



TQU Verlag

Die Bivariate Häufigkeit

Probieren und Studieren

Die Bivariate Häufigkeitsverteilung

Abhängigkeiten und Zusammenhänge erkennen und darstellen

[Autor: Dr. Konrad Reuter](#)

In der deskriptiven Statistik wird das Datenmaterial aus Erhebungen und Beobachtungen meist systematisch so zusammengefasst, dass für die Beobachtungswerte (x) angegeben wird, in welchen Häufigkeiten (y) ein einfaches oder zusammengesetztes Merkmal in allen Klassen auftritt. Diese Zusammenstellung der Beobachtungshäufigkeiten heißt Häufigkeitsverteilung.

Dividiert man dabei die einzelnen Häufigkeiten durch die Gesamtzahl der Beobachtungen, so erhält man die relative Häufigkeitsverteilung.

Die Häufigkeitsverteilung gibt an, bei wie vielen Untersuchungsobjekten unterschiedliche Ausprägungen einer oder mehrerer Variablen beobachtet wurden.

Hat eine Verteilung lineare, nicht-Zahlvariable, in beiden Richtungen von diskreten (unstetigen) Häufigkeitsverteilung.

Bei einer (theoretischen) Anzahl von endlich vielen Variablen spricht man von einer kontinuierlichen (stetigen) Häufigkeitsverteilung.

Nach der Zahl der beobachteten Variablen unterscheidet man univariate (eindimensionale), bivariate (zweidimensionale) und multivariate (mehrdimensionale) Häufigkeitsverteilungen.

Je nach der Form der grafischen Darstellung von Häufigkeitsverteilungen unterscheidet man eingipflige (unimodale) und mehrgipflige (multimodale) Verteilungen,

linksschiefe, rechtsschiefe oder symmetrische Verteilungen, links- und rechtssteile Verteilungen, dreieckige, rechteckige Verteilungen sowie glockenförmige und u-förmige Häufigkeitsverteilungen.

Häufigkeitsverteilungen lassen sich gut in einem zweidimensionalen Histogramm oder in einer Kreuztabellierung der Werte von zwei Variablen zu visualisieren.

Besteht zwischen zwei Variablen (Merkmalen) an einem Produkt ein Zusammenhang? Wenn ja, wie stark ist dieser Zusammenhang?

Kann man aus den beobachteten Ereignissen eine Hypothese ableiten? Wie brauchbar ist diese Hypothese, um Entscheidungen zu treffen?

Wie entdeckt man Ausreißer im multivariaten Fall?

Lassen sich für den multivariaten Fall Regelkarten anwenden und Prozessfähigkeitsindizes angeben?

Hat dies Vorteile gegenüber der Betrachtung von nur einzelnen Merkmalen ohne Berücksichtigung von Korrelationen?

Die Multivariate Statistik kann bei diesen und ähnlichen Fragestellungen eine äußerst hilfreiche Methode sein.

Dieses APP ist so gestaltet, dass Sie interaktiv die Grundlagen der Auswertung der Analysedaten und deren wichtigsten Kenngrößen verstehen und anwenden können.

Dieses QUALITY APP liefert dem Qualitäts- und dem Produktionsmanagement wertvolle Unterstützung bei der statistischen Bewertung von Analysedaten.

Die Applikation ist im Excel-Format und kann sofort eingesetzt werden.

Ansprechpartner: Dr. Konrad Reuter

Telefon: 0171/6006604

TQU Verlag, Magirus-Deutz-Straße 18, 89077 Ulm Deutschland, Telefon 0731/14660200, verlag@tqu-group.com, www.tqu-verlag.com

QUALITY APPS Applikationen für das Qualitätsmanagement

Lizenzvereinbarung

Dieses Produkt "Die Bivariate Häufigkeitsverteilung" wurde vom Autor Dr. Konrad Reuter mit großem Aufwand und großer Sorgfalt hergestellt. Dieses Werk ist urheberrechtlich geschützt (©). Die dadurch begründeten Rechte, insbesondere die der Weitergabe, der Übersetzung, des Kopierens, der Entnahme von Teilen oder der Speicherung bleiben vorbehalten.

