2});$ **Testentscheidung:**  $H_0$  wird abgelehnt, falls  $|TG| > \phi^{-1}(1-\frac{\alpha}{2})$ ;  $H_0$  wird angenom-Modell: Verteilung der Grundgesamtheit or Testgröße **TG** (häufig  $\bar{x}$ ) ist bekannt men, falls  $|TG| \le \phi^{-1}(1-\frac{\alpha}{2})$ bis auf einen Parameter, z.B. μ, für den 7.6 Einseitiger Gauß Test eine Hypothese aufgestellt wird. TG ~ 7.7 linksseitig  $N_{\mu,\sigma^2}$ ; Nullhypothese:  $H_0$ : Angezweifel- $H_0: \mu \ge \mu_0 \text{ gegen } H_1: \mu < \mu_0$ te Aussage, der widersprochen werden 7.8 rechtsseitig kann, wenn die Stichprobe einen Gegen- $H_0: \mu \le \mu_0 \text{ gegen } H_1: \mu > \mu_0$ beweis liefert.  $H_0: \mu = \mu_0$ ; Gegenhypo**these**  $H_1$ : Gegenteil von  $H_0$  z.B.  $H_1 \neq \mu_0$ ;  $P_{\mu 0}(\overline{X} \in C) \leq \alpha \Leftrightarrow TG = \frac{X - \mu_0}{C} \sqrt{n} < C$ 7.3 Ablehnungsbereich, Fehler 1. & 2.  $\phi^{-1}(\alpha)$ ; Testentscheidung:  $H_0$  wird abgelehnt falls,  $TG < \phi^{-1}(\alpha)$ ;  $H_0$ wird angenommen, falls  $TG \ge \phi^{-1}(\alpha)$ ;

7.9 Varianten Gauß Test,  $\sigma^2$  bekannt,  $\mu$ unbekannt Prüfgröße $tg = \frac{X - \mu_0}{\sigma_0} \sqrt{n}$  $tg < \Phi^{-1}(\alpha)$ **7.10 t-Test,**  $\mu$ ,  $\sigma^2$  unbekannt 7.11 p-Wert Wahrscheinlichkeit, bei Zutreffen von  $H_0$ 

Fehler 1. Art, mit standardisierter Prüf-

größe TG\* gilt:  $P(TG \in C) \le \alpha \Leftrightarrow TG^* \in$ 

 $]-\infty; \phi^{-1}(1-\frac{\alpha}{2})[\cup]\phi^{-1}(1-\frac{\alpha}{2}); \infty[; P(TG \in$ 

den beobachteten Wert tg der Prüfgröße or einen noch stärker von  $\mu_0$  abweichen7.12 Zusammenhang I & Hypothesentests zweiseitig zum Konfidenzniveau  $1 - \alpha$ ;  $H_0$  wird abgelehnt, falls  $\mu_0 \notin I$ ;  $H_0$  wird angenommen, falls  $\mu_0 \in I$ ; Das Konfidenzniveau ist der Annahmebereich von Ho zum Si-7.13 Zusammenfassung klass. Hy-

# gnifikanzniveau $\alpha$ ; po.test

Signifikanzniveau  $\alpha$  wird vorgegeben;  $\alpha$  & Verteilung der Testgröße unter  $H_0$ wir der Ablehnungsbereich ermittelt. Je kleiner (größer)  $\alpha$ , desto kleiner (größ-

 $H_0$  wird abgelehnt, falls der ermittelte

Wert der Testgröße (beobachteter Wert)

in C liegt. !: Die tg hängt von der konkre-

werden kann. Je kleiner der Wert, desto

kleiner ist der Fehler 1. Art & umso

signifikanter ist die Testentscheidung.

Nice to know Anhand des p-Werts kann

man für beliebige Werte von  $\alpha$  eine

Falls p - Wert < 1%: sehr hohe Signifi-

Falls  $1\% \le p - Wert < 5\%$ : hohe Signifi-

Falls  $5\% \le p - Wert \le 10\%$ : Signifikanz

Falls p - Wert > 10%: keine Signifikanz

Testentscheidung treffen;

ter) ist der Ablehnungsbereich;  $!: \alpha \& C$  hängen **nicht von** der konkreten Stichprobe ab:

ten Stichprobe ab. Sie ist eine ZV. 7.14 Test mittels p-Wert  $\alpha$  wird vorgegeben.

