Statistische Formeln und Tabellen

Kompakt
für Wirtschafts-
wissenschaftler

Joseph Bleymüller
Rafael Weißbach

Vahlen

13. Auflage

Zum Inhalt

Dieses Standardwerk umfasst die zwei Teile **Statistische Formeln** und die für praktische Berechnungen benötigten **Statistischen Tabellen**. Vom Inhalt und Umfang entsprechen sie dem Rahmen einer typischen Bachelorausbildung in Statistik für Wirtschaftswissenschaftler. Damit eignen sie sich sehr gut zur Verwendung in statistischen Prüfungen und machen die Herausgabe gesonderter Klausurhilfsblätter weitgehend überflüssig.

Zu den Autoren

Prof. Dr. Josef Bleymüller war Direktor des Instituts für Ökonometrie und Wirtschaftsstatistik der Universität Münster.

Prof. Dr. Rafael Weißbach ist Inhaber des Lehrstuhls für Statistik und Ökonometrie an der Universität Rostock.

Bis zur 12. Auflage hat Dr. Günther Gehlert an diesem Buch mitgearbeitet.

Statistische Formeln und Tabellen

Kompakt für Wirtschaftswissenschaftler

von

Prof. Dr. Josef Bleymüller

und

Prof. Dr. Rafael Weißbach

13., überarbeitete Auflage

Verlag Franz Vahlen München

Das vorliegende Taschenbuch stellt eine Ergänzung des im gleichen Verlag erschienenen Lehrbuchs

„Statistik für Wirtschaftswissenschaftler“

dar. Im **ersten Teil** sind die wichtigsten **statistischen Formeln** aus dem oben genannten Buch zusammengestellt.

Der **zweite Teil** enthält die für praktische Berechnungen benötigten **statistischen Tabellen**, und zwar in einem im Rahmen des wirtschaftswissenschaftlichen Studiums gemeinhin benötigten Umfang.

Ihrer Anlage nach dürfte sich diese Formelsammlung gut zur Verwendung in statistischen Prüfungen eignen und die Herausgabe gesonderter Klausurhilfsblätter weitgehend überflüssig machen. Für die Programmbeschreibungen zu kommerziellen Statistik-Software-Paketen sei nun auf das Lehrbuch verwiesen. Dank schulden die Verfasser Herrn *Achim Dörre* und Herrn *Benjamin Strohner*.

Für das Lektorat sind wir Herrn Brunotte dankbar.

Münster und Rostock, Josef Bleymüller Rafael Weißbach
im Dezember 2014

Teil I

Statistische Formeln

Name	Kleinbuchstabe	Großbuchstabe
Alpha	α	A
Beta	β	B
Gamma	γ	Γ
Delta	δ	Δ
Epsilon	ε	E
Zeta	ζ	Z
Eta	η	H
Theta	θ oder ϑ	Θ
Jota	ι	I
Kappa	κ	K
Lambda	λ	Λ
Mü	μ	M
Nü	ν	N
Xi	ξ	Ξ
Omikron	o	O
Pi	π	Π
Rho	ϱ	P
Sigma	σ	Σ
Tau	τ	T
Ypsilon	υ	Υ
Phi	ϕ oder φ	Φ
Chi	χ	X
Psi	ψ	Ψ
Omega	ω	Ω

Symbole

Allgemeine Symbole

Symbol	Bedeutung
$a = b$	a ist gleich b
$a < b$	a ist kleiner als b
$a \leq b$	a ist kleiner oder gleich b
$a > b$	a ist größer als b
$a \geq b$	a ist größer oder gleich b
$a \approx b$	a ist ungefähr gleich b
$\sum_{i=1}^{n} x_i$	$x_1 + x_2 + \ldots + x_n$
$\prod_{i=1}^{n} x_i$	$x_1 \cdot x_2 \cdot \ldots \cdot x_n$
$\frac{dy}{dx} = f'(x)$	1. Ableitung
$\frac{\partial y}{\partial x}$	1. partielle Ableitung
$\int$	Integral
$\lvert x \rvert$	Absolutbetrag von x
$\lim_{x \to a} f(x)$	Grenzwert von $f(x)$ für $x \to a$
$\mathbf{A}'$	Transponierte der Matrix $\mathbf{A}$
$\mathrm{sgn}(x)$	Vorzeichen von x

