

Beschreibende Statistik
Lageparameter
Arithmetisches Mittel
$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{1}{n} (x_1 + \cdots + a_n)$
Geometrisches Mittel
$\bar{x}_{geom} = \sqrt[n]{x_1 \cdot x_2 \cdot \cdots \cdot x_n}$
Median
$\tilde{x} = \begin{cases} x_{\frac{n+1}{2}} \\ \frac{1}{2}(x_{\frac{n}{2}} + x_{\frac{n}{2}+1}) \end{cases}$
Der Median \tilde{x} minimiert die Funktion
$\sum_{i=0}^n x_i - \tilde{x} $
Streuungsmaße
(empirische) Varianz
$var = \sigma^2 = s_n^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \tilde{x})^2$
Standardabweichung
$\sigma = \sqrt{\sigma^2}$
$\sigma = \sqrt{s_n^2}$
Kovarianz und Korrelationskoeffizient
Kovarianz
$cov(x, y) = S_{xy} = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x}) \cdot (y_i - \bar{y})$
Korrelationskoeffizient
$r_{xy} = \frac{S_{xy}}{S_x \cdot S_y}$
Der Korrelationskoeffizient liegt immer zwischen $-1 \leq r \leq 1$

Regressionsrechnung
Regressionsgerade
Variante 1
$y = \bar{y} + \frac{S_{xy}}{\sigma_x^2} \cdot (x - \bar{x})$
Variante 2
$y = b + a \cdot x$ $a = \frac{S_{xy}}{\sigma_x^2} \text{ und } b = \bar{y} - a \cdot \bar{x}$
Wahrscheinlichkeitstheorie
Wahrscheinlichkeitsräume
Der Wahrscheinlichkeitsbegriff
$Ergebnismenge = \Omega$ <p>Beispiel Würfel $\Omega = \{1, 2, 3, 4, 5, 6\}$ Ein Ereignis ist eine Teilmenge der Ergebnismenge</p> $\emptyset \subseteq \Omega \hat{=} \text{unmögliches Ereignis}$ $\Omega \subseteq \Omega \hat{=} \text{sicheres Ereignis}$ $A = \{1, 2, 3\} \text{ Ereignis}$ $\bar{A} = \{4, 5, 6\} \text{ Gegenereignis}$
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
$P(A) = \sum_i (P(A B_i) \cdot P(B_i))$
Allgemeine Regeln
$P(A \cap B) = P(A B) \cdot P(B) = P(B A) \cdot P(A)$ $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ $P(\bar{A}) = 1 - P(A)$ $P(A \cup B) = P(\bar{A} \cap \bar{B})$ $P(A \cap B) = P(\bar{A} \cup \bar{B})$ <p>Wenn A und B unabhängig, dann gilt</p> $P(A \cap B) = P(A) \cdot P(B)$
Zufallsvariablen
Eine Zufallsvariable ist eine Zuordnungsvorschrift die jedem möglichen Ergebnis eines Zufallsexperiments eine Größe 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 zurücklegen, Wahrscheinlichkeit für jedes Ereignis gleich
$X \sim B(n, p)$ <p>n =: Stichprobenumfang p =: Wahrscheinlichkeit (p muss bei Binomialverteilung fest bleiben)</p> $P(X = k) = \binom{n}{k} \cdot p^k \cdot (1 - p)^{n-k}$ $P(X \leq k) = \sum_{i=0}^k \binom{n}{i} \cdot p^i \cdot (1 - p)^{n-i}$ $P(X > k) = 1 - P(X \leq k)$
Eingabe Taschenrechner

$\binom{n}{k} \hat{=} n nC r k$
Binomialverteilung kann mit der Normalverteilung approximiert werden, bedingung ist
$X \sim B(n, p) \hat{=} N(n \cdot p, n \cdot p \cdot (1 - p))$ <p>falls gilt</p> $n \cdot p \cdot (1 - p) > 9$
Hypergeometrische Verteilung
Ohne zurücklegen, Wahrscheinlichkeit ändert sich nach jedem Ereignis
$X \sim H(N, M, n)$ <p>n =: Stichprobenumfang N =: Gesamtzahl M =: Anzahl der Elemente mit der Eigenschaft</p> $P(X = k) = \frac{\binom{M}{k} \cdot \binom{N-M}{n-k}}{\binom{N}{n}}$ $P(X \leq k) = \sum_{i=0}^k \frac{\binom{M}{i} \cdot \binom{N-M}{n-i}}{\binom{N}{n}}$ $P(X > k) = 1 - P(X \leq k)$
Poisson Verteilung
$X \sim Pois(\lambda)$ $P(X = k) = \pi_\lambda(k) = \frac{\lambda^k}{k!} \cdot e^{-\lambda}$
Geometrische Verteilung
$X \sim Geom(n, p)$ $P(X = n) = (1 - p)^{n-1} \cdot p^n$
Beispiel: Ein Würfel wird so lange gewürfelt bis eine 6 auftritt. Die Zufallsvariable X ist gleich Anzahl der Würfe
Stetige Verteilungen
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
 $X \sim N(\mu, \sigma^2)$
 $P(X \leq k) = \Phi(\frac{k-\mu}{\sigma})$

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
 $\beta = 0.9 \Rightarrow z_\beta = 1.28155$

Erwartungswert und Varianz

Erwartungswert

Zufallsvariable mit diskreter Verteilung

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

Zufallsvariable mit Dichtefunktion f

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

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

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

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

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

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

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

Allgemeine Regeln für Varianz

$$Var(X + Y) =$$
$$Var(X) + Var(Y) + 2 \cdot cov(X, Y)$$

Unabhängiger Zufallsvariablen

Allgemeine Regeln

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