Berechnung des p-Werts anhand der kon-kreten Stichprobe mit der Verteilung der Tg unter  $H_0$ ;

!:Der p-Wert hängt von der konkreten Stichprobe ab, ist eine ZV.  $H_0$  wird abgelehnt, falls  $p - Wert \le \alpha$ .; 8 Fehleranalyse

Derzeit ausgeklammert 9 Interpolation

Zu gegebenen Punkten  $(x_i, y_i), i = 0, ..., n$ mit  $x_i \neq x_i$  für  $i \neq j$  eine Funktion G (dies ist nicht eindeutig! Abhängig von der

0,...,n (Interpolations bedingung). Interpolation ist ungeeignet für verauschte Daten. Lösung: Approximation der kleinsten Quadrate. 9.1 Begriffe Extrapolation \(\hat{=}\) N\(\alpha\)herungwerte f\(\bar{u}\)r x-

Werte außerhalb der Stützstellen; Dividierende Differenzen 

Koeffizienten c<sub>i</sub> lassen sich rekursiv durch wie-

Funktionsklasse), so dass  $G(x_i) = y_i$ , i =

derholte Bildung von "Differenzquotienten"berechnen 9.2 Vandermonde/klassisch

Unterschiedliche Darstellungen für ein Interpolations polynom  $G(x) = p_n(x)$ vom Grad *n* haben unterschiedliche Eigenschaften bei der nume-

Berechnung. Monombasis: rischen  $x^0, x^1, x^2, x^3, ...; p_n(x) = a_n x^n + ... +$  $a_1x^1 + a_0x^0$ ; **Ziel:** Bestimmung d. Koeffizienten  $a_0, a_1, ..., a_n$  sodass

dermonde Matrix; Eigenschaften: Die

Vandermonde Matrix ist nicht singulär(

falls alle  $x_i$  verschieden); Rechenauf-

wand:  $\mathcal{O}(n^3)$ ; Für große n sehr schlecht konditioniert & als Allgemeiner Ansatz

polation ( Näherungwerte für x-Werte au-

ßerhalb der Stützstellen ) kann zu großen

Darstellung des Interpolanten, die auf

ein gestaffeltes LGS führt & einfa-

che Hinzunahme weiterer Punkte er-

laubt.  $p_n(x) = c_0 + c_1(x - x_0) + ... +$ 

Das Resultierende LGS für die Koeffizi-

enten c<sub>i</sub> hat gestaffelte Form. **Interpola**-

punktoperationen; Hinzufügen weiterer

Stützstellen ohne großen Aufwand. Andere Koeffizienten bleiben unverändert.

9.5 Dividierende Differenzen

 $\begin{vmatrix} \frac{3}{3} & \frac{(-2)}{2} & \frac{2}{2} \\ \frac{3}{1-3} & \frac{3-(-5)}{1-3} & \frac{-4-2}{1-(-2)} \end{vmatrix}$ 

Abweichungen führen.

 $c_n(x-x_0)(x-x_1)...(x-x_{n-1})$ 

Polynom vom Grad n

tionsbedingungen?

9.4 Newton

ungeeignet.

9.3 Lagrange

 $\begin{pmatrix} x_0^n & \dots & x_0^2 & x_0 & 1 \\ x_1^n & \dots & x_1^2 & x_1 & 1 \\ x_2^n & \dots & x_2^2 & x_2 & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \end{pmatrix} \begin{pmatrix} a_n \\ a_{n-1} \\ \vdots \\ \vdots \end{pmatrix} = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ \vdots \\ \vdots \end{pmatrix}$ 

Hilfszettel zur Klausur von **JD**., Seite 4 von 4

# $p_n(x_i) = y_i = a_n x_i^n + ... + a_1 x_i^1 + a_0 x^0$ für i =

0, ..., n; Für die eindeutige Lösung n+1

9.8 Horner Schema  $p_3(x) = a_3 x^3 + a_2 x^2 + a_1 + a_0 = ((a_3 + a_2)x + a_1 + a_0) = (a_3 + a_2)x + a_1 + a_0 = (a_3 + a_2)x + a_1 + a_1 + a_2 +$ 

