

Beschreibende Statistik
Lageparameter
Arithmetisches Mittel $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{1}{n} (x_1 + \cdots + a_n)$ Das Arithmetische Mittel \bar{x} minimiert die Funktion $g(t) = \sum_{i=1}^n (x_i - t)^2$
Geometrisches Mittel $\bar{x}_{geom} = \sqrt[n]{x_1 \cdot x_2 \cdot \cdots \cdot x_n}$
Median $\tilde{x} = \begin{cases} x_{\frac{n+1}{2}} & , ungerade \\ \frac{1}{2} \cdot (x_{\frac{n}{2}} + x_{\frac{n}{2}+1}) & , gerade \end{cases}$ Der Median \tilde{x} minimiert die Funktion $g(t) = \sum_{i=1}^n x_i - t $
Streuungsmaße
(empirische) Varianz $var = \sigma^2 = s_n^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$ alternativ $var = \sigma^2 = \frac{n}{n-1} \cdot (\overline{x^2} - \bar{x}^2)$
Standardabweichung $\sigma = s_n = \sqrt{\sigma^2}$ $\sigma = s_n = \sqrt{s_n^2}$
mittlere absolute Abweichung $\frac{1}{n} \sum_{i=1}^n x_i - \tilde{x} \text{ f\"ur Median}$ $\frac{1}{n} \sum_{i=1}^n x_i - \bar{x} \text{ f\"ur arithmetisches Mittel}$
Kovarianz und Korrelationskoeffizient
Kovarianz $cov(x, y) = S_{xy} = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x}) \cdot (y_i - \bar{y})$ alternativ $cov(x, y) = S_{xy} = \frac{1}{n-1} \sum_{i=1}^n (x_i \cdot y_i - n \cdot \bar{x} \cdot \bar{y})$
Korrelationskoeffizient $r_{xy} = \frac{S_{xy}}{S_x \cdot S_y}$ Der Korrelationskoeffizient liegt immer zwi-

schen $-1 \leq r \leq +1$. Je nher r_{xy} bei -1 (negative Korellation/Steigung), oder $+1$ (positive Steigung/Korrelation) liegt, desto genau-er schmiegen sich die Messwerte an eine Ge- rade an. Bei r_{xy} nahe 0 gibt es keinen <i>linearen</i> Zusammenhang zwischen den Merkmalen.
Regressionsrechnung
Regressionsgerade Variante 1 $y = \bar{y} + \frac{S_{xy}}{\sigma_x^2} \cdot (x - \bar{x})$ Variante 2 $y = b + a \cdot x$ $a = \frac{S_{xy}}{\sigma_x^2} \text{ und } b = \bar{y} - a \cdot \bar{x}$
Kleinste quadratische Abweichung Die Parameter a, b, c, \dots werden so gewhlt, dass $Q(a, b, c, \dots) = \sum_{i=1}^n (f_{a,b,c,\dots}(x_i) - y_i)^2$ minimal ist $f_{a,b,c,\dots}(x_i)$ ist die Funktion dessen Parameter gesucht werden Nullsetzen der partiellen Ableitungen: $\frac{\partial}{\partial a} Q(a, b) = 0$ $\frac{\partial}{\partial b} Q(a, b) = 0$ ber die Ableitungen lassen sich die Parameter finden welche die vorgegebene Funktion am besten annhern
Vergleich ermittelter Kurven Um Kurven zu vergleichen, einfach die ermittelten Parameter in die $Q(a, b, c, \dots)$ Funktion eingeben und Wert berechnen. Je kleiner der Wert desto besser passt die Kurve
Wahrscheinlichkeitstheorie
Wahrscheinlichkeitsrume
Der Wahrscheinlichkeitsbegriff $Ergebnismenge = \Omega$ Beispiel Wrfel $\Omega = \{1, 2, 3, 4, 5, 6\}$ Ein Ereignis ist eine Teilmenge der Ergebnismenge $\emptyset \subseteq \Omega \hat{=} \text{unmgliches Ereignis}$ $\Omega \subseteq \Omega \hat{=} \text{sicheres Ereignis}$ $A = \{1, 2, 3\}$ Ereignis $\bar{A} = \{4, 5, 6\}$ Gegenereignis

