

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

Ω: Grundgesamtheit ω : Element oder Objekt der Grundgesamtheit diskret (<30 Ausprägungen), stetig (≥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 Spannweite

$\max x_i - \min x_i$

1.8 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

1.9 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.

1.10 Quantile

R: $quantile(x, p)$. Teilt die sortierten Daten x_i ca. im Verhältnis p: (1-p) d.h. $\hat{F}(x_p) \approx p$; $\hat{F} \hat{=}$ kummul. rel. Häufigkeit;
1. Quartil = 0.25-Quantil; Median = 0.5-Quantil; 3. Quartil = 0.75-Quantil;

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

1.11 Interquartilsabstand I

$I = x_{0,75} - x_{0,25}$. Ist ein weiterer Streuungsparameter.

1.12 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 Formulierung:** $\bar{S}_k = \{i | |x_i - \bar{x}| \geq k \cdot s\}$; $\frac{N(\bar{S}_k)}{n} \leq \frac{1}{k^2}$;

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 $\bar{x} \pm 3s$.
1.13 Korrelation
Grafische Zusammenhang zwischen multivariaten Daten x und y durch ein Streudiagramm. Kennzahlen zur Untersuchung des Zusammenhangs:

1.14 Empirische Kovarianz

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

1.15 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).

1.16 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 $[-0,7]$ bis $[-1] \Rightarrow$ linearer Zusammenhang.

2 Wahrscheinlichkeitsrechnung

2.1 Begriffe

Ergebnisraum Ω: Menge aller möglichen Ergebnisse eines Experiments

Elementarereignis $\omega \in \Omega$: einzelnes Element von Ω

Ereignis $E \subseteq \Omega$: beliebige Teilmenge des Ergebnisraums Ω 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. **Ge-**

genereignis $\bar{E} = \Omega \setminus E$: 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:

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

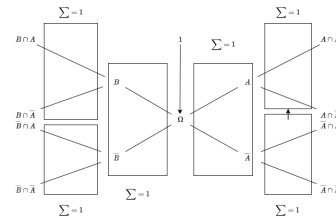
2.6 Bedingte Wahrscheinlichkeit

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

2.7 Satz 2.2

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

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

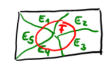


2.8 Satz der totalen Wahrscheinlichkeit

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

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

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** $P(E_k|F) =$

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

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

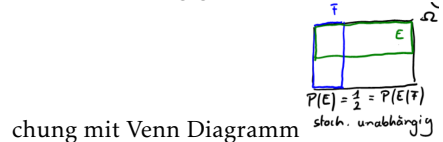
$P(E|F) = P(E)$ or $P(E \cap F) = P(E) \cdot P(F)$

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

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 Veranschauli-



chung mit Venn Diagramm

$$P(\bar{E}) = \frac{1}{2} < P(E|F)$$

$\circ A, B \neq \emptyset$ 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 Ω 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.

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 Ω. Für jedes $X \in \mathbb{R}$ ist die Verteilungsfunktion F: $\mathbb{R} \rightarrow [0, 1]$ einer ZV X definiert durch:

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

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

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

\circ 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 f: $\mathbb{R} \rightarrow [0, \infty]$ definiert durch

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

Es gilt:

$\circ F(x) = P(X \leq x) = \int_{-\infty}^x f(t) dt$ 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 Dichtefunktion $f_x \int_{-\infty}^{\infty} f(x) dx = 1$

\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 μ ist ein quadratisches Streuungsmaß. $\sigma^2 = Var[X] =$

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

$g(X)$

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

- $Var[b] = 0$
- $Var[aX + b] = a^2 Var[X]$

3.11 Satz 3.2

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

3.12 Z-Transformation, Standardisierung

Sei X eine ZV mit μ und σ . Dann ist

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

3.13 Kovarianz

Eigenschaften:

- $Cov[X, Y] = Cov[Y, X]$
- $Cov[X, X] = Var[X]$
- $Cov[aX, Y] = aCov[X, Y]$

Die Kovarianz zweier ZV (X, Y) ist definiert durch $Cov[X, Y] = E[(X - E[X])(Y - E[Y])]$ 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 $\Rightarrow Cov[X, Y] = 0$

3.14 Satz 3.3

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

3.15 Varianz einer Summe von ZV

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

3.16 Overview $\mu \sigma$

3.17 $E[X]$

$$\begin{aligned} E[aX + b] &= aE[X] + b; E[X_1 + \dots + X_n] = \sum_{i=1}^n E[X_i] \\ \text{Falls } X_1, X_2 \text{ unabhängig:} \\ E[X_i] &= \mu \Rightarrow E[\bar{X}] = E\left[\frac{1}{n}(X_1 + \dots + X_n)\right] = \end{aligned}$$