9.6 Effizienz

9.7 klasisch

 $a_1$ ) $x + a_0$ ; Allg.:  $p_n(x) = (...(a_n x + a_{n-1})x +$ ... +  $a_1$ ) $x + a_0$ ; **Aufwand:** n Mult. 9.9 Interpolationsfehler f hinreichend glatt ist &

 $p_n(x) = a_n x^n + ... + a_0$ ; Aufwand: 2n-1

Gleichungen: Interpolationsbedingundas eindeutige Interpolati-

onspolynom von Gradn *n*, dann gilt fürn den Interpolationsfehler:  $f(x) - p_n(x) = \frac{f^{(n+1)}(\theta)}{(n+1)!}(x - x_0)...(x - x_n)$ 

Die Koeffizientenmatrix ist die sog. Vandermonde Matrix; Eigenschaften: Die Vandermonde Matrix ist nicht singulär (
falls alle 
$$x_i$$
 verschieden); Rechenaufwand:  $\mathcal{O}(n^3)$ ; Für große n sehr schlecht konditioniert & als Allgemeiner Ansatz ungegengt.

mit  $\theta \in [x_0; x_n]$  Vergleichbar zum Restglied bei der Taylorreihenentwicklung; Bemerkung:  $\theta$  unbekannt, daher nur Fehlerabschätzung; Fehler ist Abhängig von der Verteilung der Stützstellen; Der Fehler ist bei großen n an den Intervallrändern

deutlich größer, als in der Intervallmitte 9.10 Wahl der Stüztstellen **2 Formeln**;  $p_n(x) = y_0 L_0(x) + y_1 L_1(x) + ... +$ Runge Funktion  $(f) = \frac{1}{1+25x^2}$  äquidistan $y_n L_n(x)$ ;  $L_k(x) \prod_{j=0; j \neq k}^n \frac{x - x_j}{x_k - y_j}$ ; Jede Basiste Stützstellen das Interpolationspolynom nicht immer gegen die zugrundeliefunktion  $L_k(x)$  ist ein Polynom vom Grad gende stetige Funktion konvergiert, wenn  $\leq n$ ; **Bemerkung:** Findet Anwendung bei die Anzahl der Stützstellen & damit Numerischer Integration; Wenn Stützstelder Grad des Polynoms wächst. Lösung: len  $x_i$  gleich bleiben & nur  $y_i$  ändern  $\Rightarrow$ Nicht-aquidistante Verteilung der Stützkeine Neuberechnung; Rechenaufwand stellen, dichter an den Intervallgrenzen.  $\mathcal{O}((n+1)^2)$ ; Kommen neue Stützpunkte

hinzu ⇒ Neuberechnung!; Die Interpolahaben die Eigenschaft; senkrechte Protionspolynome liefern nur sinnvolle Näjektion von gleichverteilten Punkten auf herungswerte für x-Werte, die zwischen dem Einheitskreis.  $t_k = cos \frac{(2k-1)\pi}{2n}, k =$ den gegebenen Stützstellen liegen; Extra-

9.11 Chebyshev-Punkte

1, ..., n, auf ] - 1, 1[; Invtervall:  $]a, b[: x_k =$  $\frac{a+b}{2} + \frac{b-a}{2}t_k$ .  $\Rightarrow$  Fehler wird gleichmäßiger verteiltund Konvergenz erreicht. 9.12 Schwächen der Polynominterpola-

Hoher Rechenaufwand bei meist keiner hoher Differenzierbarkeitsgrad benötigt wird; RB kann Interpolationsfehler sehr groß sein; Bei wachsenden n ist es unmöglich eine Konvergenz gegen die zu interpolierenden Funktion sicherzustellen;  $\mathbf{R}$ : approx  $\hat{=}$  lin Interpolation; Spline ≜ Spline interpolation; Bibliotheken für Polynominterpolation; **Vorteile:** Rechenaufwand  $\mathcal{O}(n^2)$  Gleit-9.13 Spline

Jede Funktion  $S_i$  ist ein Polynom vom Grad  $n \le k$ ; S(x) ist (k-1) - mal stetig differenzierbar, d.h. für alle  $x_i$  (i = 1, ..., n-1) gilt:  $S_{i-1}(x_i) = S_i(x_i)$ ; 9.14 Kubisch

