

## 1 Beschreibende Statistik

### 1.1 Beschreibende/Deskriptive Statistik

Beobachtete Daten werden durch geeignete statistische Kennzahlen charakterisiert und durch geeignete Grafiken anschaulich gemacht.

### 1.2 Schließende/Induktive Statistik

Aus beobachtete Daten werden Schlüsse gezogen und diese im Rahmen vorgegebener Modelle der Wahrscheinlichkeitstheorie bewertet.

### 1.3 Grundgesamtheit

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

### Lagemaße

### 1.4 Modalwerte $x_{mod}$

Am häufigsten auftretende Ausprägungen (insbesondere bei qualitativen Merkmalen)

### 1.5 Mittelwert, quantitativ

$R: mean(x)$   
Schwerpunkt der Daten.  
Empfindlich gegenüber Ausreißern.  
 $\bar{X} = \frac{1}{n} \sum_{i=1}^n x_i$

### 1.6 Median, quantitativ

$R: median(x)$   
Liegt in der Mitte der sortierten Daten  $x_i$ .  
Unempfindlich gegenüber Ausreißern.

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

### Streuungsmaße

### 1.7 Stichprobenvarianz $s^2$

$R: var(x)$   
Verschiebungssatz:  
 $s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{n-1} (\sum_{i=1}^n x_i^2 - n\bar{x}^2)$  Gemittelte Summe der quadratischen Abweichung vom Mittelwert

### 1.8 Stichpr. standardabw.

$R: sd(x)$   
 $s = \sqrt{s^2}$  Streuungsmaß mit gleicher Einheit wie beobachteten Daten  $x_i$ .  $\bar{x}$  minimiert die "quadratische Verlustfunktion" oder die Varianz gibt das Minimum der Fehlerquadrate an.

### 1.9 p-Quantile

$R: quantile(x, p)$ . Teilt die sortierten Daten  $x_i$  ca. im Verhältnis  $p$ :  $(1-p)$  d.h.

$\hat{F}(x_p) \approx p$ ;  $\hat{F}$   $\approx$  kummulierte Häufigkeit;  
1. Quartil = 0.25-Quantil; Median = 0.5-Quantil; 3. Quartil = 0.75-Quantil;

$$x_p \begin{cases} x_{\lfloor np \rfloor + 1}, & np \in \mathbb{N} \\ \frac{1}{2}(x_{np} + x_{np+1}), & np \notin \mathbb{N} \end{cases}$$

### 1.10 Boxplot

Interquartilsabstand  $I = x_{0.75} - x_{0.25}$ . Innerhalb der Box 50% aller Stichproben;  $1/4$  je zu  $I_{min}$  & zu  $I_{max}$  Whiskers zeigen die Spannweite =  $\max x_i - \min x_i$

### 1.11 Chebyshev

$\frac{N(S_k)}{n} > 1 - \frac{1}{k^2}$ , für alle  $k \geq 1$   $\bar{x}$  der Durchschnitt,  $s > 0$  die Stichproben-Standardabweichung von Beobachtungswerten  $x_1, \dots, x_n$ . Sei  $S_k = \{i, 1 \leq i \leq n : |x_i - \bar{x}| < k \cdot s\}$ ; Für eine beliebige Zahl  $k \geq 1$  liegen mehr als  $100 \cdot (1 - \frac{1}{k^2})$  Prozent der Daten im Intervall von  $\bar{x} - ks$  bis  $\bar{x} + ks$ . **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 Formulierungen:**  $\bar{S}_k = \{i : |x_i - \bar{x}| \geq k \cdot s\}$ ;  $\frac{N(\bar{S}_k)}{n} \leq \frac{1}{k^2}$ ;

Die Ungleichheit liefert nur eine **sehr grobe Abschätzung**, ist aber unabhängig von der Verteilung der Daten. **Empirische Regeln** 68% der Daten im Bereich um  $\bar{x} \pm s$ . 95% um  $\bar{x} \pm 2s$ . 99.7% um  $\bar{x} \pm 3s$ .

### 1.12 Korrelation

Grafische Zusammenhang zwischen multivariaten Daten  $x$  und  $y$  durch ein Streudiagramm. Kennzahlen zur Untersuchung des Zusammenhangs:

### 1.13 Empirische Kovarianz

$R: cov(x, y)$ ;  $s_{xy} = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = \frac{1}{n-1} (\sum_{i=1}^n x_i y_i - n\bar{x}\bar{y})$ ;  $S_{xy} > 0$  steigend;  $S_{xy} < 0$  fallend;

### 1.14 Empir. Korrelkoeff. $r$

$R: cor(x, y)$ ;  $r = \frac{s_{xy}}{s_x s_y}$ ; 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).

### 1.15 Regressionsgerade $y$

$y = mx + t$  mit  $m = r \cdot \frac{s_y}{s_x}$  und  $t = \bar{y} - m \cdot \bar{x}$ ; Für den Bereich  $|\pm 0,7|$  bis  $\pm 1 \Rightarrow$  linearer Zusammenhang.

## 2 Wahrscheinlichkeitsrechnung

### 2.1 Begriffe

**Ergebnisraum  $\Omega$ :** Menge aller möglichen Ergebnisse eines Experiments  
**Elementarereignis  $\omega \in \Omega$ :** einzelnes Element von  $\Omega$

$Eignis E \subseteq \Omega$ : beliebige Teilmenge des Ergebnisraums  $\Omega$  heißt sicheres Ereignis,  $\emptyset$  heißt unmögliches Ereignis  
**Vereinigung  $E \cup F$ :** Ereignis  $E$  oder Ereignis  $F$  treten ein.  $\bigcup_{i=1}^n E_i$ : mindestens ein Ereignis  $E_i$  tritt ein.

**Schnitt  $E \cap F$ :** Ereignis  $E$  und Ereignis  $F$  treten ein.  
 $\bigcap_{i=1}^n E_i$  alle Ereignisse  $E_i$  treten ein. **Generereignis  $\bar{E} = \Omega \setminus E$ :** Ereignis  $E$  tritt nicht ein (Komplement von  $E$ )

**Disjunkte Ereignisse  $E$  und  $F$ :**  $E \cap F = \emptyset$

### 2.2 De Morgan'schen Regeln

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

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

### 2.3 Wahrscheinlichkeit

$0 \leq P(E) \leq 1$ ;  $P(\Omega) = 1$ ;  
 $P(\bigcup_{i=1}^n E_i) = \sum_{i=1}^n P(E_i)$ , falls  $E_i \cap E_j = \emptyset$  für  $i \neq j$

### 2.4 Satz 2.1

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

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

(Übungsaufgabe!!! Ergänzen)

### 2.5 Laplace-Experiment

Zufallsexperimente mit  $n$  gleich wahrscheinlichen Elementarereignissen. Dann berechnet sich die Wahrscheinlichkeit  $P(E)$  für  $E \subseteq \Omega$  aus:

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

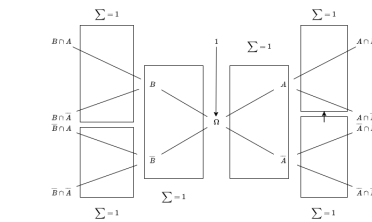
### 2.6 Bedingte Wahrscheinlichkeit

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

### 2.7 Satz 2.2

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

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

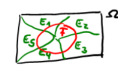


### 2.8 Satz der totalen Wahrscheinlichkeit

Sei  $\Omega = \bigcup_{i=1}^n E_i$  mit  $E_i \cap E_j = \emptyset$  für  $i \neq j$  d.h. die Ereignisse bilden eine disjunkte Zerlegung bzw. eine Partition von  $\Omega$ . Somit gilt:

$$P(F) = \sum_{i=1}^n P(F \cap E_i) = \sum_{i=1}^n P(F|E_i) \cdot P(E_i)$$

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



### 2.9 Vierfeldertafel

$$P(F) = P(F \cap E) + P(F \cap \bar{E})$$

$$P(\bar{F}) = P(\bar{F} \cap E) + P(\bar{F} \cap \bar{E})$$

	E	$\bar{E}$	
F	$P(F \cap E)$	$P(F \cap \bar{E})$	$P(F)$
$\bar{F}$	$P(\bar{F} \cap E)$	$P(\bar{F} \cap \bar{E})$	$P(\bar{F})$
	$P(E)$	$P(\bar{E})$	1