Symbole der Mengenlehre

Symbol	Bedeutung
$\{a, b, c\}$	Menge, bestehend aus den Elementen a, b, und c
$x \in M$	x ist Element der Menge M
$x \notin M$	x ist nicht Element von M
$\{x \in M / x \text{ hat Eigenschaft } E\}$	Menge derjenigen Elemente von M, die die Eigenschaft E haben
$A \subset B$	A ist Teilmenge von M
$A \not\subset B$	A ist nicht Teilmenge von M
$\emptyset$	Leere Menge
$\mathfrak{P}(M)$	Potenzmenge von M, d. h. Menge aller Teilmengen von M
$A \cup B$	Vereinigungsmenge von A und B
$A \cap B$	Schnittmenge von A und B
$\overline{A}$	Komplement von A
$\mathbb{N}$	Menge der natürlichen Zahlen
$\mathbb{N}_0$	Menge der natürlichen Zahlen einschließlich 0
$\mathbb{Z}$	Menge der ganzen Zahlen
$\mathbb{Q}$	Menge der rationalen Zahlen
$\mathbb{R}$	Menge der reellen Zahlen
$\mathbb{R}^+$	Menge der nicht-negativen reellen Zahlen
(a, b)	$\{x \in \mathbb{R} / a < x < b\}$
$[a, b]$	$\{x \in \mathbb{R} / a \leq x \leq b\}$
$(a, b]$	$\{x \in \mathbb{R} / a < x \leq b\}$
$[a, b)$	$\{x \in \mathbb{R} / a \leq x < b\}$
$\lvert M \rvert$	Anzahl der Elemente von M

Symbole der Aussagenlogik

Symbol	Bedeutung
A	A ist eine Aussage, die *wahr* (w) oder *falsch* (f) sein kann.
$v(A)$	$v(A)$ wird als der Wahrheitswert der Aussage A bezeichnet; $v(A) = 1$ heißt, dass A *wahr* und $v(A) = 0$, dass A *falsch* ist.
$\neg A$	Die *Negation* $\neg A$ (bzw. $\overline{A}$) der Aussage A ist *wahr*, wenn A *falsch* ist, und *falsch*, wenn A *wahr* ist.
$A \wedge B$	Die *Konjunktion* $A \wedge B$ ist *wahr*, wenn beide Aussagen *wahr* sind, und *falsch*, wenn wenigstens eine der beiden Aussagen falsch ist.
$A \vee B$	Die *Disjunktion* $A \vee B$ ist *wahr*, wenn wenigstens eine der beiden Aussagen wahr ist, und *falsch*, wenn beide Aussagen falsch sind.
$A \Rightarrow B$	Die *Implikation* $A \Rightarrow B$ bedeutet: Wenn A wahr ist, dann ist auch B wahr. A wird als Voraussetzung (Prämisse), B als Folgerung (*Konklusion*) bezeichnet. $A \Rightarrow B$ ist *nur dann falsch*, wenn aus einer wahren Voraussetzung eine falsche Folgerung gezogen wird.
$A \Leftrightarrow B$	Die *Äquivalenz* $A \Leftrightarrow B$ bedeutet: Wenn A wahr ist, dann ist auch B wahr und umgekehrt. $A \Leftrightarrow B$ ist *nur dann falsch*, wenn eine der beiden Aussagen wahr und die andere falsch ist.
$\exists$	"Es gibt" (z. B.: $\exists x \in \mathbb{Q}\colon x^2 = 4$ heißt: Es gibt eine rationale Zahl x mit $x^2 = 4$).
$\forall$	"Für alle" (z. B.: $\forall x \in \mathbb{Q}\colon x^2 \geq 0$ heißt: Für alle rationalen Zahlen x gilt $x^2 \geq 0$).

Häufigkeiten

Kommt ein statistisches Merkmal in k verschiedenen

Merkmalsausprägungen $x_1, x_2, \ldots, x_k$

vor, für die bei insgesamt N Beobachtungen

absolute Häufigkeiten $h_1, h_2, \ldots, h_k \quad \text{mit} \quad \sum_{i=1}^{k} h_i = N$

beobachtet werden, so ergeben sich daraus entsprechende

relative Häufigkeiten $f_1, f_2, \ldots, f_k \quad \text{mit} \quad \sum_{i=1}^{k} f_i = 1$

und $f_i = \dfrac{h_i}{N} \quad (i = 1, \ldots, k).$

Summenhäufigkeiten

Bei *ordinal- und metrischskalierten Merkmalen* ergeben sich durch Summierung über alle Merkmalsausprägungen x_j mit $x_j \leq x_i$

absolute Summenhäufigkeiten $H_i = \sum_{x_j \leq x_i} h_j \quad (i = 1, \ldots, k)$

und

relative Summenhäufigkeiten $F_i = \sum_{x_j \leq x_i} f_j \quad (i = 1, \ldots, k)$

