

<b>Beschreibende Statistik</b>
<b>Lageparameter</b>
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 $\tilde{x} = \begin{cases} x_{\frac{n+1}{2}} & , \text{ungerade} \\ \frac{1}{2} \cdot (x_{\frac{n}{2}} + x_{\frac{n+1}{2}}) & , \text{gerade} \end{cases}$ Der Median $\tilde{x}$ minimiert die Funktion $g(t) = \sum_{i=1}^n  x_i - t $
<b>Streungsmaße</b>
(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 $\frac{1}{n} \sum_{i=1}^n  x_i - \tilde{x}  \text{ f\"ur Median}$ $\frac{1}{n} \sum_{i=1}^n  x_i - \bar{x}  \text{ f\"ur arithmetisches Mittel}$
<b>Kovarianz und Korrelationskoeffizient</b>
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.
<b>Regressionsrechnung</b>
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
<b>Wahrscheinlichkeitstheorie</b>
<b>Wahrscheinlichkeitsrume</b>
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

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

<b>Zufallsvariablen</b> Eine Zufallsvariable ist eine Zuordnungsvorschrift die jedem mglichen Ergebnis eines Zufallsexperiments eine Groe zuordnet $X = k \hat{=} \{\omega \in \Omega   X(\omega) = k\}$ $X = 3 \hat{=} \{\omega \in \Omega   X(\omega) = 3\}$ $X \leq k \hat{=} \{\omega \in \Omega   X(\omega) \leq k\}$
<b>Diskrete Verteilungen</b>
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 $\binom{n}{k} \hat{=} n   nCr   k$
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, <b>bedingung ist</b> $X \sim B(n, p) \approx N(n \cdot p, n \cdot p \cdot (1 - p))$ 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 $P(X \leq k) \approx F_N(R + 0,5)$ $P(X < k) \approx F_N(R - 0,5)$ $P(a \leq X \leq b) \approx F_N(b + 0,5) - F_N(a - 0,5)$
Hypergeometrische Verteilung Ohne zurucklegen, Wahrscheinlichkeit andert sich nach jedem Ereignis $X \sim H(N, M, n)$ $n =$ : Stichprobenumfang $N =$ : Gesamtzahl

$M$ =: Anzahl der Elemente mit der Eigenschaft $P(X = k) = \frac{\binom{M}{k} \cdot \binom{N-M}{n-k}}{\binom{N}{n}}$ $P(X \leq k) = \sum_{i=0}^k \frac{\binom{M}{i} \cdot \binom{N-M}{n-i}}{\binom{N}{n}}$ $P(X > k) = 1 - P(X \leq k)$
<b>Hypergeometrische Vert. approximieren</b> Die hypergeometrische Verteilung kann mit der <b>Binomialverteilung</b> approximiert werden. Dabei muss folgende Bedingung gelten $\frac{n}{N} < 0,05$
Poisson Verteilung Schlüsselwörter sind <b>Ereignisse pro Zeiteinheit</b> , zum Beispiel Anrufe innerhalb bestimmter Zeitspanne $X \sim Pois(\lambda)$ $P(X = k) = \pi_{\lambda}(k) = \frac{\lambda^k}{k!} \cdot e^{-\lambda}$
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
<b>Stetige Verteilungen</b>
Dichtefunktion Die Dichtefunktion ist ein Hilfsmittel zur <b>Beschreibung einer stetigen Wahrscheinlichkeitsverteilung</b> Bedingungen der Dichtefunktion $f(x) \geq 0$ $\int_{-\infty}^{\infty} f(x) dx = 1$ Die Dichtefunktion muss <b>nicht</b> 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 <b>muss</b> stetig sein Die Verteilungsfunktion <b>muss</b> 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 $P(X \leq k) = \Phi(k)$ $P(X = k) = \Phi(k) = 0$ $P(X \leq -k) = 1 - \Phi(+k)$ allgemein gilt $X \sim N(\mu, \sigma^2)$ $P(X \leq k) = \Phi(\frac{k-\mu}{\sigma})$ $P(a \leq X \leq b) = \Phi(\frac{b-\mu}{\sigma}) - \Phi(\frac{a-\mu}{\sigma})$
Quantile der Normalverteilung Tabelliert ist das $\beta$ -Quantil $z_{\beta}$ 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) $f(t) = \begin{cases} \frac{1}{b-a} & , t \in [a, b] \\ 0 & , sonst \end{cases}$