**Satz 2.2 oben:**  $P(E \cap F) = P(E) \cdot P(F|E) = P(F) \cdot P(E|F)$  **Tafel**  
 $= P(F) - P(F \cap \bar{E}) = P(E) - P(\bar{F} \cap E)$ ;  $P(\bar{F}|E) = 1 - P(F|E)$

### 2.10 Formel von Bayes

Hilfreich, wenn man  $P(F|E_i)$  kennt, aber nicht  $P(E_k|F)$  **Satz 2.4**  $P(E_k|F) = \frac{P(F|E_k) \cdot P(E_k)}{\sum_{i=1}^n P(F|E_i) \cdot P(E_i)}$

$$\sum_{i=1}^n P(F|E_i) \cdot P(E_i)$$

**Nur Nenner!**  $P(F)$  aus dem Satz der totalen Wahrscheinlichkeit.

### 2.11 Stochastische Unabhängigkeit

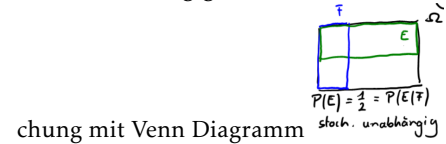
**Übung** Die Ereignisse  $E$  und  $F$  heißen (stochastisch) unabhängig, wenn die Information über das Eintreten des einen Ereignisses die Wahrscheinlichkeit für das Eintreten des anderen Ereignisses nicht ändert, d.h. falls  $P(E|F) = P(E)$  or  $P(E \cap F) = P(E) \cdot P(F)$

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

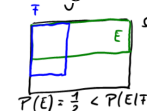
**Es gilt** Falls die Ereignisse  $E, F$  unabhängig sind, dann sind auch:  $\circ E, \bar{F}$ ;  $\circ \bar{E}, F$ ;  $\circ \bar{E}, \bar{F}$  unabhängig

### Bemerkung

$\circ$  Stochastische Unabhängigkeit bedeutet nicht notwendigerweise eine kausale Abhängigkeit;  $\circ$  Veranschaulichung mit Venn Diagramm



chung mit Venn Diagramm



$\circ A, B \neq \emptyset$  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  $P(B) > 0$   
 $\Rightarrow A, B$  stochastisch abhängig

### 3 Zufallsvariable

Abbildung des **abstrakten** Ergebnisraums  $\Omega$  auf  $\mathbb{R}$ . Eine Abbildung  $X: \Omega \rightarrow \mathbb{R}$ ,  $\omega \mapsto X(\omega)$  heißt Zufallsvariable (ZV).  $x \in \mathbb{R}$  heißt Realisation der ZV  $X$ .

$\circ$  Diskrete ZV:  $X(\Omega) = x_1, \dots, x_n (n \in \mathbb{N})$ ;  
z.B.  $X$  = "Augensumme beim Würfeln"

$\circ$  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  $\mathbb{R}$  wird zurückgeführt auf die Wahrscheinlichkeit der entsprechenden Ereignisse in  $\Omega$ . Für jedes  $X \in \mathbb{R}$  ist die Verteilungsfunktion  $F: \mathbb{R} \rightarrow [0, 1]$  einer ZV  $X$  definiert durch:

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

$$\circ 0 \leq F(x) \leq 1$$

$$\circ \lim_{x \rightarrow -\infty} F(x) = 0 \quad \lim_{x \rightarrow \infty} F(x) = 1$$

$$\circ \text{monoton wachsend}$$

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

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

### 3.2 Diskrete ZVs

Für eine diskrete ZV  $X$  mit  $X(\Omega) = x_1, \dots, x_n$  ( $n$  endlich oder abzählbar unendlich) ist die Wahrscheinlichkeitsfunktion definiert durch:

$$p(x) = \begin{cases} P(X = x_i), & \text{falls } x_i \in X(\Omega) \\ 0, & \text{sonst} \end{cases} \quad (1)$$

**Es gilt:**

$$\circ F(x) = P(X \leq x) = \sum_{x_i \leq x} p(x_i)$$

$\circ F(x)$  ist eine rechtseitig stetige **Treppenfunktion** mit **Sprüngen** bei der Realisation von  $x_i$ .

### 3.3 Stetige ZVs

Stetige ZV  $X$  ist die Wahrscheinlichkeitsdichte  $f: \mathbb{R} \rightarrow [0, \infty]$  definiert durch

$$P(a < X < b) = \int_a^b f(x) dx$$

**Es gilt:**

$$\circ F(x) = P(X \leq x) = \int_{-\infty}^x f(t) dt \text{ und } F'(x) = f(x)$$

$\circ F(x)$  ist stetig &  $P(a < X \leq b) = P(a \leq X \leq b)$  wegen  $P(X = a) = 0$

### 3.4 Verteilungsfunktion

**Untergrenze** Es wird normal mit - integriert.

### 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.

$\circ$  Verteilungsfunktion  $F(x)$  ist rechtsseitig stetige

**Treppenfunktion. Sprunghöhen:**  $P(X = x_i) = F(x_i) - \lim_{x \rightarrow x_i^-} F(x) \neq 0$

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

### 3.7 Stetige ZV

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

$\circ$  Verteilungsfunktion  $F(x)$  ist stetig mit  $F'(x) = f(x)$ ;  $P(X = x_i) = 0$

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

### 3.8 Erwartungswert

Der Erwartungswert  $E[X] = \mu$  einer ZV  $X$  ist der **Schwerpunkt** ihrer Verteilung oder der durchschnittliche zu erwartende Wert der ZV.

◦ diskrete ZV:  $E[X] = \sum_{i=1}^n x_i \cdot p(x_i)$

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

ZV ist konstant.  $E[X]$  verhält sich linear.

Eigenschaften von  $E[X]$ :

◦  $E[b] = b$

◦  $E[aX + b] = aE[X] + b$

◦  $E[X_1 + \dots + X_n] = \sum_{i=1}^n E[X_i]$

◦  $\sum_{i=1}^n x_i$

### 3.9 Satz 3.1

Sei  $Y = g(X)$  eine Funktion der ZV  $X$ . Dann gilt:

◦ für diskrete ZV:  $E[g(X)] = \sum_{i=1}^n g(x_i) \cdot p(x_i)$

◦ für stetige ZV:  $E[g(X)] = \int_{-\infty}^{\infty} g(x) \cdot f(x) dx$ . Das vertauschen von  $E$  und  $g$  nur bei **linearen** Funktionen möglich.  $\Rightarrow g(E[X])$

### 3.10 Varianz

Die Varianz einer ZV  $X$  mit  $\mu$  ist ein quadratisches Streuungsmaß.  $\sigma^2 = \text{Var}[X] =$

$$E[(X - \mu)^2] \stackrel{\text{falls } x \text{ stetig}}{=} \int_{-\infty}^{\infty} (x - \mu)^2 \cdot f(x)$$

$g(X)$

Die Standardabweichung  $\sigma = \sqrt{\text{Var}[X]}$  hat im Gegensatz zur Varianz die gleiche Dimension von der ZV  $X$ .

◦  $\text{Var}[b] = 0$

◦  $\text{Var}[aX + b] = a^2 \text{Var}[X]$

### 3.11 Satz 3.2

$\text{Var}[X] = E[X^2] - (E[X])^2$  Beim Minuend wird beim Erwartungswert nur das einfache stehende  $x$  quadriert **nicht**  $f(x)$ !!!