$$F_i = \frac{H_i}{N} \quad (i = 1, \ldots, k).$$

Arithmetisches Mittel μ

Bei N *Einzelwerten*

$$a_1, a_2, \ldots, a_N$$

ist das arithmetische Mittel definiert als

$$\mu = \frac{1}{N} \sum_{i=1}^{N} a_i \, .$$

Bei einer *Häufigkeitsverteilung* mit k verschiedenen Werten

$$x_1, x_2, \ldots, x_k$$

ergibt sich das (gewogene) arithmetische Mittel zu

$$\mu = \frac{1}{N} \sum_{i=1}^{k} x_i h_i \quad \text{bzw.} \quad \mu = \sum_{i=1}^{k} x_i f_i \, .$$

Bei einer *Häufigkeitsverteilung klassifizierter Daten* ergibt sich mithilfe der Klassenmitten

$$x'_1, x'_2, \ldots, x'_k$$

näherungsweise

$$\mu = \frac{1}{N} \sum_{i=1}^{k} x'_i h_i \quad \text{bzw.} \quad \mu = \sum_{i=1}^{k} x'_i f_i \, .$$

Für eine Grundgesamtheit, die aus k *Teilgesamtheiten* mit den Umfängen $N_1, N_2, \ldots, N_k$ und den arithmetischen Mitteln $\mu_1, \mu_2, \ldots, \mu_k$ besteht, ergibt sich das arithmetische Mittel zu

$$\mu = \sum_{i=1}^{k} \frac{N_i}{N} \mu_i \quad \text{bzw.} \quad N = \sum_{i=1}^{k} N_i \, .$$

Median Me und Quartile Q_1, Q_2 und Q_3

Zunächst werden die *Einzelwerte* $a_1, a_2, \ldots, a_N$ so umgeordnet, dass gilt

$$a_{[1]} \leq a_{[2]} \leq \ldots \leq a_{[N]} \, .$$

Dann ist bei ungeradem N

$$\mathrm{Me} = a_{\left[\frac{N+1}{2}\right]}$$

und bei geradem N

$$\mathrm{Me} = \frac{1}{2}\left(a_{\left[\frac{N}{2}\right]} + a_{\left[\frac{N}{2}+1\right]}\right) .$$

Für großes N kann als Median der *größte* Merkmalswert $a_{[k]}$ verwendet werden, für den

$$\mathrm{F}(a_{[k]}) \leq 0{,}5$$

gilt, wobei $\mathrm{F}(a_{[k]})$ der Wert der Summenhäufigkeitsfunktion für $a_{[k]}$ ist.

Analog ist das 1. Quartil Q_1 der größte Merkmalswert $a_{[j]}$, für den

$$\mathrm{F}(a_{[j]}) \leq 0{,}25$$

und das 3. Quartil Q_3 der *größte* Merkmalswert $a_{[l]}$, für den

$$\mathrm{F}(a_{[l]}) \leq 0{,}75$$

gilt.

Bei *klassifizierten Daten* ergibt sich der feinberechnete Median aus der Klassenuntergrenze x_i^u und der Klassenobergrenze x_i^o derjenigen Klasse i, in der die Summenhäufigkeitsfunktion den Wert 0,5 erreicht:

$$\mathrm{Me} = x_i^u + \frac{0{,}5 - \mathrm{F}(x_i^u)}{\mathrm{F}(x_i^o) - \mathrm{F}(x_i^u)}(x_i^o - x_i^u) .$$

In analoger Weise ergeben sich die feinberechneten Quartile zu

$$Q_1 = x_k^u + \frac{0{,}25 - \mathrm{F}(x_k^u)}{\mathrm{F}(x_k^o) - \mathrm{F}(x_k^u)}(x_k^o - x_k^u) ,$$

wobei k diejenige Klasse ist, in der die Summenhäufigkeitsfunktion den Wert 0,25 erreicht, und

$$Q_3 = x_l^u + \frac{0{,}75 - \mathrm{F}(x_l^u)}{\mathrm{F}(x_l^o) - \mathrm{F}(x_l^u)}(x_l^o - x_l^u) ,$$

wobei l diejenige Klasse ist, in der die Summenhäufigkeitsfunktion den Wert 0,75 erreicht.

Q_2 entspricht dem Median Me.

Modus Mo

Der Modus Mo ist als die häufigste Merkmalsausprägung definiert. Bei klassifizierten Daten wird als Modus die Klassenmitte der Klasse mit der größten Säulenhöhe im Histogramm gewählt.