Elementarereignis einelementige Teilmenge von Ω Ereignis, eine 3 werfen $B = \{3\}$ $P(\{3\}) = \frac{1}{6}$																
Laplace-Versuch Jedes Elementarereignis ist gleich wahrscheinlich $P(\{\omega_i\}) = \frac{1}{ \Omega }$ $P(A) = \frac{ A }{ \Omega } = \frac{3}{6} = \frac{1}{2}$																
Bedingte Wahrscheinlichkeit																
Bedingte Wahrscheinlichkeit Wahrscheinlichkeit für A unter der Bedingung B $P(A B) = \frac{P(A \cap B)}{P(B)}$ $P(\bar{A} B) = 1 - P(A B)$																
Formel von Bayes $P(A B) = \frac{P(B A) \cdot P(A)}{P(B)}$																
Satz der totalen Wahrscheinlichkeit $P(A) = \sum_{i=1}^n (P(A B_i) \cdot P(B_i))$																
Viel Felder Tafel <table><tr><td></td><td>A</td><td>\bar{A}</td><td>Σ</td></tr><tr><td>B</td><td>$P(A \cap B)$</td><td>$P(\bar{A} \cap B)$</td><td>$P(B)$</td></tr><tr><td>\bar{B}</td><td>$P(A \cap \bar{B})$</td><td>$P(\bar{A} \cap \bar{B})$</td><td>$P(\bar{B})$</td></tr><tr><td>Σ</td><td>$P(A)$</td><td>$P(\bar{A})$</td><td>1</td></tr></table> <p>Die Ränder sind immer die Summen der zugehörigen Zeilen oder Spalten</p>		A	\bar{A}	Σ	B	$P(A \cap B)$	$P(\bar{A} \cap B)$	$P(B)$	\bar{B}	$P(A \cap \bar{B})$	$P(\bar{A} \cap \bar{B})$	$P(\bar{B})$	Σ	$P(A)$	$P(\bar{A})$	1
	A	\bar{A}	Σ													
B	$P(A \cap B)$	$P(\bar{A} \cap B)$	$P(B)$													
\bar{B}	$P(A \cap \bar{B})$	$P(\bar{A} \cap \bar{B})$	$P(\bar{B})$													
Σ	$P(A)$	$P(\bar{A})$	1													
Allgemeine Regeln $P(A \cap B) = P(A B) \cdot P(B) = P(B A) \cdot P(A)$ $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ $P(\bar{A}) = 1 - P(A)$ $P(\bar{A} \cup \bar{B}) = P(\bar{A} \cap \bar{B})$ $P(A \cap B) = P(\bar{A} \cup \bar{B})$ Wenn A und B unabhängig, dann gilt $P(A \cap B) = P(A) \cdot P(B)$ $P(A B) = P(A)$																