### 3.12 Z-Transformation, Standardisierung

Sei  $X$  eine ZV mit  $\mu$  und  $\sigma$ . Dann ist

$$Z = \frac{X - \mu}{\sigma} = \frac{x}{\sigma} - \frac{\mu(\text{konstant})}{\sigma}$$

### 3.13 Kovarianz

Eigenschaften:

◦  $\text{Cov}[X, Y] = \text{Cov}[Y, X]$

◦  $\text{Cov}[X, X] = \text{Var}[X]$

◦  $\text{Cov}[aX, Y] = a \text{Cov}[X, Y]$

Die Kovarianz zweier ZV ( $X, Y$ ) ist definiert durch  $\text{Cov}[X, Y] = E[(X - E[X])(Y - E[Y])]$ ; Die Kovarianz beschreibt die Abhängigkeit zweier ZV  $X$  und  $Y$ . Je stärker diese Korrelieren, desto (betragsmäßig) größer ist die Kovarianz. Falls  $X, Y$  (stochastisch) unabhängig  $\Rightarrow \text{Cov}[X, Y] = 0$

### 3.14 Satz 3.3

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

### 3.15 Varianz einer Summe von ZV

$$\begin{aligned} \circ \text{Var}[X_1 + \dots + X_n] &= \sum_{i=1}^n \sum_{j=1}^n \text{Cov}[X_i, X_j]; \text{Var}[X_1 + X_2] = \\ &\text{Var}[X_1] + \text{Var}[X_2] + 2\text{Cov}[X_1, X_2] \\ \circ \text{Falls } X_i, X_j \text{ paarweise unabhängig !!!} &: \\ \text{Var}[X_1 + \dots + X_n] &= \sum_{i=1}^n \text{Var}[X_i] \end{aligned}$$

### 3.16 Overview $\mu \sigma$

#### 3.17 $E[X]$

$$E[aX + b] = aE[X] + b; E[X_1 + \dots + X_n] = \sum_{i=1}^n E[X_i]$$

Falls  $X_1, X_2$  unabhängig:

$$E[X_i] = \mu \Rightarrow E[\bar{X}] = E\left[\frac{1}{n}(X_1 + \dots + X_n)\right] = \frac{1}{n} \sum_{i=1}^n E[x_i] = \frac{1}{n} \cdot n \cdot \mu = \mu;$$

$\mu;$

#### 3.18 Varianz

$$\text{Var}[aX + b] = a^2 \text{Var}[X]$$

Falls  $X_i, X_j$  paarweise unabhängig:

$$\begin{aligned} \text{Var}[X_1 + \dots + X_n] &= \sum_{i=1}^n \text{Var}[X_i] \\ \text{Var}[X_i] &= \sigma^2 \Rightarrow \text{Var}[\bar{X}] = \text{Var}\left[\frac{1}{n}(x_1 + \dots + x_n)\right] = \\ &\frac{1}{n^2} \sum_{i=1}^n \text{Var}[X_i] = \frac{1}{n^2} \cdot n \cdot \sigma^2 = \frac{\sigma^2}{n} \end{aligned}$$

### 3.19 Quantile

Sei  $X$  eine ZV mit Verteilungsfunktion  $F(x)$  und  $0 < p < 1$ . Dann ist das  $p$ -Quantil definiert als der Wert  $x_p \in \mathbb{R}$  für den gilt:

$F(x_p) \geq p$ .  $p$ -Quantil einer stetigen ZV mit **streng monoton wachsenden**

$F(x): x_p = F^{-1}(p)$  d. h. umkehrbar. Zuerst  $p$  dann  $e^{xp}$

### 4 Spezielle Verteilung

#### 4.1 Diskrete Verteilung

#### 4.2 Bernouilliverteilung

Indikatorvariable mit den Werten 1 bei Erfolg und 0 bei Misserfolg; **Wahrscheinlichkeit**:  $P(X = 1) = p, P(X = 0) = 1 - p$ ; **Verteilung**:  $X \sim B_{1,p}$   $p$  ist Erfolgswahrscheinlichkeit;  $E[X] = p = \sum x_i \cdot p(x_i) = 1 \cdot p(1); \text{Var}[X] = p(1-p) = E[X^2] - (E[X])^2 = p - p^2 = p(1-p)$ ;

#### 4.3 Binominalverteilung

Anzahl der Erfolge beim  $n$ -maligen Ziehen mit Zurücklegen; **Wahrscheinlichkeit**  $P(x = k) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}, k \in \{0, 1, \dots, n\}$ ; **Verteilung**  $X \sim B_{n,p}$ ;  $E[X] = np$ ;  $\text{Var}[X] = np(1-p)$ ; **R**:  $\text{dbinom}(k, n, p) = P(X=k) \hat{=}$  Wahrscheinlichkeits-/Dichtefunktion;  $\text{pbinom}(k, n, p) = F(k) \hat{=}$  Verteilungsfunktion;  $\text{qbinom}(q, n, p) \hat{=}$   $q$ -Quantil;  $\text{rbinom}(q, n, p) \hat{=}$   $q$ -binomialverteilte Zufallszahlen;

### 4.4 Hypergeometrische Verteilung

Anzahl der Erfolge beim  $n$ -maligen Ziehen ohne Zurücklegen aus einer Menge mit  $M$  Elementen, die Erfolg bedeuten, und  $N$  Elementen, die Misserfolg bedeuten. **Gesamtumfang**  $= M + N$ ; **Wahrscheinlichkeit**  $P(X = k) = \frac{\binom{M}{k} \cdot \binom{N-k}{n-k}}{\binom{M+N}{n}}, k \in \{0, 1, \dots, \min\{n, M\}\}$ ; **Verteilung**  $X \sim H_{M, N, n}$ ;  $E[X] = n \frac{M}{M+N}$ ;

$\frac{M}{M+N} \hat{=}$  Trefferwahrscheinlichkeit;  $\text{Var}[X] = n \frac{M}{M+N} (1 - \frac{M}{M+N}) \frac{M+N-n}{M+N-1}$ ;  $\rightarrow 1$  falls  $n$  klein im Verhältnis zu  $M+N$ ; **R**:  $\text{dhyper}(k, M, N, n) = P(X = k)$ ;  $\text{phyper}(k, M, N, n) = F(k)$ ;

### 4.5 Poisson-Verteilung

**Verteilung der seltenen Ereignisse** Häufigkeit punktförmiger Ereignisse in einem Kontinuum. Die durchschnittlich zu erwartende Anzahl der Erfolge  $\lambda$  pro Maßeinheit (i. a. Zeiteinheit) sei bekannt.  $k \in \mathbb{N}_0 \rightarrow$  **diskret Wahrscheinlichkeit**  $P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$  mit  $\sum_{k=0}^{\infty} P(X = k) = 1$ , da  $\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{\lambda}$ ; **Verteilung**  $X \sim P_{\lambda}$ ;  $E[X] = \lambda$ , da  $\sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=1}^{\infty} \lambda \frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda} \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} = \lambda$ ;  $\text{Var}[X] = \lambda$  **R**:  $\text{dpois}(k, \lambda) = P(X = k)$ ;  $\text{ppois}(k, \lambda) = F(k)$ ;

### 4.6 Gleichverteilung

Alle Werte  $\{x_1, \dots, x_n\}$  einer ZV  $X$  sind gleich wahrscheinlich; **Wahrscheinlichkeit**  $P(X = x_k) = \frac{1}{n}$ ; **Verteilung**  $X \sim U_{\{x_1, \dots, x_n\}}$ ;  $E[X] = \frac{1}{n} \sum_{k=1}^n x_k = \bar{x}$ ;  $\text{Var}[X] = \frac{1}{n} \sum_{k=1}^n x_k^2 - \bar{x}^2$ ; **R**:  $\text{sample}(1 : N, n) \hat{=}$   $n$  Zufallszahlen zwischen 1 und  $N$

### 4.7 Gleichverteilung

**4.8 Stetige Gleichverteilung/Rechteck** Zufallszahlen aus einem Intervall  $[a, b]$ ; **Dichte**:  $f(x) = \frac{1}{b-a}$  für  $x \in [a, b]$ ; **Verteilung**:  $X \sim U_{[a, b]}$ ;  $E[X] = \frac{a+b}{2}$ ;