$F(t) = \begin{cases} 0 & , t < a \\ \frac{t-a}{b-a} & , t \in [a, b] \\ 1 & , t > b \end{cases}$
<b>Erwartungswert und Varianz</b>
<b>Erwartungswert</b> 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 $f$ $\mu = E(X) = \int_{-\infty}^{\infty} x \cdot f(x) dx$
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 Rechteckverteilung $E(T_i) = \frac{a+b}{2}$
Allgemeine Regeln für den Erwartungswert $a, b \in \mathbb{R}$ $E(aX + b) = a \cdot E(X) + b$ $E(X + Y) = E(X) + E(Y)$ $E(aX + bY) = a \cdot E(X) + b \cdot E(Y)$
<b>Varianz</b>
Zufallsvariable mit diskreter Verteilung $\sigma^2 = Var(X) = \sum (x_i - \mu)^2 \cdot p_i$
Zufallsvariable mit Dichtefunktion $f$ $Var(X) = E(X^2) - (E(X))^2$
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 $Var(S_n) = n \cdot \frac{M}{N} \cdot (1 - \frac{M}{N}) \cdot \frac{N-n}{N-1}$
Für Rechteckverteilung $Var(T_i) = \frac{(b-a)^2}{12}$
Allgemeine Regeln für Varianz $Var(X + Y) = Var(X) + Var(Y) + 2 \cdot cov(X, Y)$
<b>Unabhängiger Zufallsvariablen</b>
Allgemeine Regeln $E(X \cdot Y) = E(X) \cdot E(Y)$ $Var(X + Y) = Var(X) + Var(Y)$
<b>Wichtige Sätze der Stochastik</b>
Zentraler Grenzwertsatz n groß (Anzahl der Zufallsvariablen) $n \geq 30$ $X_i$ unabhängig und identisch verteilt $\hat{=}$ haben die gleiche Verteilung $E(X_i) = \mu$ $Var(X_i) = \sigma^2$ $\sum X_i \sim N(n \cdot \mu, n \cdot \sigma^2)$ $\frac{1}{n} \sum_{i=1}^n = \bar{x} \sim N(\mu, \frac{\sigma^2}{n})$ Manche Verteilungen verhalten sich in der Summe anders, zum Beispiel die Rechteckverteilung ist nicht mehr R-Verteilt. Dann wird der Zentrale Grenzwertsatz verwendet
<b>Induktive Statistik - Schätztheorie</b>
<b>Schätzfunktionen</b>
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  $\alpha$   
ableiten und Nullsetzen  
Das Ergebnis ist der  
Maximum-Likelihood-Schätzer

Konfidenzintervalle

Intervall für  $E(X)$  einer Normalverteilung

Bei bekannter Standardabweicheung

Ist  $X \sim N(\mu, \sigma^2)$  verteilt, dann ist  
$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \sim N(0, 1)$$
  
 $\bar{X}$  = erwartungsstreu Schätzer bei  $n$   
unabhängigen Stichproben  
Ist  $\alpha$  gegeben, berechne das Quantil  
$$z_{1-(\alpha/2)}$$
  
Daraus erhält man das Konfidenzintervall  
$$\left[ \bar{x} - z_{1-(\alpha/2)} \cdot \frac{\sigma}{\sqrt{n}}, \bar{x} + z_{1-(\alpha/2)} \cdot \frac{\sigma}{\sqrt{n}} \right]$$

Allgemeine Matheregeln

Potenzen und Logarithmen

Potenzgesetze

$a^0 = 1$
$a^1 = a$
$a^m \cdot a^n = a^{m+n}$
$(a^n)^m = a^{n \cdot m}$
$a^n \cdot b^n = (a \cdot b)^n$
$\frac{a^n}{a^m} = a^{n-m}$
$a^{\frac{b}{n}} = \sqrt[n]{b}$
$\prod_{i=1}^n a^{x_i} = a^{\sum_{i=1}^n x_i}$

Logarithmusregeln

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

Ableitungen und Integrale

Grundlegende Ableitungsregeln

$f(x)$	$f'(x)$
$c = const$	0
$x^n$	$n \cdot x^{n-1}$
$\sqrt{x}$	$\frac{1}{2\sqrt{x}}$
$e^x$	$e^x$
$a^x$	$\ln a \cdot a^x$
$\ln x$	$\frac{1}{x}$
$\log_a x$	$\frac{1}{\ln a \cdot x}$
$\sin x$	$\cos x$
$\cos x$	$-\sin x$
$\tan x$	$\frac{1}{\cos^2 x}$
$\cot x$	$\frac{1}{\sin^2 x}$

Verknüpfte Ableitungsregeln

$f(x)$	$f'(x)$
$(f(x) + g(x))$	$(f'(x) + g'(x))$
$(f(x) \cdot g(x))$	$(f'(x) \cdot g(x)) + (f(x) \cdot g'(x))$
$\frac{f(x)}{g(x)}$	$\frac{(f'(x) \cdot g(x)) - (f(x) \cdot g'(x))}{g(x)^2}$
$f(g(x))$	$f'(g(x)) \cdot g'(x)$

wichtige Stammfunktionen

$f(x)$	$F(x)$
$x^n, n \neq 1$	$\frac{1}{n+1} \cdot x^{n+1} + c$
$\frac{1}{x}, x \neq 0$	$\ln  x  + c$
$\sqrt{x}$	$\frac{2}{3} \cdot x^{\frac{3}{2}} + c$
$e^x$	$e^x + c$