

# Formelsammlung Statistik für Wirtschaftswissenschaftler

## 2. Häufigkeitsverteilungen

Relative Häufigkeit	$\frac{\text{Absolute Häufigkeit}}{\text{Gesamtzahl der Beobachtungen}}$
Prozentuale Häufigkeit	Relative Häufigkeit · 100
Kumulierte Häufigkeit	$H_i = \sum_{j=1}^i h_j$
Winkel eines Kreisdiagramms	Relative Häufigkeit · 360
Klassenbreite	(Maximalwert – Minimalwert)/Klassenzahl

## 3. Lagemaße statistischer Verteilungen

	Grundgesamtheit	Stichprobe
Arithmetisches Mittel	$\mu_a = \frac{1}{N} \sum_{i=1}^N x_i$	$\bar{x}_a = \frac{1}{n} \sum_{i=1}^n x_i$
Arithmetisches Mittel mit absoluten Häufigkeiten	$\mu_a = \frac{1}{N} \sum_{i=1}^N h_i \cdot x_i$	$\bar{x}_a = \frac{1}{n} \sum_{i=1}^n h_i \cdot x_i$
Geometrisches Mittel	$\mu_g = \sqrt[n]{\prod_{i=1}^N x_i}$	$\bar{x}_g = \sqrt[n]{\prod_{i=1}^n x_i}$
P-tes Perzentil bzw. q-tes Quantil	$i = (p/100) \cdot n$ <p>Ist i ganzzahlig, dann ist <math>P_p = Q_q = \frac{1}{2} (x_i + x_{i+1})</math></p> <p>Ist i keine ganze Zahl, dann ist <math>P_p = Q_q = x_j</math>, wo j die auf i folgende ganze Zahl ist</p>	

## 4. Streuungsmaße statistischer Verteilungen

Spannweite	Größter Wert – kleinster Wert	
Interquartilsabstand	3. Quartil – 1. Quartil	
	Grundgesamtheit	Stichprobe
Varianz	$\sigma^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2$	$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$
Standardabweichung	$\sigma = \sqrt{\sigma^2}$	$s = \sqrt{s^2}$
Variationskoeffizient	$\frac{\sigma}{ \mu } \cdot 100$	$\frac{s}{ \bar{x} } \cdot 100$

## 5. Weitere Maße statistischer Verteilungen

	Grundgesamtheit	Stichprobe
Schiefe	$\gamma = \frac{1}{N} \sum_{i=1}^N \left( \frac{x_i - \mu}{\sigma} \right)^3$	$g = \frac{n}{(n-1) \cdot (n-2)} \sum_{i=1}^n \left( \frac{x_i - \bar{x}}{s} \right)^3$
Lorenzkurve	$a_k$ werden aufsteigend geordnet $x_i = \frac{i}{n} \quad y_i = \frac{1}{S} \sum_{k=1}^i a_k = \frac{1}{S} \cdot S_i$ mit $S = \sum_{k=1}^n a_k$	
Gini-Koeffizient	$\frac{n}{n-1} \left( \sum_{i=1}^n \frac{2i-1}{n} \cdot \frac{a_i}{S} - 1 \right)$	
Gini-Koeffizient mit absoluten Häufigkeiten	$\frac{S_x}{S_x - 1} \left( \sum_{i=1}^n (x_i + x_{i-1})(y_i - y_{i-1}) - 1 \right)$ mit $S_x = \sum_{k=1}^n h_k$ , $S_y = \sum_{k=1}^n h_k \cdot a_k$ , $x_i = \frac{1}{S_x} \sum_{k=1}^i h_k$ , $y_i = \frac{1}{S_y} \sum_{k=1}^i h_k \cdot a_k$	

## 6. Wahrscheinlichkeitsrechnung

Wahrscheinlichkeit eines Komplementärereignisses	$P(\neg A) = 1 - P(A)$
--	------------------------

Additionssatz	$P(A \cup B) = P(A) + P(B) - P(A \cap B)$
Bedingte Wahrscheinlichkeit	$P(A B) = \frac{P(A \cap B)}{P(B)}$
Multiplikationssatz	$P(A \cap B) = P(B A) \cdot P(A)$
Satz von Bayes	$P(A_i B) = \frac{P(B A_i) \cdot P(A_i)}{P(B A_1) \cdot P(A_1) + P(B A_2) \cdot P(A_2) + \dots + P(B A_n) \cdot P(A_n)}$
Unabhängige Ereignisse	$P(A B) = P(A), P(B A) = P(B)$
Multiplikationssatz für unabhängige Ereignisse	$P(A \cap B) = P(A) \cdot P(B)$
Additionssatz für sich ausschließende Ereignisse	$P(A \cup B) = P(A) + P(B)$