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

### 4.9 Normalverteilung

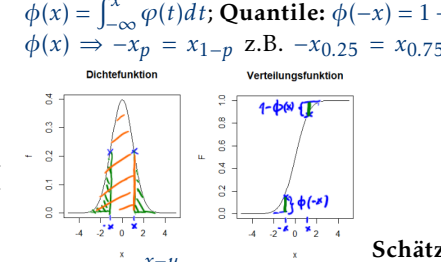
Beschreibt viele reale Situationen, ist insbesondere Grenzverteilung unabhängiger Summen; **Dichte**:  $f(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}$ ; **Verteilung**:  $X \sim N_{\mu, \sigma^2}$ ;  $E[X] = \mu$ ;  $\text{Var}[X] = \sigma^2$ ; **R**:  $\text{dnorm}(x, \mu, \sigma) = f(x)$ ;  $\text{pnorm}(x, \mu, \sigma) = F(x)$ ;  $\text{qnorm}(q, \mu, \sigma) : q$ -Quantil; **Maximalstelle** von  $f(x)$  bei  $x = \mu$ ; **Wende-**

stelle von  $f(x)$  bei  $x = \mu \pm \sigma$ ;  $E[aX + b] = aE[X] + b$ ;  $\text{Var}[aX + b] = a^2 \text{Var}[X]$ ;  $X \sim N_{\mu, \sigma^2} \Rightarrow aX + b \sim N_{a\mu+b, a^2\sigma^2}$  und  $\frac{X-\mu}{\sigma} \sim N_{0,1}$ ;  $X_1 \sim N_{\mu_1, \sigma_1^2}$  und  $X_2 \sim N_{\mu_2, \sigma_2^2} \Rightarrow X_1 + X_2 \sim N_{\mu_1+\mu_2, \sigma_1^2+\sigma_2^2}$ ;

$X_1, X_2$  stochastisch unabhängig

### 4.10 Standardnormalverteilung

**Dichte**:  $\varphi(x) = \frac{1}{\sqrt{2}} e^{-\frac{1}{2}x^2}$ ; **Verteilung**  $\Phi(x) = \int_{-\infty}^x \varphi(t) dt$ ; **Quantile**:  $\Phi(-x) = 1 - \Phi(x) \Rightarrow -x_p = x_{1-p}$  z.B.  $-x_{0.25} = x_{0.75}$ ;



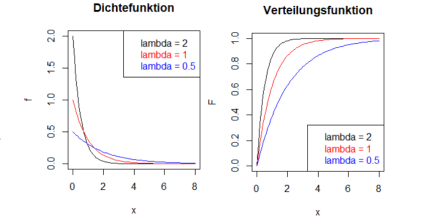
**werte**:  $Z = \frac{x-\mu}{\sigma} \sim N_{0,1}$ ;

$P(\mu - \sigma \leq X \leq \mu + \sigma) = P(-1 \leq Z \leq 1) \approx 68\%$ ;  
 $P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) = P(-2 \leq Z \leq 2) \approx 95\%$ ;  
 $P(\mu - 3\sigma \leq X \leq \mu + 3\sigma) = P(-3 \leq Z \leq 3) \approx 99.7\%$

### 4.11 Exponentialverteilung

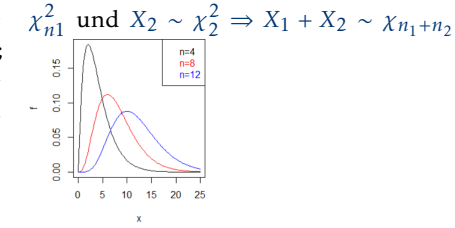
Modellierung von Lebensdauern, Wartezeiten Sei  $Y_t \sim P_{\lambda t}$  im Intervall  $[0, t]$  von  $t$  Zeiteinheiten, dann beschreibt die Exponentialverteilung die Wartezeit  $X$  bis zum Eintreten eines Ereignisses; **Dichte- und Verteilungsfunktion**:  $f(x) = \lambda e^{-\lambda x} (x \geq 0)$  und  $F(x) = 1 - e^{-\lambda x}$ ; **Verteilung**:  $X \sim \text{Exp}_{\lambda}$ ;  $E[X] = \frac{1}{\lambda} \Rightarrow$  Berechnung mit partieller Integration;  $\text{Var}[X] = \frac{1}{\lambda^2}$ ; **R**:  $\text{dexp}(x, \lambda) = f(x)$ ;

$\text{pexp}(x, \lambda) = F(x)$ ; **Eigenschaft**: Eine exponentialverteilte ZV  $X$  ist gedächtnislos, d.h.  $P(X > s + t) | X > t = P(X > s)$ ;



### 4.12 Chiquadrat-Verteilung

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



### 4.13 t-Verteilung

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

verteilt mit  $n$  Freiheitsgraden; **Anwendungsmodell**: Schätz- und Testverfahren bei unbekannter Varianz; **Verteilung**:  $Y \sim t_n$ ;  $E[Y] = 0$  für  $n > 1$ ;  $\text{Var}[Y] = \frac{n}{n-2}$  für  $n > 2$ ; **R**:  $\text{dt}(y, n) \hat{=}$   $f(x)$ ;  $\text{pt}(y, n) \hat{=}$   $F(x)$ ; **Eigenschaften**: Für  $n \rightarrow \infty : t_n \rightarrow N_{0,1}$ ; Achsensymmetrie der Dichtefunktion  $\Rightarrow -y_p = x_{1-p}$

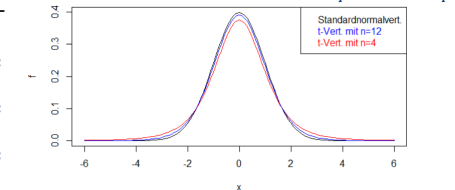


Abbildung Dichtefunktion

### 5 Zentraler Grenzwertsatz

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

#### 5.1 ZGWS

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

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

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

Der Satz gilt sogar allgemeiner, wenn die  $X_i$  abhängig und nicht identisch verteilt sind, vorausgesetzt kein  $X_i$  ist 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  $\bar{X}$  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);  $n \leq 15$ : falls die unbekannte Verteilung annähernd normalverteilt ist;



**5.2  $\phi$**   
  
 $\phi(-a) = 1 - \phi(a)$ ;  $\phi(a) = 1 - \phi(-a)$ ;  $P(-a < Z < a) = \phi(a) - \phi(-a) = \phi(a) - (1 - \phi(a)) = 2\phi(a) - 1$  or  $1 - \phi(-a) - \phi(-a) = 1 - 2\phi(-a)$

**5.3  $\phi^{-1}$**   
  
 $-qnorm(p) = qnorm(1 - p)$   $\Leftrightarrow$   
 $-\phi^{-1}(p) = \phi^{-1}(1 - p)$  **Zusammenhang**

$\phi^{-1}(p) = x_p$ ;  
**Aufgabentypen:** Seien  $X_i$  i.i.d. ZV mit  $\mu$  und  $\sigma^2$ , aber unbekannter Verteilung.

Dann sind  $Z_1 = \frac{\sum X_i - n\mu}{\sqrt{n}\sigma}$  und  $Z_2 = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$  näherungsweise standardnormalverteilt.