**Ansatz:**  $S_i = a_i + b_i(x - x_i) + c_i(x - x_i)^2 + c_i(x - x_i)^2$  $d_i(x-x_i)^3$ ; Gleichungssystem: 4n Parameter  $a_i, b_i, c_i, d_i (i = 0, ..., n - 1)$ ; 2n Interpolationsbedingungen: am Rand je

Basierend auf äquidistanten Knoten  $t_j = \frac{1}{(k)}$  Methode  $(i = 0, 1, ..., n - 1) \Rightarrow$  Stetigkeit; **Stetigkeit der 1. Abl:**  $S_{i}(x_{i+1}) = S_{i+1}(x_{i+1}); \Leftrightarrow$  $S'_{i}(x_{i+1}) - S'_{i+1}(x_{i+1}) = 0$ ; für i = 0, 1, ..., n -2; Stetigkeit der 2. Abl.:  $S_i''(x_{i+1}) =$  $S_{i+1}^{"}(x_{i+1}); S_{i}^{"}(x_{i+1}) - S_{i+1}^{"}(x_{i}+1) = 0;$  $f\ddot{u}r^{i} = 0, 1, ..., n-2$ ); natürlicher Rand**bedingungen:**  $S_0''(x_0) = 0$ ;  $S_{n-1}''(x_n) = 0$ ; nach geschickter Umformung der Gleichungen hat das LGS Tridiagonalform.

der Fehler? 10.5 Ordnung Integrationsregel Eine Integrationsregel hat Ordnung p, **Rechenaufwand**  $\mathcal{O}(n)$  Gleitpunktopera-

> 1 exakte Werte liefert;  $T_1$  Ordnung 2  $\Rightarrow$  exakt für Polynome Grad  $\leq$  1; Ordnung Newton-Cotes Regeln: mind. Ord-

 $p_k = \text{Interpolationspolynom}$ ;  $I_n = \text{Quadra-}$ turformel; K =Fehlerkonstante des Verfahrens.; Singularität \(\hat{=}\) isolierter Punkt, der ungewöhnliches Verhalten zeigt;

# 10.2 Newton-Cotes

10 NumInt

krete Punkte.

10.1 Def

das Int. von f(x);  $\int_0^1 f(t)dt \approx \int_0^1 p_k(t)dt =$  $\sum_{i=0}^{k} \alpha_i f(t_i)$  Das Interpolationspolynom muss nicht explizit aufgestellt werden, es dient vorab der Bestimmung der Gewichte  $\alpha_j$ ;  $\int_0^1 p_k(t) = \int_0^1 \sum f(t_j) L_j(t) dt =$  $\sum f(t_i) \int_0^1 L_i(t) dt$ 

nur eine.  $S_i x_i = y_i$ ;  $S_i(x_{i+1}) = y_{i+1}$  für

Verbesserung der Näherung: Aufteilung

in kleine Teilintervalle & Summe von

Rechtecksflächen bilden; Interpolations

mit Polynom höheren Gredes durch dis-

# 10.3 Trapezregel

 $T_1: \int_0^1 f(t)dt \approx \frac{1}{2}(f(0)+f(1)); \int_a^b f(x)dx \approx$  $\frac{(b-a)}{2}(f(a)+f(b));$  $T_n$ : Für Teilintervalle mit gleicher Länge:  $h = \frac{b-a}{n}$ ;  $T_n = h(\frac{f(x_0)}{2} + f(x_1) + ... + f(x_{n-1}) +$ 

# 10.4 SimpsonRegel

 $\int_{a}^{b} f(x)dx \approx \frac{b-a}{6} (f(a) + 4f(\frac{a+b}{2}) + f(b));$ Für n = 1:  $\frac{(b-a)}{2\cdot 1} \frac{1}{3} (f(a) + 4f(\frac{a+b}{2}) + f(b));$ Für n allg.:  $\frac{(b-a)}{2n} \frac{1}{3} (f(a) + 4(a+h) +$ ... + 4f(b-h) + f(b)  $S_n$ : Beachte gerade Anzahl an Teilinvervallen!; Für 2n Teilintervalle, 2n+1 Knoten

mit gleicher Länge  $h = \frac{b-a}{2n}$ ;  $S_2 =$ 

 $\frac{h}{3}(f(x_0)+4f(x_1)+2f(x_2)+4f(x_3)+f(x_4));$ 

 $S_1: \int_0^1 f(t)dt \approx \frac{1}{6}(f(0) + 4f(0.5) + f(1));$ 

3-Rule Falls  $\alpha_i$  positiv. Integrations regel stabil;  $k \le 7\&k = 9 \Rightarrow$  positive Gewichte; Bei halbierung der Intervalle Nachfrage vervierfacht or versechszehnfacht sich