Zufallsvariablen Eine Zufallsvariable ist eine Zuordnungsvorschrift die jedem mglichen Ergebnis eines Zufallsexperiments eine Groe zuordnet $X = k \hat{=} \{\omega \in \Omega X(\omega) = k\}$ $X = 3 \hat{=} \{\omega \in \Omega X(\omega) = 3\}$ $X \leq k \hat{=} \{\omega \in \Omega X(\omega) \leq k\}$
Diskrete Verteilungen
Binomialverteilung Mit zurucklegen, Wahrscheinlichkeit fur jedes Ereignis gleich $X \sim B(n, p)$ $n =$: Stichprobenumfang $p =$: Wahrscheinlichkeit (p muss bei Binomialverteilung fest bleiben) $P(X = k) = \binom{n}{k} \cdot p^k \cdot (1 - p)^{n-k}$ $P(X \leq k) = \sum_{i=0}^k \binom{n}{i} \cdot p^i \cdot (1 - p)^{n-i}$ $P(X > k) = 1 - P(X \leq k)$ Eingabe Taschenrechner $\binom{n}{k} \hat{=} n nCk k$ $Mode \rightarrow 4 \rightarrow \downarrow \rightarrow 1 \rightarrow 2 \rightarrow k \rightarrow n \rightarrow p$
Binomialverteilung approximieren Die Binomialverteilung kann mit der Poisson Verteilung approximiert werden, dann gilt $\lambda = n \cdot p$ Die Binomialverteilung kann auch mit der Normalverteilung approximiert werden, wobei gilt $n \cdot p = \mu$ und $n \cdot p \cdot (1 - p) = \sigma^2$, bedingung ist $X \sim B(n, p) \approx N(n \cdot p, n \cdot p \cdot (1 - p))$ falls gilt $n \cdot p \cdot (1 - p) > 9$ Bei der approximation mit der Normalverteilung kann man eine Stetigkeitskorrektur verwenden um ein besseres Ergebnis zu erhalten $P(X \leq k) \approx F_N(R + 0,5)$ $P(X < k) \approx F_N(R - 0,5)$ $P(a \leq X \leq b) \approx F_N(b + 0,5) - F_N(a - 0,5)$ Zusammengefasst Falls np und n(1 - p) gro genug sind: $\mu = n \cdot p$ und $\sigma^2 = n \cdot p \cdot (1 - p)$

$$F_B(x) \approx F_N(x + 0.5) = \Phi\left(\frac{x+0.5-np}{\sqrt{np(1-p)}}\right)$$

Hypergeometrische Verteilung

Ohne zurücklegen, Wahrscheinlichkeit ändert sich nach jedem Ereignis
 $X \sim H(N, M, n)$
 $n =$: Stichprobenumfang
 $N =$: Gesamtzahl
 $M =$: Anzahl der Elemente mit der Eigenschaft
$$P(X = k) = \frac{\binom{M}{k} \cdot \binom{N-M}{n-k}}{\binom{N}{n}}$$
$$P(X \leq k) = \sum_{i=0}^k \frac{\binom{M}{i} \cdot \binom{N-M}{n-i}}{\binom{N}{n}}$$
$$P(X > k) = 1 - P(X \leq k)$$

Hypergeometrische Vert. approximieren

Die hypergeometrische Verteilung kann mit der **Binomialverteilung** approximiert werden. Dabei muss folgende Bedingung gelten
 $\frac{n}{N} < 0,05$

Poisson Verteilung

Schlüsselwörter sind **Ereignisse pro Zeiteinheit**, zum Beispiel Anrufe innerhalb bestimmter Zeitspanne
 $X \sim Pois(\lambda)$
$$P(X = k) = \pi_\lambda(k) = \frac{\lambda^k}{k!} \cdot e^{-\lambda}$$
Eingabe Taschenrechner
 $Mode \rightarrow 4 \rightarrow \downarrow \rightarrow 3 \rightarrow 2 \rightarrow k \rightarrow \lambda$

Näherung an Normalverteilung

Wenn λ groß genug ist kann die Verteilungsfunktion $F_P(x)$ der Poissonverteilung durch die Verteilungsfunktion der Normalverteilung $F_N(x)$ mit den Parametern $\mu = \lambda$ und $\sigma^2 = \lambda$ genähert werden:
$$F_P(x) \approx F_N(x + 0.5) = \Phi\left(\frac{x+0.5-\lambda}{\sqrt{\lambda}}\right)$$

Geometrische Verteilung

$X \sim Geom(n, p)$
$$P(X = n) = (1 - p)^{n-1} \cdot p$$
Beispiel: Ein Würfel wird so lange gewürfelt bis eine 6 auftritt. Die Zufallsvariable X ist gleich Anzahl der Würfe