- Es lassen sich Wahrscheinlichkeiten für  $\sum X_i, \bar{X}, Z_1$  oder  $Z_2$  berechnen.
- Es lässt sich  $n$  bestimmen, so dass, zu vorgegebener Schranke  $k$  und Wahrscheinlichkeit  $p$  gilt:  $P(Z_i > k) \geq p$  or  $P(-k \leq Z_i \leq k) \geq p$

**5.4 Stichprobenvert.normalvert. Grundgesamt.**

**5.5 Stichprobenmittel**

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

**5.6 Stichprobenvarianz**

Die Stichprobenfunktion  $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 = \frac{1}{n-1} (\sum_{i=1}^n X_i^2 - n\bar{X}^2)$  ist eine erwartungstreue Schätzfunktion für die Varianz  $\sigma^2$ , d. h.  $E[S^2] = \sigma^2$ ;  $E[\bar{X}] = E[\frac{1}{n} \sum X_i] = \frac{1}{n} E[\sum X_i] = \frac{1}{n} n\mu = \mu$ ;  $Var[\bar{X}] = Var[\frac{1}{n} \sum X_i] = \frac{1}{n^2} Var[\sum X_i] = \frac{1}{n^2} n\sigma^2 = \frac{\sigma^2}{n}$ ; Seien  $X_i (i = 1, \dots, n)$  unabhängige normalverteilte ZV mit Erwartungswert  $\mu$  und Varianz  $\sigma^2$ . Dann gilt:

**bei bekannter Varianz:**  $\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N_{0,1}$ ;

$\frac{(n-1)S^2}{\sigma^2} \Rightarrow \text{Standardisierung} \sim \chi^2_{n-1}$ ; Bei unbekannter Varianz:  $\frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} \sim t_{n-1}$ ;

**6 Konfidenzintervall**

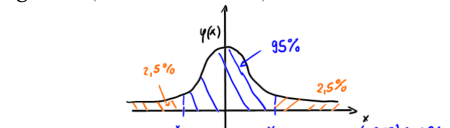
**6.1 Begriffe**  
Irrtumswahrscheinlichkeit =  $\alpha$ ; Konfidenzniveau =  $1 - \alpha$ ; Konfidenzintervall =  $I$

**6.2 Punktschätzer**  
 $E[X]$ : Stichprobenmittel:  $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ ;

Varianz: Stichprobenvarianz:  $s^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ ; Schätzwert für wahren Parameter, aber keine Aussage über Unsicherheit der Schätzung, Geringe Sicherheit für wahren Parameter;

**6.3 Intervallschätzer**

Intervall für wahren Parameter, mit vorgegebener Sicherheit; Vorgabe (95% or 99%); Dichtefunktion:



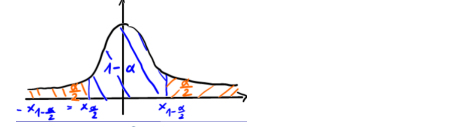
on:  $P(-a \leq \bar{x} \leq a) > 0.95$ ;  $\sigma$  ist unbekannter Parameter

$P(x_{0.025} < \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} < x_{0.975}) \geq 0.95$   
 $-1.96; N_{0,1}; 1.96$ ;

**6.4  $\mu$ , unbekannt,  $\sigma^2$ , bekannt**

$I = ]\bar{X} - \phi^{-1}(1 - \frac{\alpha}{2}) \frac{\sigma}{\sqrt{n}}, \bar{X} + \phi^{-1}(1 - \frac{\alpha}{2}) \frac{\sigma}{\sqrt{n}}[$   
 $qnorm(1 - \frac{\alpha}{2})$

$1 - \alpha$	$\frac{\alpha}{2}$	$\phi^{-1}(1 - \frac{\alpha}{2})$
90%	5%	$\phi^{-1}(0.95) \approx 1.645$
95%	2.5%	$\phi^{-1}(0.975) \approx 1.96$
99%	0.5%	$\phi^{-1}(0.995) \approx 2.576$



**6.5  $\mu$  &  $\sigma^2$ , unbekannt**  
 $I = ]\bar{X} - t_{n-1}^{-1}(1 - \frac{\alpha}{2}) \frac{S}{\sqrt{n}}, \bar{X} + t_{n-1}^{-1}(1 - \frac{\alpha}{2}) \frac{S}{\sqrt{n}}[$

**6.6 Zusammenfassung**  
Wie verändert sich das  $(1 - \alpha)$ -Konfidenzintervall, n größer  $\Rightarrow$  I kürzer;  $1 - \alpha$  größer  $\Rightarrow$  I länger; Für  $\frac{L}{2} = 2\phi^{-1}(1 - \frac{\alpha}{2}) \frac{\sigma}{\sqrt{n}} \frac{1}{2} = 2\phi^{-1}(1 - \frac{\alpha}{2}) \frac{\sigma}{\sqrt{4n}}$

**6.7 Aufgabentypen**

**Geg:** n,  $1 - \alpha$ ; **Ges:** I s.o. **Geg:**  $\bar{X}, \sigma, 1 - \alpha, L$ ;  $L = 2\phi^{-1}(1 - \frac{\alpha}{2}) \frac{\sigma}{\sqrt{n}}$ ; **Ges:** n;  $\sqrt{n} > 2\phi^{-1}(1 - \frac{\alpha}{2}) \frac{\sigma}{L}$  **Geg:** n, I, L; **Ges:**  $1 - \alpha$ ;  $1 - \frac{\alpha}{2} =$

$\phi(\frac{L\sqrt{n}}{2\sigma})$

**7 Hypothesentests**

Basierend auf n unabhängig und identisch Verteilte (i.i.d) Zufallsvariablen  $X_1, \dots, X_n$  (Messungen) soll eine Entscheidung getroffen werden, ob eine Hypothese für einen unbekannten Erwartungswert  $\mu$  gültig ist or nicht.

**7.1 Def**

$\alpha$  = Signifikanzniveau/ Fehlerwahrscheinlichkeit TG = Prüfgröße; TG\* = standardisierte Prüfgröße; signifikante Schlussfolgerung =  $H_0$  verworfen  $\rightarrow$  klassischer Parametertest; schwache Schlussfolgerung =  $H_0$  wird nicht verworfen  $\rightarrow$  klassischer Parametertest. p-Wert = beobachtetes Signifikanzniveau

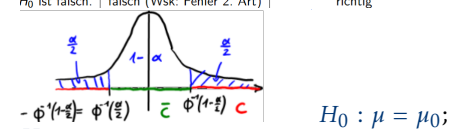
**7.2 Null- und Gegenhypothese**

**Modell:** Verteilung der Grundgesamtheit or Testgröße TG (häufig  $\bar{x}$ ) ist bekannt bis auf einen Parameter, z.B.  $\mu$ , für den eine Hypothese aufgestellt wird.  $TG \sim N_{\mu, \sigma^2}$ ; **Nullhypothese:**  $H_0$ : Angezweifelte Aussage, der widersprochen werden kann, wenn die Stichprobe einen Gegenbeweis liefert.  $H_0 : \mu = \mu_0$ ; **Gegenhypothese**  $H_1$ : Gegenteil von  $H_0$  z.B.  $H_1 \neq \mu_0$ ;

**7.3 Ablehnungsbereich, Fehler 1. & 2.**

Treffen der Testentscheidung, basierend auf einer konkreten Stichprobe  $\{x_1, \dots, x_n\}$ ; Berechnung der Realisation  $tg = TG(x_1, \dots, x_n)$  der Prüfgröße TG; **Ablehnungsbereich / Kritischer Bereich C:** Werte der Testgröße, die für  $H_1$ , sprechen & 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  $\bar{C}$  des Ablehnungsbereichs.  $H_0$  kann nicht abgelehnt werden, falls  $tg \in \bar{C} (P(tg \in \bar{C}) \geq 1 - \alpha)$ . **Fehler 2. Art:** Die Wahrscheinlichkeit, dass  $H_0$  nicht abgelehnt wird, obwohl sie falsch ist.