Bei Fehlern, die zu einer wesentlichen Beeinträchtigung der Nutzung dieses Softwareproduktes führen, leisten wir kostenlos Ersatz. Beschreibungen und Funktionen verstehen sich als Beschreibung von Nutzungsmöglichkeiten und nicht als rechtsverbindliche Zusicherung bestimmter Eigenschaften. Wir übernehmen keine Gewähr dafür, dass die angebotenen Lösungen für bestimmte vom Kunden beabsichtigte Zwecke geeignet sind.

Sie erklären sich damit einverstanden, dieses Produkt nur für Ihre eigene Arbeit und für die Information innerhalb Ihres Unternehmens zu verwenden. Sollten Sie es in anderer Form, insbesondere in Schulungs- und Informationsmaßnahmen bei anderen Unternehmen (Beratung, Schulungseinrichtung etc.) verwenden wollen, setzen Sie sich unbedingt vorher mit uns wegen einer möglichen Lizenzvereinbarung in Verbindung. Dieses Produkt wird kontinuierlich weiterentwickelt. Bitte melden Sie sich, wenn Sie ein Update wünschen.

Alle Ergebnisse basieren auf den vom Autor eingesetzten Formeln und müssen vom Anwender sorgfältig geprüft werden. Die berechneten Ergebnisse sind als Hinweise und Anregungen zu verstehen.

Wir wünschen viel Spaß und Erfolg mit dieser Applikation

TQU Verlag, Magirus-Deutz-Straße 18, 89077 Ulm Deutschland, Telefon 0731/14660200, verlag@tqu-group.com, www.tqu-verlag.com

Anwendungshinweise

1. Statistik

Multivariate Statistik beginnt mit der Betrachtung zweier Merkmale.
Die Darstellung der Häufigkeit zweier Merkmale ist eine Unterstützung der Interpretation der damit verbundenen Kennwerte.
Räumliche Diagramme können vom Benutzer in EXCEL gedreht und somit von allen Seiten betrachtet werden.
(Einfach an einem Eckpunkt mit der Maus "anfassen" und ausprobieren).
Wird jede Variable einzeln betrachtet erhält man die sog. Randverteilungen.
Zusätzliche Informationen erhält man aus den Höhenlinien des "Häufigkeitsgebirges".

Blattname

Histogramm

Klassierung

Häufigkeitslinien

Diese Schnitte stellen Linien gleicher Häufigkeit dar und werden sich oft als Ellipsen präsentieren.

Eine wichtige Möglichkeit, Wahrscheinlichkeiten für Entscheidungen z. B. für Prozessrisikoprüfung, ist die Schnittmenge der sich überlagernde "Wahrscheinlichkeitsmasse" äquivalent zu 6 Sigma ist (99,73%).

Als Kennwerte sind die Kovarianz und die Korrelationskoeffizienten zuzunehmen zu den Kennwerten der einzelnen Merkmale zu ermitteln.

Die Formeln sind in der EXCEL-Hilfe zu finden.

Die Varianzen und Kovarianzen werden als Matrix zusammengefasst.

Diese Form ist eine Vorbereitung für die weiteren Anwendungen der multivariaten Statistik auch für mehr als 2 Variable.

Die Kovarianzmatrix ist quadratisch und symmetrisch. Auf der Hauptdiagonalen sitzen die jeweiligen Varianzen.

Eine wichtige Umformung der Kovarianzmatrix ist die Kehrmatrix (inverse Matrix).

Für diese Operation stellt EXCEL ebenfalls eine Matrixformel bereit.

Die Hauptdiagonale der Korrelationsmatrix ist jeweils mit 1 besetzt (Effekt der Normierung auf die jeweilige Varianz).

In der Literatur wird diese Matrix wegen ihrer Symmetrie meist nur halb dargestellt.

Die Berechnung der Determinante der Kovarianzmatrix für die weitere Verwendung kann ebenfalls mit EXCEL erfolgen.