Stetige Verteilungen

Dichtefunktion

Die Dichtefunktion ist ein Hilfsmittel zur **Beschreibung einer stetigen Wahrscheinlichkeitsverteilung**
Bedingungen der Dichtefunktion
$$f(x) \geq 0$$
$$\int_{-\infty}^{\infty} f(x)dx = 1$$
Die Dichtefunktion muss **nicht** stetig sein
Die Dichtefunktion ist die Ableitung der Verteilungsfunktion $F(x)$

Verteilungsfunktion

Eine Verteilungsfunktion ist eine Funktion F , die jedem x einer Zufallsvariable X genau eine Wahrscheinlichkeit $P(X \leq x)$ zuordnet
$$F(x) \rightarrow P(X \leq x)$$
Bedingungen der Verteilungsfunktion
Die Verteilungsfunktion **muss** stetig sein
Die Verteilungsfunktion **muss** monoton steigend sein
$$\lim_{x \rightarrow \infty} F(x) = 1$$
$$\lim_{x \rightarrow -\infty} F(x) = 0$$

Normalverteilung

$X \sim N(\mu, \sigma^2)$
Ist $X \sim N(0, 1)$ dann heißt sie Standardnormalverteilt
Jede Normalverteilung kann standardisiert werden, das heißt die Mitte der Kurve wird auf den Nullpunkt gesetzt
Wenn $X \sim N(\mu, \sigma^2)$ verteilt ist dann ist die standardisierte Zufallsvariable $Z = \frac{x-\mu}{\sigma} \sim N(0, 1)$ standardnormalverteilt
Ist die Zufallsvariable standardverteilt kann die Wahrscheinlichkeit aus der Tabelle abgelesen werden

Regeln für den Phi-Wert

$$P(X \leq k) = \Phi(k)$$
$$P(X \leq -k) = 1 - \Phi(+k)$$
$$P(X = k) = 0$$
 (“Integral ohne Breite!”)
allgemein folgt daraus, wenn
 $X \sim N(\mu, \sigma^2)$
dann gilt
$$P(X \leq k) = \Phi\left(\frac{k-\mu}{\sigma}\right)$$
$$P(a \leq X \leq b) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

Additionssatz der Normalverteilung

Seien X und Y unabhängig und Normalverteilt, dann gilt:
 $X + Y = N(\mu_X + \mu_Y; \sigma_X^2 + \sigma_Y^2)$
Ihre Summe ist ebenfalls Normalverteilt!
Quantile der Normalverteilung
Tabelliert ist das β -Quantil z_β der Normalverteilung $N(0, 1)$
$$P(X \leq z_\beta) = \beta$$
$$z_{1-\beta} = -z_\beta$$
Beispiel
$$\beta = 0.9 \Rightarrow z_\beta = 1.28155$$

Exponentialverteilung

Eine exponentialverteilte Zufallsvariable T hat die Dichte
$$f(t) = \begin{cases} \lambda \cdot e^{-\lambda \cdot t} & , t \geq 0 \\ 0 & , t < 0 \end{cases}$$
und daraus ergibt sich die Verteilungsfunktion
$$F(x) = P(T \leq x) = \int_{-\infty}^x f(t)dt = \begin{cases} 1 - e^{-\lambda \cdot x} & , x \geq 0 \\ 0 & , x < 0 \end{cases}$$

Die Exponentialverteilung ist Gedächtnislos

Gleichverteilung (Rechteckverteilung)

$$f(t) = \begin{cases} \frac{1}{b-a} & , t \in [a, b] \\ 0 & , \text{sonst} \end{cases}$$
$$F(t) = \begin{cases} 0 & , t < a \\ \frac{t-a}{b-a} & , t \in [a, b] \\ 1 & , t > b \end{cases}$$

Erwartungswert und Varianz

Erwartungswert

Erwartungswert und Mittelwert sind prinzipiell gleichwertig, der Erwartungswert entspricht der theoretischen Erwartung, der Mittelwert entspricht den tatsächlichen Werten