Realität	Testentscheidung $H_0$ wird (nicht) abgelehnt	$H_0$ wird abgelehnt.
$H_0$ ist wahr.	richtig	falsch (Wsk: Fehler 1. Art)
$H_0$ ist falsch.	falsch (Wsk: Fehler 2. Art)	richtig



$H_0 : \mu = \mu_0$ ;

$H_1 : \mu \neq \mu_0$ ;

**7.4 Klassischer Parametertest**

$H_0$  wird abgelehnt, falls  $tg = TG(x_1, \dots, x_n) \in C$ ;  $H_0$  wird angenommen falls  $tg = TG(x_1, \dots, x_n) \in \bar{C}$ ; Der kritische Bereich ergibt sich analog zu den Konfidenzintervallen durch die Vorgabe eines kleinen Signifikanzniveau  $\alpha$  d.h. max. Wahrscheinlichkeit für

Fehler 1. Art: mit standardisierter Prüfgröße TG\* gilt:  $P(TG \in C) \leq \alpha \Leftrightarrow TG^* \in ]-\infty; \phi^{-1}(1 - \frac{\alpha}{2})] \cup [\phi^{-1}(1 - \frac{\alpha}{2}); \infty[$ ;  $P(TG \in \bar{C}) \geq 1 - \alpha \Leftrightarrow TG^* \in [\phi^{-1}(\frac{\alpha}{2}), \phi^{-1}(1 - \frac{\alpha}{2})]$ ; Wird dann  $H_0$  verworfen, spricht man von einer signifikanten Schlussfolgerung. Kann  $H_0$  nicht verworfen werden, dann lässt sich keine Aussage über den Fehler 2. Art treffen & man spricht von einer schwachen Schlussfolgerung.

**7.5 Zweiseitiger Gauß Test**

$H_0 : \mu = \mu_0$  gegen  $H_1 : \mu \neq \mu_0$ ;  $\bar{X} \sim N_{\mu_0, \sigma_0^2/n} \Rightarrow \frac{\bar{X} - \mu_0}{\frac{\sigma_0}{\sqrt{n}}} \sim N_{0,1}$ ;  $P_{\mu_0}(\bar{X} \in C) \leq \alpha \Leftrightarrow |TG| = \frac{|\bar{X} - \mu_0|}{\frac{\sigma_0}{\sqrt{n}}} \sqrt{n} > \phi^{-1}(1 - \frac{\alpha}{2})$ ; **Testentscheidung:**  $H_0$  wird abgelehnt, falls  $|TG| > \phi^{-1}(1 - \frac{\alpha}{2})$ ;  $H_0$  wird angenommen, falls  $|TG| \leq \phi^{-1}(1 - \frac{\alpha}{2})$

**7.6 Einseitiger Gauß Test**

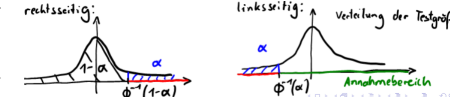
**7.7 linksseitig**

$H_0 : \mu \geq \mu_0$  gegen  $H_1 : \mu < \mu_0$

**7.8 rechtsseitig**

$H_0 : \mu \leq \mu_0$  gegen  $H_1 : \mu > \mu_0$

$P_{\mu_0}(\bar{X} \in C) \leq \alpha \Leftrightarrow TG = \frac{\bar{X} - \mu_0}{\frac{\sigma_0}{\sqrt{n}}} \sqrt{n} < \phi^{-1}(\alpha)$ ; **Testentscheidung:**  $H_0$  wird abgelehnt falls,  $TG < \phi^{-1}(\alpha)$ ;  $H_0$  wird angenommen, falls  $TG \geq \phi^{-1}(\alpha)$ ;



**7.9 Varianten Gauß Test,  $\sigma^2$  bekannt,  $\mu$  unbekannt**

Prüfgröße  $tg = \frac{\bar{X} - \mu_0}{\frac{\sigma_0}{\sqrt{n}}} \sqrt{n}$ ;

$H_0$	$H_1$	$H_0$ ablehnen, falls	p-Wert
$\mu = \mu_0$	$\mu \neq \mu_0$	$ tg  > \phi^{-1}(1 - \frac{\alpha}{2})$	$2(1 - \Phi( tg ))$
$\mu \leq \mu_0$	$\mu > \mu_0$	$tg > \phi^{-1}(1 - \alpha)$	$1 - \Phi(tg)$
$\mu \geq \mu_0$	$\mu < \mu_0$	$tg < \phi^{-1}(\alpha)$	$\Phi(tg)$

**7.10 t-Test,  $\mu, \sigma^2$  unbekannt**

Prüfgröße  $tg = \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}} \sqrt{n} \sim t_{n-1}$

$H_0$	$H_1$	$H_0$ ablehnen, falls	p-Wert
$\mu = \mu_0$	$\mu \neq \mu_0$	$ tg  > t_{n-1}^{-1}(1 - \frac{\alpha}{2})$	$2(1 - t_{n-1}(\frac{ tg }{t_{n-1}^{-1}(1 - \frac{\alpha}{2})}))$
$\mu \leq \mu_0$	$\mu > \mu_0$	$tg > t_{n-1}^{-1}(1 - \alpha)$	$1 - t_{n-1}(tg)$
$\mu \geq \mu_0$	$\mu < \mu_0$	$tg < t_{n-1}^{-1}(\alpha)$	$t_{n-1}(tg)$

**7.11 p-Wert**

Wahrscheinlichkeit, bei Zutreffen von  $H_0$  den beobachteten Wert tg der Prüfgröße or einen noch stärker von  $\mu_0$  abweichenden Wert zu bekommen. Der p-Wert zu einer Hypothese  $H_0$  ist der kleinste Wert von  $\alpha$ , für den  $H_0$  noch abgelehnt

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 Testentscheidung treffen; Falls  $p - \text{Wert} < 1\%$  : sehr hohe Signifikanz; Falls  $1\% \leq p - \text{Wert} < 5\%$  : hohe Signifikanz; Falls  $5\% \leq p - \text{Wert} \leq 10\%$  : Signifikanz; Falls  $p - \text{Wert} > 10\%$  : keine Signifikanz

**7.12 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  $H_0$  zum Signifikanzniveau  $\alpha$ ;

**7.13 Zusammenfassung klass. Hypo.test**

Signifikanzniveau  $\alpha$  wird vorgegeben;  $\alpha$  & Verteilung der Testgröße unter  $H_0$  wir der Ablehnungsbereich ermittelt. Je kleiner (größer)  $\alpha$ , desto kleiner (größer) ist der Ablehnungsbereich;  $! : \alpha$  & C hängen nicht von der konkreten Stichprobe ab;  $H_0$  wird abgelehnt, falls der ermittelte Wert der Testgröße (beobachteter Wert) in C liegt.  $! : \text{Die tg hängt von der konkreten Stichprobe ab. Sie ist eine ZV.}$

**7.14 Test mittels p-Wert**

$\alpha$  wird vorgegeben. Berechnung des p-Werts anhand der konkreten Stichprobe mit der Verteilung der Tg unter  $H_0$ ;  $! : \text{Der p-Wert hängt von der konkreten Stichprobe ab, ist eine ZV.}$   $H_0$  wird abgelehnt, falls  $p - \text{Wert} \leq \alpha$ ;

**8 Fehleranalyse**

**8.1 Auslöschung**

wenn ungefähr gleich große, bereits mit Fehlern behaftete Zahlen voneinander abgezogen werden & signifikante Mantissenstellen wörtlich ausgelöscht werden.

**8.2 Addition**

große signifikante Stellen schlucken kleine signifikante Stellen.

**9 Interpolation**

Zu gegebenen Punkten  $(x_i, y_i), i = 0, \dots, n$  mit  $x_i \neq x_j$  für  $i \neq j$  eine Funktion G (dies ist nicht eindeutig! Abhängig von der Funktionsklasse), so dass  $G(x_i) = y_i, i = 0, \dots, n$  (Interpolationsbedingung). Interpolation ist **ungeeignet** für verauschte Daten. Lösung: Approximation der kleinsten Quadrate.