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

μ

3.18 Varianz

$$\begin{aligned} Var[aX + b] &= a^2 Var[X] \\ \text{Falls } X_i, X_j \text{ paarweise unabhängig:} \\ Var[X_1 + \dots + X_n] &= \sum_{i=1}^n Var[X_i] \\ Var[X_i] &= \sigma^2 \Rightarrow Var[\bar{X}] = Var\left[\frac{1}{n}(x_1 + \dots + x_n)\right] = \end{aligned}$$

$$\frac{1}{n^2} \sum_{i=1}^n Var[X_i] = \frac{1}{n^2} \cdot n \cdot \sigma^2 = \frac{\sigma^2}{n}$$

3.19 Quantile

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

$$\begin{aligned} F(x_p) &\geq p. \text{ p-Quantil einer stetigen ZV mit } \textbf{streng monoton wachsenden} \\ F(x): x_p &= F^{-1}(p) \text{ d. h. umkehrbar.} \end{aligned}$$

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; $E[X] = p = \sum x_i \cdot p(x_i) = 1 \cdot p(1); Var[X] = p(1-p) = E[X^2] - (E[X])^2 = p - p^2 = p(1-p)$;

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** $X \sim B_{n,p}$; $E[X] = np$; $Var[X] = np(1-p)$; **R**: $d\text{binom}(k, n, p) = P(X=k) \triangleq$ Wahrscheinlichkeits-/Dichtefunktion; $p\text{binom}(k, n, p) = F(k) \triangleq$ Verteilungsfunktion; $q\text{binom}(q, n, p) \triangleq$ q -Quantil; $r\text{binom}(q, n, p) \triangleq$ k 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** $P(X = k) = \frac{\binom{M}{k} \cdot \binom{N-M}{n-k}}{\binom{M+N}{n}}, k \in \{0, 1, \dots, \min\{n, M\}\}$; **Verteilung** $X \sim H_{M, N, n}$; $E[X] = n \frac{M}{M+N}$;

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

4.5 Poisson-Verteilung

Verteilung der seltenen Ereignisse Häufigkeit punktförmiger Ereignisse in einem Kontinuum. Die durchschnittlich zu erwartende Anzahl der Erfolge λ 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 $\sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=1}^{\infty} \lambda \frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda} \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} = \lambda$; $Var[X] = \lambda$ **R**: $d\text{pois}(k, \lambda) = P(X = k)$; $p\text{pois}(k, \lambda) = F(k)$;

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

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

$Var[X] = \frac{(b-a)^2}{12}$ **R**: $d\text{unif}(x, a, b) = f(x)$; $p\text{unif}(x, a, b) = F(x)$; $r\text{unif}(n) \triangleq n$ Zufallszahlen zwischen 0 und 1; $r\text{unif}(n, a, b) \triangleq n$ Zufallszahlen zwischen a und b ;

4.9 Normalverteilung

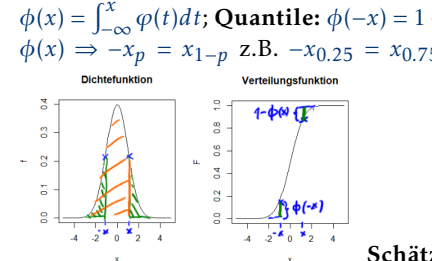
Beschreibt viele reale Situationen, ist insbesondere Grenzverteilung unabhängiger Summen; **Dichte**: $f(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}$; **Verteilung**: $X \sim N_{\mu, \sigma^2}$; $E[X] = \mu$; $Var[X] = \sigma^2$; **R**: $d\text{norm}(x, \mu, \sigma) = f(x)$;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

stelle von $f(x)$ bei $x = \mu \pm \sigma$; $E[aX + b] = aE[X] + b$; $Var[aX + b] = a^2 Var[X]$; $X \sim N_{\mu, \sigma^2} \Rightarrow aX + b \sim N_{a\mu + b, a^2 \sigma^2}$ und $\frac{X - \mu}{\sigma} \sim N_{0,1}$; $X_1 \sim N_{\mu_1, \sigma_1^2}$ 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. $-x_{0.25} = x_{0.75}$;



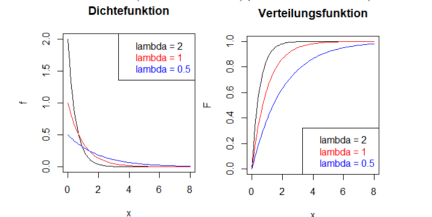
werte: $Z = \frac{x - \mu}{\sigma} \sim N_{0,1}$;