Zufallsvariable mit diskreter Verteilung

$$\mu = E(X) = \sum_{i=0}^n (x_i \cdot p_i)$$

Zufallsvariable mit Dichtefunktion f

$$\mu = E(X) = \int_{-\infty}^{\infty} x \cdot f(x)dx$$

$$E(X^2) = \int_{-\infty}^{\infty} x^2 \cdot f(x)dx$$

Exponentialverteilung mit Zufallsvariable T

$$E(T) = \sigma_T = \frac{1}{\lambda}$$

Für Binomialverteilung

$$\mu = E(X) = n \cdot p$$

Für geometrische Verteilung

$$\mu = E(X) = \frac{1}{p}$$

Für Poissonverteilung

$$\mu = E(X) = \lambda$$

Für Hypergeometrischeverteilung

$$E(S_n) = E(X_1 + \dots + X_n) = n \cdot E(X_1) = n \cdot \frac{M}{N}$$

Für Gleichverteilung (Rechteckverteilung)

$$E(T_i) = \frac{a+b}{2}$$

Allgemeine Regeln für den Erwartungswert

$$a, b \in \mathbb{R}$$
$$E(aX + b) = a \cdot E(X) + b$$
$$E(X + Y) = E(X) + E(Y)$$
$$E(aX + bY) = a \cdot E(X) + b \cdot E(Y)$$

Varianz

Zufallsvariable mit diskreter Verteilung

$$\sigma^2 = Var(X) = \sum (x_i - \mu)^2 \cdot p_i$$

Zufallsvariable mit Dichtefunktion f

$$Var(X) = E(X^2) - (E(X))^2$$

Varianz aus Erwartungswert berechnen

$$\sigma^2 = E(X^2) - \mu^2$$

Exponentialverteilung mit Zufallsvariable T

$$Var(T) = \frac{1}{\lambda^2}$$

Für Binomialverteilung

$$\sigma^2 = n \cdot p \cdot (1 - p)$$

Für geometrische Verteilung

$$\sigma^2 = \frac{1}{p^2} - \frac{1}{p}$$

Für Poissonverteilung

$$\sigma^2 = Var(X) = E(X^2) - E(X)^2 = \lambda$$

Für Hypergeometrischeverteilung

$$Var(S_n) = n \cdot \frac{M}{N} \cdot \left(1 - \frac{M}{N}\right) \cdot \frac{N-n}{N-1}$$

Für Gleichverteilung (Rechteckverteilung)

$$Var(T_i) = \frac{(b-a)^2}{12}$$

Allgemeine Regeln für Varianz

$Var(X + b) = Var(X)$
 $Var(aX + b) = a^2 \cdot Var(X)$
 $Var(X + Y) =$
 $Var(X) + Var(Y) + 2 \cdot cov(X, Y)$
wobei gilt:
 $Cov(X, Y) = E((X - \mu_X)(Y - \mu_Y)) =$
 $E(X \cdot Y) - \mu_X \mu_Y$
bei unabhängigen Zufallsvariablen X und Y
ist $Cov(X, Y) = 0$ siehe unten.

Unabhängige Zufallsvariablen

Allgemeine Regeln

$Var(X + const) = Var(X)$
 $Var(X + Y) = Var(X) + Var(Y)$
 $E(X \cdot Y) = E(X) \cdot E(Y)$

Wichtige Sätze der Stochastik

Zentraler Grenzwertsatz

n groß (Anzahl der Zufallsvariablen) $n \geq 30$
 X_i unabhängig und identisch verteilt
 $\hat{=}$ haben die gleiche Verteilung
 $E(X_i) = \mu$
 $Var(X_i) = \sigma^2$
 $\sum X_i \sim N(n \cdot \mu, n \cdot \sigma^2)$
 $\bar{X}_n = \frac{X_1 + \dots + X_n}{n} = \bar{x} \sim N(\mu, \frac{\sigma^2}{n})$
Manche Verteilungen verhalten sich in der
Summe anders, zum Beispiel die
Rechteckverteilung ist nicht mehr R-Verteilt.
Dann wird der Zentrale Grenzwertsatz
verwendet