**9.1 Begriffe**

Extrapolation  $\hat{=}$  Näherungswerte für x-Werte außerhalb der Stützstellen; Dividierende Differenzen  $\hat{=}$  Koeffizienten  $c_i$  lassen sich rekursiv durch wie-

derholte Bildung von "Differenzquotienten" berechnen

9.2 Lagrange, quer

2 Formeln;  $p_n(x) = y_0 L_0(x) + y_1 L_1(x) + \dots + y_n L_n(x)$ ;  $L_k(x) = \prod_{j=0, j \neq k}^n \frac{x-x_j}{x_k-x_j}$ ; Jede Basisfunktion  $L_k(x)$  ist ein Polynom vom Grad  $\leq n$ ; **Bemerkung:** Findet Anwendung bei Numerischer Integration; Wenn Stützstellen  $x_i$  gleich bleiben & nur  $y_i$  ändern  $\Rightarrow$  keine Neuberechnung; Rechenaufwand  $\mathcal{O}((n+1)^2)$ ; Kommen neue Stützpunkte hinzu  $\Rightarrow$  Neuberechnung!; Die Interpolationspolynome liefern nur sinnvolle Näherungswerte für x-Werte, die zwischen den gegebenen Stützstellen liegen; Extrapolation (Näherungswerte für x-Werte außerhalb der Stützstellen) kann zu großen Abweichungen führen.

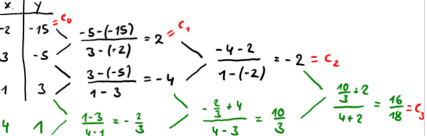
9.3 Newton

Darstellung des Interpolanten, die auf ein gestaffeltes LGS führt & einfache Hinzunahme weiterer Punkte erlaubt.  $p_n(x) = c_0 + c_1(x-x_0) + \dots + c_n(x-x_0)(x-x_1)\dots(x-x_{n-1})$

Polynom vom Grad n  
Das Resultierende LGS für die Koeffizienten  $c_i$  hat gestaffelte Form. **Interpolationsbedingungen?**

**Vorteile:** Rechenaufwand  $\mathcal{O}(n^2)$  Gleitpunktoperationen; Hinzufügen weiterer Stützstellen ohne großen Aufwand. Andere Koeffizienten bleiben unverändert.

9.4 Dividierende Differenzen



9.5 Quiz

Newton & Lagrange ermöglichen ohne großen Berechnungsaufwand die Änderung der Werte  $y_i$  für gleichbleibende Stützstellen  $x_i$ ; Newton ermöglicht ohne großen zusätzlichen Berechnungsaufwand die Hinzunahme weiterer Stützstellen, zur Verbesserung der Genauigkeit

9.6 Effizienz

9.7 klasisch

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

9.8 Horner Schema

$p_3(x) = a_3 x^3 + a_2 x^2 + a_1 x + a_0 = ((a_3 + a_2)x + a_1)x + a_0$ ; Allg.:  $p_n(x) = (\dots(a_n x + a_{n-1})x + \dots + a_1)x + a_0$ ; Aufwand: n Mult.

9.9 Interpolationsfehler

Falls f hinreichend glatt ist &  $p_n$  das eindeutige Interpolati-

onspolynom aus Grad n, dann gilt für den Interpolationsfehler:

$$f(x) - p_n(x) = \frac{f^{(n+1)}(\theta)}{(n+1)!} (x-x_0)\dots(x-x_n)$$

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 Stützstellen

Mit äquidistante Stützstellen konvergiert das Interpolationspolynom nicht immer gegen die zugrundeliegende stetige Funktion, wenn die Anzahl der Stützstellen & damit der Grad des Polynoms wächst. **Lösung:** Nicht-äquidistante Verteilung der Stützstellen, dichter an den Intervallgrenzen.

9.11 Chebyshev-Punkte

haben die Eigenschaft; senkrechte Projektion von gleichverteilten Punkten auf dem Einheitskreis.  $t_k = \cos \frac{(2k-1)\pi}{2n}$ ,  $k = 1, \dots, n$ , auf  $[-1, 1]$ ; Intervall:  $[a, b]$ :  $x_k = \frac{a+b}{2} + \frac{b-a}{2} t_k$ .  $\Rightarrow$  Fehler wird gleichmäßiger verteilt und Konvergenz erreicht.

9.12 Schwächen der Polynominterpolation

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 interpolierende Funktion sicherzustellen; **R:** approx  $\hat{=}$  lin Interpolation; Spline  $\hat{=}$  Spline interpolation; Bibliotheken für Polynominterpolation;

9.13 Spline

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

9.14 Kubisch

**Ansatz:**  $S_i = a_i + b_i(x-x_i) + c_i(x-x_i)^2 + d_i(x-x_i)^3$ ; **Gleichungssystem:** 4n Parameter  $a_i, b_i, c_i, d_i (i = 0, \dots, n-1)$ ; **2n Interpolationsbedingungen:** am Rand je nur eine.  $S_i x_i = y_i$ ;  $S_i(x_{i+1}) = y_{i+1}$  für  $(i = 0, 1, \dots, n-1) \Rightarrow$  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  $i = 0, 1, \dots, n-2$ ; **Stetigkeit der 2. Abl:**  $S''_i(x_{i+1}) = S''_{i+1}(x_{i+1})$ ;  $S''_i(x_{i+1}) - S''_{i+1}(x_{i+1}) = 0$ ; für  $i = 0, 1, \dots, n-2$ ; **natürlicher Randbedingungen:**  $S''_0(x_0) = 0$ ;  $S''_{n-1}(x_n) = 0$ ; nach geschickter Umformung der Gleichungen

chungen hat das LGS die Form  $\begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix} \begin{pmatrix} a_0 \\ b_0 \\ c_0 \\ d_0 \end{pmatrix} = \begin{pmatrix} y_0 \\ y_1 \\ y_2 \\ y_3 \end{pmatrix}$ . **Rechenaufwand**  $\mathcal{O}(n)$  Gleitpunktoperationen.

10 NumInt

Verbesserung der Näherung: Aufteilung in kleine Teilintervalle & Summe von Rechtecksflächen bilden; Interpolation mit Polynom höheren Grades durch diskrete Punkte.

10.1 Ansatz[a,b]

$$\int_a^b f(x) dx \approx (b-a) \sum_{j=0}^i \alpha_j f(x_j)$$

10.2 Def

$p_k \hat{=}$  Interpolationspolynom;  $I_n \hat{=}$  Quadraturformel;  $K \hat{=}$  Fehlerkonstante des Verfahrens; Singularität  $\hat{=}$  isolierter Punkt, der ungewöhnliches Verhalten zeigt;

10.3 Newton-Cotes

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

10.4 Trapezregel

$$T_1: \int_0^1 f(t) dt \approx \frac{1}{2}(f(0)+f(1)); \int_a^b f(x) dx \approx \frac{(b-a)}{1} \frac{1}{2}(f(a)+f(b));$$

$T_n$ : Für Teilintervalle mit gleicher Länge:  
 $h = \frac{b-a}{n}$ ;  $T_n = h(\frac{f(x_0)}{2} + f(x_1) + \dots + f(x_{n-1}) + \frac{f(x_n)}{2})$ ;

10.5 SimpsonRegel

$$S_1: \int_0^1 f(t) dt \approx \frac{1}{6}(f(0) + 4f(0.5) + f(1));$$
$$\int_a^b f(x) dx \approx \frac{b-a}{6}(f(a) + 4f(\frac{a+b}{2}) + f(b));$$

Für n = 1:  $\frac{(b-a)}{2} \frac{1}{3}(f(a) + 4f(\frac{a+b}{2}) + f(b))$ ;  
Für n allg.:  $\frac{(b-a)}{2n} \frac{1}{3}(f(a) + 4(a+h) + \dots + 4f(b-h) + f(b))$   $S_n$ : **Beachte gerade Anzahl an Teilintervallen!**; Für 2n Teilintervalle, 2n+1 Knoten mit gleicher Länge  $h = \frac{b-a}{2n}$ ;  $S_2 = \frac{h}{3}(f(x_0) + 4f(x_1) + 2f(x_2) + 4f(x_3) + f(x_4))$ ;