Simpsor

# wenn sie für Polynome vom Grad ≤ p

nung k+1 (k: GRad des Interpolationspolynoms); Beweis der Ordnung: 1 =  $\int_0^1 x^0 dx \stackrel{!}{=} ; \frac{1}{2} = \int_0^1 x dx \stackrel{!}{=} ; \frac{1}{3} = \int_0^1 x^2 \stackrel{!}{=} ;$  $\frac{1}{4} = \int_0^1 x^3 \stackrel{!}{=};$ 10.6 Fehler Quadratur

Für (globalen ) Fehler  $e_{In} = \int_{a}^{b} f(x)dx - I_{n}$ einer Quadraturformel  $I_n$  der Ordnung pauf [a, b] gilt:  $|e_{In}| = (b-a)h^p K|f^{(p)}(\xi)|.\xi \in$ 

 $|a,b|, h = \frac{b-a}{n} \& |e_{In}| \le (b-a)h^p K$ . Das Intergral des  $p_k$  diens al Appr. für  $\max_{a \le x < b} |f^{(p)}(x)|$ ;

### 10.7 Fehler $T_n$ Der Fehler ist proportional zu $h^2$ ; Eine

Halbierung der Intervalllänge reduziert den Fehler um den Faktor 1/4; Ein In- 11.6 Potenzen tegral kann beliebig genau approx. werden, falls h entsprechend klein gewählt wird. Aber Rundungsfehler bei vielen Rechenoperationen, verschlechtert wieder das Ergebnis. Vorteil von Verfahren höherer Ordnung: Weniger Teilintervalle nötig.  $|e_{T_n}| \le \frac{h^2}{12}(b-a)max_{a \le x \le b}|f''(x)|, K =$  $\frac{1}{12}$ ,  $h = \frac{b-a}{n}$ 

### 10.8 Fehler $S_n$ Der Fehler ist proportional zu $h^4$ ; Eine

Halbierung der Intervalllänge reduziert den Fehler um den Faktor  $\frac{1}{16}$ ;  $|e_{Sn}| \leq$  $\frac{h^4}{180}(b-a)max_{a\leq x\leq b}|f^4(x)|, h=\frac{(b-a)}{2n}, K=$ 

# 10.9 Grenzen NeCo

viele äquidistante Knoten → Gewichte negativ → Verfahren instabil; geschlossene NeCoRe → Funktionsauswertung an RB → Problem mit Singularitäten. größtmögliche Ordnung unerreichbar wegen äquidistanten Knoten; Lösung: 10.10 GauQua

11 Allgemein

11.1 Symbole

Standardabweichung  $\hat{=}\sigma$ 

 $x^n \hat{=} nx^{n-1}$ sinx = cosx; cosx = -sinx;  $tanx = \frac{1}{cos^2x} = 1 + \frac{1}{cos^2x}$  $tan^2x$ ;  $cotx = -\frac{1}{sin^2x} = -1 - cot^2x$ ;

 $f_n(x) \Rightarrow y' = f_1'(x) + f_2'(x) + ... + f_n'(x)$ ; **Pro**-

Kettenregel  $f'(x) = F'(u)u'(x) = \hat{F}'(u)$ :

 $e^{x} = e^{x}$ ;  $a^{x} = (\ln a) \cdot a^{x}$ ;  $\ln x = \frac{1}{x}$ ;  $\log_a x = \frac{1}{(\ln a) \cdot x}$ ; 11.3 Abl.Regeln