ZGS - Definition

Seien X_1, \dots, X_n unabhängige und identisch
verteilte Zufallsvariablen (nicht zwangsläufig
Normalverteilt) mit Erwartungswert μ und
Varianz σ^2 . Ihre Summe sei
 $S = X_1 + \dots + X_n$ mit Erwartungswert $n\mu$
und Varianz $n\sigma^2$. Es gilt für die zugehörige
Zufallsvariable
 $Z = \frac{S - n\mu}{\sqrt{n}\sigma} = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$ gilt
 $\lim_{n \rightarrow \infty} P(Z \leq z) = \Phi(z)$

Induktive Statistik - Schätztheorie

Schätzfunktionen

Maximum-Likelihood-Schätzer

$L(x_1, \dots, x_n, \alpha) = \prod_{i=1}^n f(x_i)$

$f(x_i)$ muss eine **Dichtefunktion** sein
 $\frac{\partial \ln L(x_1, \dots, x_n, \alpha)}{\partial \alpha} = 0$
Die Funktion nach dem Parameter α
ableiten und Nullsetzen
Das Ergebnis ist der
Maximum-Likelihood-Schätzer

Konfidenzintervalle

Intervall für $E(X)$ einer Normalverteilung

Ist $X \sim N(\mu, \sigma^2)$ verteilt, dann ist
 $Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1)$

Bei bekannter Standardabweichung σ

$\left[\bar{x} - z_{1-(\alpha/2)} \cdot \frac{\sigma}{\sqrt{n}}, \bar{x} + z_{1-(\alpha/2)} \cdot \frac{\sigma}{\sqrt{n}} \right]$
 \bar{X} = arithmetisches Mittel, bzw
erwartungstreuer Schätzer bei n
unabhängigen Stichproben
 α = Signifikanzwahrscheinlichkeit
(Irrtumswahrscheinlichkeit)
 $1 - \alpha$ = Vertrauensniveau
Ist α gegeben, berechne das Quantil
 $z_{1-(\alpha/2)}$

Bei unbekannter Standardabweichung σ

$\left[\bar{x} - t_{1-(\alpha/2)} \cdot \frac{s}{\sqrt{n}}, \bar{x} + t_{1-(\alpha/2)} \cdot \frac{s}{\sqrt{n}} \right]$
Anstatt σ^2 wird der erwartungstreue
Schätzer s verwendet
 $s^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$
beziehungsweise
 $s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2}$

Induktive Statistik - Hypothesentest

Tests für Lageparameter

Gauß-Test (z -Test, u -Test)

Ist ein Test für den Erwartungswert einer
Normalverteilung bei bekannter
Standardabweichung σ
Formuliere die Hypothesen:
a) $H_0 : \mu = \mu_0$ gegen $H_1 : \mu \neq \mu_0$
b) $H_0 : \mu \geq \mu_0$ gegen $H_1 : \mu < \mu_0$
c) $H_0 : \mu \leq \mu_0$ gegen $H_1 : \mu > \mu_0$
Wähle ein Signifikanzniveau α (z.B.: 0.05)

Ziehe eine Stichprobe vom Umfang n ,
berechne \bar{x} und den zugehörigen
standardisierten Prüfwert:
 $z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$
Bestimme das entsprechende Quantil der
Standartnormalverteilung:
a) $z_{1-\frac{\alpha}{2}}$ bzw. b), c) $z_{1-\alpha}$
Entscheidungsregel: H_0 ist zu verwerfen,
falls
a) $|z| > z_{1-\frac{\alpha}{2}}$
b) $z < -z_{1-\alpha}$
c) $z > z_{1-\alpha}$