## 7. Wahrscheinlichkeitsverteilungen

	Diskret	Stetig
Erwartungswert einer Zufallsvariablen	$E(X) = \mu = \sum_{i=1}^n x_i \cdot p(x_i)$	$E(X) = \mu = \int_{-\infty}^{\infty} x \cdot f(x) dx$
Varianz einer Zufallsvariablen	$V(X) = \sigma^2 = \sum_{i=1}^n (x_i - \mu)^2 \cdot p(x_i)$	$V(X) = \sigma^2 = \int_{-\infty}^{\infty} (x - \mu)^2 \cdot f(x) dx$
Stetige Gleichverteilung	$f(x) = \begin{cases} \frac{1}{b-a} & \text{für } a \leq x \leq b \\ 0 & \text{sonst} \end{cases}$ $E(X) = \frac{a+b}{2} \quad V(X) = \frac{(b-a)^2}{12}$	
Binomialverteilung	$P(X = k) = B_{n,p}(k) = \binom{n}{k} p^k \cdot (1-p)^{n-k}$ <p>mit <math>\binom{n}{k} = \frac{n!}{k!(n-k)!}</math></p> $B_{n,p}(k) = B_{n,1-p}(n-k)$ $F_{n,p}(k) = 1 - F_{n,1-p}(n-k-1)$ $E(X_{\text{Binom},n;p}) = n \cdot p \quad V(X_{\text{Binom},n;p}) = n \cdot p \cdot (1-p)$	
z-Transformation	$z = \frac{x - \mu}{\sigma}$	

Poisson-Verteilung	$p_{\text{Poisson},\mu}(x) = e^{-\mu} \cdot \frac{\mu^x}{x!} \text{ für } x = 0, 1, 2, \dots$ $E(X_{\text{Poisson},\mu}) = \mu \quad V(X_{\text{Poisson},\mu}) = \mu$
Exponentialverteilung	$P(X \leq x_0) = 1 - e^{-\lambda \cdot x_0} \quad (x_0 \geq 0)$ $E(X_{\text{Exponential},\lambda}) = \frac{1}{\lambda} \quad V(X_{\text{Exponential},\lambda}) = \left(\frac{1}{\lambda}\right)^2$

## 8. Punkt- und Intervallschätzungen

	Mittelwert	Varianz
Punktschätzung für Parameter der Grundgesamtheit	$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$	$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$
	Erwartungswert	Varianz
Stichprobenmittel	$E(\bar{X}) = \mu$	$V(\bar{X}) = \frac{\sigma^2}{n}$
Intervallschätzung für Mittelwert der Grundgesamtheit, wenn $\sigma$ bekannt ist	$\bar{x} \pm z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$	
Intervallschätzung für Mittelwert der Grundgesamtheit, wenn $\sigma$ unbekannt ist	$\mu = \bar{x} \pm t_{\alpha/2} \cdot \frac{s}{\sqrt{n}}$ $n - 1$ Freiheitsgrade	
Stichprobenumfang, wenn $\sigma$ bekannt ist	$n \geq \left( \frac{z_{\alpha/2} \cdot \sigma}{e} \right)^2$ $n$ wird immer aufgerundet	

## 9. Hypothesentests über Mittelwerte

Schritte eines Hypothesentests

1. Entscheidung, ob ein zweiseitiger oder einseitiger Test angemessen ist.
2. Formulierung der Nullhypothese und der Alternativhypothese.
3. Festlegung des Signifikanzniveaus.
4. Ermittlung der Testgröße.
5. Entscheidung über Ablehnung oder vorläufige Annahme der Nullhypothese.