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

4.11 Exponentialverteilung

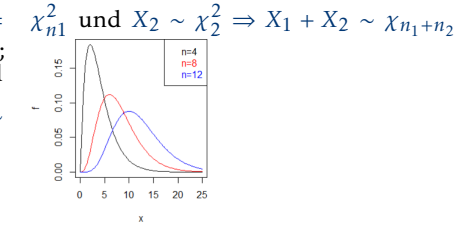
Modellierung von Lebensdauern, Wartezeiten Sei $Y_t \sim P_{\lambda t}$ im Intervall $[0, t]$ von t Zeiteinheiten, dann beschreibt die Exponentialverteilung die Wartezeit X bis zum Eintreten eines Ereignisses; **Dichte- und Verteilungsfunktion**: $f(x) = \lambda e^{-\lambda x} (x \geq 0)$ und $F(x) =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

$p\text{exp}(x, \lambda) = F(x)$; **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**: $X \sim \chi_n^2$; $E[X] = n$; $Var[X] = 2n$; **R**: $d\text{chisq}(x, n) = f(x)$; $p\text{chisq}(x, n) = F(x)$; **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 $n > 1$; $Var[Y] = \frac{n}{n-2}$ für $n > 2$; **R**: $d\text{t}(y, n) \triangleq f(x)$; $p\text{t}(y, n) \triangleq F(x)$; **Eigenschaften**: Für $n \rightarrow \infty : t_n \rightarrow N_{0,1}$; Achsensymmetrie der Dichtefunktion $\Rightarrow -y_p = x_{1-p}$

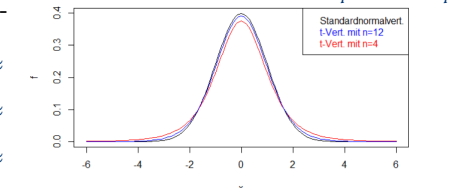


Abbildung Dichtefunktion

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 μ und Varianz σ^2 . Dann gilt für hinreichend große n (> 30) und $\bar{X} = \frac{1}{n} \sum_{i=1}^n$ näherungsweise:

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

$\sum X_i$ 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 ϕ

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

5.3 ϕ^{-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 μ und σ^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 μ , 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 σ^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 μ und Varianz σ^2 . Dann gilt:
bei bekannter Varianz: $\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N_{0,1}$;

$\frac{(n-1)S^2 = \sum (x_i - \bar{x})^2}{\sigma^2} \Rightarrow \text{Standardisierung} \sim \chi^2_{n-1}$; Bei unbekannter Varianz: $\frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} \sqrt{n} \sim t_{n-1}$;
6 Konfidenzintervall
6.1 Begriffe
 Irrtumswahrscheinlichkeit = α ; 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:
 $P(-a \leq \bar{x} \leq a) > 0.95$; σ 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 μ , unbekannt, σ^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 μ & σ^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 μ gültig ist or nicht.
7.1 Def
 α = 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. μ , 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:** α 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) α wird vorgeg
H_0 ist falsch.	falsch (Wsk: Fehler 2. Art)	richtig

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 α 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 $H_1 : \mu \neq \mu_0$; $\bar{X} \sim N_{\mu_0, \sigma_0^2/n} \Rightarrow \frac{\bar{X} - \mu_0}{\frac{\sigma_0}{\sqrt{n}}} \sqrt{n} \sim N_{0,1}$; $P_{\mu_0}(\bar{X} \in C) \leq \alpha \Leftrightarrow |TG| = \frac{|\bar{X} - \mu_0|}{\frac{\sigma_0}{\sqrt{n}}} \sqrt{n} > \phi^{-1}(1 - \frac{\alpha}{2})$;
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, σ^2 bekannt, μ 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, μ, σ^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 μ_0 abweichenden Wert zu bekommen. Der p-Wert zu einer Hypothese H_0 ist der kleinste Wert von α , 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 α 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 α ;
7.13 Zusammenfassung klass. Hypo.test
 Signifikanzniveau α wird vorgegeben; α & Verteilung der Testgröße unter H_0 wir der Ablehnungsbereich ermittelt. Je kleiner (größer) α , desto kleiner (größer) ist der Ablehnungsbereich;
 !: α & C hängen nicht von der konkreten Stichprobe ab;
 H_0 wird abgelehnt, falls der ermittelte Wert der Testgröße (beobachteter Wert) in C liegt. !: Die tg hängt von der konkreten Stichprobe ab. Sie ist eine ZV.
7.14 Test mittels p-Wert
 α wird vorgegeben.
 Berechnung des p-Werts anhand der konkreten 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 $p - \text{Wert} \leq \alpha$;
8 Fehleranalyse
 Derzeit ausgeklammert
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 vertauschte 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 wiederholte Bildung von "Differenzquotienten" berechnen
9.2 Vandermonde/klassisch
 Unterschiedliche Darstellungen für ein Interpolationspolynom $G(x) = p_n(x)$ vom Grad n haben unterschiedliche Eigenschaften bei der numerischen

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