t-Test

Formuliere die Hypothesen:
a) $H_0 : \mu = \mu_0$ gegen $H_1 : \mu \neq \mu_0$
b) $H_0 : \mu \geq \mu_0$ gegen $H_1 : \mu < \mu_0$
c) $H_0 : \mu \leq \mu_0$ gegen $H_1 : \mu > \mu_0$
Wähle ein Signifikanzniveau α (z.B.: 0.05)
Ziehe eine Stichprobe vom Umfang n ,
berechne daraus \bar{x} und s sowie den
zugehörigen Prüfwert:
 $t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$
Bestimme das entsprechende Quantil der
 t -Verteilung:
a) $t_{n-1; 1-\frac{\alpha}{2}}$ bzw. b), c) $t_{n-1; 1-\alpha}$
Entscheidungsregel: H_0 ist zu verwerfen,
falls
a) $|t| > t_{n-1; 1-\frac{\alpha}{2}}$
b) $t < -t_{n-1; 1-\alpha}$
c) $t > t_{n-1; 1-\alpha}$

Tests für Streuungsmaße

χ^2 - Anpassungstest

Der χ^2 -Anpassungstest überprüft ob eine
unbekannte Wahrscheinlichkeitsverteilung
einem bestimmten Verteilungsmodell folgt
 $T = \frac{1}{n} \left(\sum_{i=1}^r \frac{N_i^2}{p_i} \right) - n$
wobei N_i absolute Häufigkeit in Kategorie r ,
 n ist der Stichprobenumfang. Alternativ
 $T = \sum_{i=1}^r \frac{(N_i - np_i)^2}{n \cdot p_i}$
Das Ergebnis T mit $\chi^2_{r-1; 1-\alpha}$ Wert aus der
Tabelle vergleichen
 $T < \chi^2 =$ Hypothese wird nicht verworfen
Algorithmus
1. Hypothese aufstellen

2. n und r festlegen (n = Stichprobenumfang,
 r = Anzahl der Klassen)
3. Verteilung auf welche getestet werden soll
bestimmen
4. Alle p_i berechnen (zu jeder Kategorie r
eines nach in 3. festgelegter Verteilung)
5. T berechnen.
6. χ^2 mit T vergleichen.

Übersicht: Induktive Statistik

GEGEBENENFALLS FOLGENDES
NOCH ZU DEN EINZELNEN
POSITIONEN VERSCHIEBEN

Konfidenzbereich/Test für Erwartungswert

Varianz σ^2 bekannt

Zweiseitige Konfidenzintervalle

Wegen ZGS:
 $Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \cdot \sqrt{n} \sim N(0, 1)$
ist zu gegebenen Konfidenzniveau $1 - \alpha$ das
Konfidenzintervall gleich
 $\left[\bar{x} - z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}, \bar{x} + z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \right]$
wobei $z_{1-\frac{\alpha}{2}}$ ein Quantil der
Standartnormalverteilung ist.
Das Konfidenzintervall überdeckt den
gesuchten Erwartungswert μ mit der
Wahrscheinlichkeit $1 - \alpha$.

Einseitige Konfidenzintervalle

$\left] -\infty, \bar{x} + z_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}} \right]$ bzw.
 $\left[-\infty, \bar{x} - z_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}} \right[$

Zweiseitige Tests

Für μ_0 sei
 $Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$
 $H_0 : \mu = \mu_0$ Bereich: $-z_{1-\frac{\alpha}{2}} \leq Z \leq z_{1-\frac{\alpha}{2}}$

Einseitige Tests

$H_0 : \mu \leq \mu_0$ Bereich: $Z \leq z_{1-\alpha}$
 $H_0 : \mu \geq \mu_0$ Bereich: $Z \geq -z_{1-\alpha}$

Varianz σ^2 unbekannt

Zweiseitige Konfidenzintervalle

Allgemeine Matheregeln

Logarithmusregeln
$x = \log_a y \Leftrightarrow y = a^x$
$\log 1 = 0$
$\log x \cdot y = \log x + \log y$
$-\log x = \log \frac{1}{x}$
$\log \frac{x}{y} = \log x - \log y$
$\log x^n = n \cdot \log x$
$\log_a x = \frac{\log x}{\log a}$
$\log \left(\prod_{i=1}^n x_i \right) = \sum_{i=1}^n \log x_i$

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

$$\int_a^b a \cdot f(x)dx = a \cdot \int_a^b f(x)dx$$