Test über einen Mittelwert, wenn $\sigma$ bekannt ist	$\bar{z} = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$
Test über einen Mittelwert, wenn $\sigma$ unbekannt ist	$\bar{t} = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$ $n - 1$ Freiheitsgrade

## 10. Statistische Analyse der Differenz von zwei Mittelwerten

Intervallschätzung für Mittelwertdifferenz, wenn $\sigma_1$ und $\sigma_2$ bekannt sind	$\bar{x}_1 - \bar{x}_2 \pm z_{\alpha/2} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$
Intervallschätzung für Mittelwertdifferenz, wenn $\sigma_1$ und $\sigma_2$ unbekannt sind	$\bar{x}_1 - \bar{x}_2 \pm t_{\alpha/2} \cdot \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$ $fg = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{1}{n_1 - 1} \cdot \left(\frac{s_1^2}{n_1}\right)^2 + \frac{1}{n_2 - 1} \cdot \left(\frac{s_2^2}{n_2}\right)^2}$ fg wird immer abgerundet
Hypothesentest über eine Mittelwertdifferenz, wenn $\sigma_1$ und $\sigma_2$ bekannt sind	$z = \frac{\bar{x}_1 - \bar{x}_2 - \delta_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$
Hypothesentest über eine Mittelwertdifferenz, wenn $\sigma_1$ und $\sigma_2$ unbekannt sind	$t = \frac{\bar{x}_1 - \bar{x}_2 - \delta_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$ $fg = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{1}{n_1 - 1} \cdot \left(\frac{s_1^2}{n_1}\right)^2 + \frac{1}{n_2 - 1} \cdot \left(\frac{s_2^2}{n_2}\right)^2}$ fg wird immer abgerundet

## 11. Auswertung von zweidimensionalen Daten

	Grundgesamtheit	Stichprobe
Kovarianz	$\sigma_{xy} = \frac{\sum_{i=1}^N (x_i - \mu_x) \cdot (y_i - \mu_y)}{N}$	$s_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x}) \cdot (y_i - \bar{y})}{n - 1}$
Pearsons Korrelationskoeffizient	$\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \cdot \sigma_y}$	$r_{xy} = \frac{s_{xy}}{s_x \cdot s_y}$
Häufigkeiten der Randverteilung	$f_{i.} = \sum_{j=1}^m f_{ij} \text{ für } i = 1, \dots, n$ $f_{.j} = \sum_{i=1}^n f_{ij} \text{ für } j = 1, \dots, m$	
$\chi^2$ -Test auf Unabhängigkeit	$\chi^2 = \sum_{i=1}^n \sum_{j=1}^m \frac{(f_{ij} - e_{ij})^2}{e_{ij}}$ <p>mit <math>e_{ij} = (f_{i.} \cdot f_{.j}) / f_{..}</math></p> <p><math>(n - 1) \cdot (m - 1)</math> Freiheitsgrade</p>	

## 12. Einfache lineare Regression

Lineares Regressionsmodell	$y = \beta_0 + \beta_1 \cdot x + \varepsilon$
Lineare Regressionsschätzung	$\hat{y} = b_0 + b_1 \cdot x$
Steigung der Regressionsgeraden	$b_1 = \frac{\sum_{i=1}^n (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}$
Achsenabschnitt der Regressionsgeraden	$b_0 = \bar{y} - b_1 \cdot \bar{x}$
Beziehung zwischen SQT, SQE und SQR	$SQT = SQE + SQR$ $\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 + \sum_{i=1}^n (y_i - \hat{y}_i)^2$
Bestimmtheitsmaß	$r^2 = \frac{SQE}{SQT} = 1 - \frac{SQR}{SQT}$

### 13. Multiple lineare Regression

Lineares Regressionsmodell	$y = \beta_0 + \beta_1 \cdot x_1 + \beta_2 \cdot x_2 + \dots + \beta_p \cdot x_p + \varepsilon$
Lineare Regressionsschätzung	$\hat{y} = b_0 + b_1 \cdot x_1 + b_2 \cdot x_2 + \dots + b_p \cdot x_p$
Adjustiertes Bestimmtheitsmaß	$r^2_{adj} = 1 - \frac{\frac{SQR}{n-p-1}}{\frac{SQT}{n-1}} = 1 - \frac{SQR}{SQT} \cdot \frac{n-1}{n-p-1} = 1 - (1-r^2) \cdot \frac{n-1}{n-p-1}$
F-Test	$F = \frac{MQE}{MQR} = \frac{\frac{SQE}{p}}{\frac{SQR}{n-p-1}}$ <p>p Freiheitsgrade im Zähler und n – p – 1 Freiheitsgrade im Nenner</p>
t-Test	$t_j = \frac{b_j}{s_{b_j}}$ <p>n – p – 1 Freiheitsgrade</p>

Statistik für Wirtschaftswissenschaftler

Ein Lehr- und Übungsbuch für das Bachelor-Studium

Schuster, Th.; Liesen, A.

2017, XII, 261 S. 80 Abb., Softcover

ISBN: 978-3-662-49835-4