11.2 Abl.

**Faktorregel**  $y = C \cdot f(x) \Rightarrow y' = C \cdot f'(x)$ ; Summerregel  $y = f_1(x) + f_2(x) + ... +$ 

**duktregel**  $y = u \cdot v \Rightarrow y' = u' \cdot v + v' \cdot u;$  $y = u \cdot v \cdot x \Rightarrow y' = u' \cdot v \cdot w + u \cdot v' \cdot w + u \cdot v \cdot x';$ Quotientenregel  $y = \frac{u}{v} \Rightarrow y' = \frac{u' \cdot v - u \cdot v'}{v}$ ;

Ableitung der Äußeren Funktion; u'(x): Ableitung der Inneren Funktion

11.4 Integralregel, elementar

**Faktorregel**  $\int_a^b C \cdot f(x) dx = C \cdot \int_a^b f(x) dx$ ;

**Summenregel**  $\int_a^b [f_1(x) + ... + f_n(x)]dx =$  $\int_a^b f_1(x)dx + \dots + \int_a^b f_n(x)dx$ ; Vertau-

schungsregel  $\int_{a}^{a} f(x)dx = -\int_{a}^{b} f(x)dx$ ;

 $m, n \in \mathbb{N}^*$ ;

a > 0, b > 0:

Exponenten

beliebig reele

 $a,b \in \mathbb{R}$ 

 $a > 0 : a^b$ 

 $=e^{b \ln a}$ 

 $\int_{a}^{a} f(x)dx = 0; \int_{a}^{b} f(x)dx = \int_{a}^{c} f(x)dx +$  $\int_{a}^{b} f(x)dx \text{ für } (a \le c \le b);$ 

11.5 Berechnung best. Integr.  $\int_{a}^{b} f(x)dx = [F(x)]_{a}^{b} = F(b) - F(a)$ 

 $a^0 = 1, a^{-n} = \frac{1}{a^n}$  $\frac{a^m}{a^n} = a^{m-n} text f ra \neq 0$  $!(a^m)^n = (a^n)^m = a^{m \cdot n}$  $a^n \cdot b^n = (a \cdot b)^n$ 

 $\frac{a^n}{b^n} = (\frac{a}{b})^n$  für  $b \neq 0$ 

 $\sqrt{a^2} = |a|$ ;  $b = a^n \Leftrightarrow a = \sqrt[n]{b}$ ;  $\sqrt[n]{a} = a^{\frac{1}{n}}$ ;

 $\sqrt[n]{a \pm b} \neq \sqrt[n]{a} \pm \sqrt[n]{b}$ 

 $\sqrt[n]{a^m} = (a^m)^{\frac{1}{n}} = a^{\frac{m}{n}} = (a^{\frac{1}{n}})^m = (\sqrt[n]{a})^m$  $\sqrt[m]{\sqrt[n]{a}} = \sqrt[m]{a^{\frac{1}{n}}} = (a^{\frac{1}{n}})^{\frac{1}{m}} = a^{\frac{1}{m \cdot n}} = \sqrt[m]{a}$ 

 $\sqrt[n]{a} \cdot \sqrt[n]{b} = (a^{\frac{1}{n}}) \cdot (b^{\frac{1}{n}}) = (ab)^{\frac{1}{n}} = \sqrt[n]{ab}$ 

 $\frac{\sqrt[n]{a}}{\sqrt[n]{b}} = \frac{a^{\frac{1}{n}}}{b^{\frac{1}{n}}} = \left(\frac{a}{b}\right)^{\frac{1}{n}} = \sqrt[n]{\frac{a}{b}} \text{ für } b > 0$  $\Rightarrow m, n \in \mathbb{N}^* : a > 0, b > 0$ 

11.8 Abc-Formel

 $x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$ ;  $x_{1,2} = \frac{2a}{1 - \sqrt{b^2 - 4ac}}$ 

# 11.9 Bin.Formel

$$(a+b)^2 = a^2 + 2ab + b^2$$
 1. Binom;  $(a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$ ;  $(a+b)^4 = a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4$ 

$$(a-b)^2 = a^2 - 2ab + b^2$$
; 2. Binom;  $(a-b)^3 = a^3 - 3a^2b + 3ab^2 - b^3$ ;  $(a-b)^4 = a^4 - 4a^3b + 6a^2b^2 - 4ab^3 + b^4$ 

$$(a+b)(a-b) = a^2 - b^2$$
 3. Binom;

# 11.10 Einigungen

o Beim Runden mind. eine Nachkommastelle.