Newton-Cotes Regeln

Basierend auf äquidistanten Knoten $t_j = \frac{j}{n}$			
$\alpha_j$	Methoden	Ordnung p	
1	2	1	2
2	1	4	1
3	1	3	8
4	7	92	96

*d.h. exakt für  $\int_0^1 x^k dx (k=0,1,\dots,9)$*

Falls  $\alpha_j$  positiv. Integrationsregeln stabil;  $k \leq 7$  &  $k = 9 \Rightarrow$  positive Gewichte;

10.6 Ordnung Integrationsregel

Eine Integrationsregel hat Ordnung p, wenn sie für Polynome vom Grad  $\leq p-1$  exakte Werte liefert;  $T_1$  Ordnung 2

$\Rightarrow$  exakt für Polynome Grad  $\leq 1$ ; Ordnung Newton-Cotes Regeln: mind. Ordnung k+1 (k: Grad des Interpolationspolynoms); **Beweis der Ordnung:**  $1 = \int_0^1 x^0 dx \hat{=}$ ;  $\frac{1}{2} = \int_0^1 x dx \hat{=}$ ;  $\frac{1}{3} = \int_0^1 x^2 dx \hat{=}$ ;  $\frac{1}{4} = \int_0^1 x^3 dx \hat{=}$ ;

10.7 Fehler Quadratur

Für (globalen) Fehler  $e_{In} = \int_a^b f(x) dx - I_n$  einer Quadraturformel  $I_n$  der Ordnung p auf  $[a, b]$  gilt:  $|e_{In}| = (b-a)h^p K |f^{(p)}(\xi)|$ ,  $\xi \in [a, b]$ ,  $h = \frac{b-a}{n}$  &  $|e_{In}| \leq (b-a)h^p K \cdot \max_{a \leq x \leq b} |f^{(p)}(x)|$ ;

10.8 Grenzen NeCo

viele äquidistante Knoten  $\rightarrow$  Gewichte negativ  $\rightarrow$  Verfahren instabil; geschlossene NeCoRe  $\rightarrow$  Funktionsauswertung an RB  $\rightarrow$  Problem mit Singularitäten. größtmögliche Ordnung unerreichbar wegen äquidistanten Knoten; **Lösung:**

10.9 GauQua

Gauß-Quadraturformeln

k	$\alpha_j$	$t_j$	Ordnung
0	1	$\frac{1}{2}$	2
1	$\frac{1}{2}$	$\frac{1}{2}, \frac{\sqrt{3}}{2}$	4
2	$\frac{1}{18}, \frac{8}{18}, \frac{1}{18}$	$\frac{1}{2} - \frac{\sqrt{15}}{10}, \frac{1}{2}, \frac{1}{2} + \frac{\sqrt{15}}{10}$	6

Nur positive Gewichte!

11 Allgemein

11.1 Symbole

Stichprobenstandardabweichung  $\hat{=}$  s; Standardabweichung  $\hat{=}$   $\sigma$

11.2 Abl.

$$x^n \hat{=} nx^{n-1}$$
$$\sin x \hat{=} \cos x; \cos x \hat{=} -\sin x; \tan x \hat{=} \frac{1}{\cos^2 x} = 1 + \frac{\tan^2 x}{1}$$
$$\frac{e^x}{e^x} \hat{=} 1; \frac{a^x}{a^x} \hat{=} \ln a$$
$$\ln x \hat{=} \frac{1}{x}; \log_a x \hat{=} \frac{\ln x}{(\ln a)}$$

11.3 Abl.Regeln

**Faktorregel**  $y = C \cdot f(x) \Rightarrow y' = C \cdot f'(x)$ ;  
**Summenregel**  $y = f_1(x) + f_2(x) + \dots + f_n(x) \Rightarrow y' = f'_1(x) + f'_2(x) + \dots + f'_n(x)$ ;  
**Produktregel**  $y = u \cdot v \Rightarrow y' = u' \cdot v + u \cdot v'$ ;  
**Quotientenregel**  $y = \frac{u}{v} \Rightarrow y' = \frac{u' \cdot v - u \cdot v'}{v^2}$ ;

**Kettenregel**  $f'(x) = F'(u)u'(x) \hat{=} F'(u)$  : Ableitung der Äußeren Funktion;  $u'(x)$  : Ableitung der Inneren Funktion

11.4 Integralregel, elementar

$$\text{Faktorregel } \int_a^b C \cdot f(x) dx = C \cdot \int_a^b f(x) dx;$$

$$\text{Summenregel } \int_a^b [f_1(x) + \dots + f_n(x)] dx = \int_a^b f_1(x) dx + \dots + \int_a^b f_n(x) dx;$$

$$\text{Vertauschungsregel } \int_a^b f(x) dx = - \int_b^a f(x) dx;$$

$$\int_a^a f(x) dx = 0; \int_a^b f(x) dx = \int_c^c f(x) dx + \int_c^b f(x) dx$$

$$\int_c^b f(x) dx \text{ für } (a \leq c \leq b)$$

11.5 Berechnung best. Integr.

$$\int_a^b f(x) dx = [F(x)]_a^b = F(b) - F(a)$$

11.6 Potenzen

$$x^{-n} = \frac{1}{x^n}; a^0 = 1, a^{-n} = \frac{1}{a^n}; a^m \cdot a^n = a^{m+n}; \frac{a^m}{a^n} = a^{m-n} \text{ für } a \neq 0; !(a^m)^n = (a^n)^m = a^{m \cdot n}; a^n \cdot b^n = (a \cdot b)^n; \frac{a^n}{b^n} = (\frac{a}{b})^n \text{ für } b \neq 0; a > 0: a^b = e^{b \ln a}; 0^0 = 1; x_1^1 = x_1;$$

11.7 Wurzel

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

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

$$\sqrt[n]{\sqrt[n]{a}} = \sqrt[n]{a^{\frac{1}{n}}} = (a^{\frac{1}{n}})^{\frac{1}{n}} = a^{\frac{1}{n^2}} = \sqrt[n^2]{a}$$

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

$$\frac{\sqrt[n]{a}}{\sqrt[n]{b}} = \frac{a^{\frac{1}{n}}}{b^{\frac{1}{n}}} = (\frac{a}{b})^{\frac{1}{n}} = \sqrt[n]{\frac{a}{b}} \text{ für } b > 0$$

$$\Rightarrow m, n \in \mathbb{N}^*; a \geq 0, b \geq 0$$

11.8 Abc-Formel

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

11.9 Bin.Formel

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

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

$$(a+b)(a-b) = a^2 - b^2 \text{ 3. Binom;}$$

11.10 Einigungen

o Beim Runden mind. eine Nachkommastelle.

11.11 Trigonometrischer Pythagoras

$$\sin^2 x + \cos^2 x = 1$$

11.12 e

$y = a^x = e^{\lambda x} (\lambda = \ln a)$ ; Def.Ber.:  $-\infty < x < \infty$ ; Wert.ber.:  $0 < y < \infty$ ; Mon.:  $\lambda > 0$  d.h.  $a > 1$ : str. mon. wachst;  $\lambda < 0$  d.h.  $0 < a < 1$ : str. mon. fallt; Asymp.:  $y = \log_a x$  ist Spiegel von  $y = a^x$  an der y-Achse;  $y(0) = 1$  (alle Kurven schneiden die y-Achse bei y = 1);  $y = a^{-1}$  entsteht durch Spiegelung von  $y = a^x$  an der y-Achse.

11.13 Logarithm.

$y = \log_a x$  mit  $x > 0$  ist Umkehrfunktion von  $y = a^x$ ; Def.Ber.:  $x > 0$ ; Wert.Ber.:  $-\infty < y < \infty$ ; Nullst.:  $x_1 = 1$ ; Monot.:  $0 < a < 1$ : str.mon. fallt;  $a > 1$ : str.mon.wachst.; Asymp.:  $x = 0$  (y-Achse);  $\log_a 1 = 0, \log_a a = 1$ ;  $y = \log_a x$  ist Spiegel von  $y = a^x$  an Wink.halb. d. 1. Quadr.