Kennwerte

2. Anwendung

Im Blatt Merkmal sind zugehörigen Angaben einzutragen.

In das Blatt Daten sind die Daten einzutragen oder aus Anwendungen zu übernehmen.

Merkmale

Daten

Überprüfen Sie unbedingt den Datensatz auf Datenfehler!!!

Die Stellenanzahl im Blatt Kennwerte ist nicht begrenzt.

Die Anzahl der Klassen ist auf 14 beschränkt und kann bei Bedarf reduziert werden.

Die Klassierung sollte nicht mit zu wenigen Daten vorgenommen werden!

Die Angabe von Toleranzen ist nicht zwingend erforderlich.

Klassierung

Merkmale

3. EXCEL

Die EXCEL-Lösung stützt sich auf folgende Prinzipien:

Funktionelle Aufteilung auf verschiedene Blätter.

Optische Hervorhebung von Zellen in Abhängigkeit von ihrer Funktion.

Kommentierung wesentlicher Zellen

Vergaben von Namen für Variable.

Die Liste muss nach Änderungen mit F3 manuell aktualisiert werden.

Textliche Kommentierungen in rot kursiv.

Bezug auf Zellen mit Funktionen

Reagieren auf Bedingungen /Verzweigungen

Ausblenden von Zellinhalten, die nicht zutreffend sind

Erzwingung einer geeigneten Zahlenformatierung im Protokoll mit vorgegebener Stellenzahl.

Merkmal

Eingabe Daten

errechnete Werte

Bezeichnungen

Namen

Bemerkungen

VERGLEICH(;;0)

INDEX(;;)

INDIREKT()

ISTLEER()

ISTZAHL()

WENN(;;)

Verknüpfung von Zellinhalten über "&"

Zellen werden in Berechnungsblättern und Protokollen nicht über "Zellen verbinden" formatiert!

Als Lösung dient die Formatierung schmaler Spalten und die Formatierung benachbarter Zellen mit "Über Auswahl zentrieren".

Die Nachteile verbundener Zellen sind damit vermieden..

FEST(;
= "text1"&BEZUG

Ausblenden von Blättern, die nicht ständig gebraucht werden.

Blattschutz gegen versehentliches Überschreiben, bitte nutzerintern anpassen.

QUALITY APPS im TQU VERLAG

Die Datei wird in dem Wasserzug für II

Matrixfunktionen (VEGLICH) wird die Zughörigkeit zu den Klassen festgelegt.

Die Verknüpfung zu den gemeinsamen Häufigkeiten erfolgt über eine Matrixfunktion.

`{=SUMME(WENN(Y&F&A<N(H Y L 2 1;)))}`

Matrixfunktionen werden aus "normalen" Funktionen insofern erzeugt, als beim

Eingeben die Funktion mit "strg", "shift" und "enter" gleichzeitig (beidhändig arbeiten) abgeschlossen wird.

Die geschweiften Klammern werden bei dieser Prozedur von EXCEL automatisch eingefügt, sind also nicht mit der Tastatur erzeugbar!

Matrixfunktionen arbeiten gewissermaßen als "Schleife" und ersparen so häufig eine sonst notwendige VBA Prozedur!

(Achtung, die Berechnung dauert allerdings ein wenig länger!).

Das Diagramm ist ein räumliches Säulendiagramm mit Null Abständen.

Zusätzlich werden die jeweiligen Randverteilungen dargestellt.

Histogramm

Bereitstellung von Testdaten zu Überprüfung der Funktion der Datei,

Protokoll bitte in geeigneten Abständen aktualisieren.

Testdaten

Anmerkung:

Eine EXCEL Ermittlung der Häufigkeiten wäre auch mit der Pivottable möglich.

Allerdings gibt es hierbei Schwierigkeiten mit nicht besetzten Klassen.

4. Quellen

Rinne "Taschenbuch Statistik", Verlag Harri Deutsch, 2003

Jahn/Braun "Praxisleitfaden Qualität", Hanserverlag 2006

Bezeichnung	Testteil	
Zeichnungsnummer	4711	
Merkmale	Maß1	Maß2
Auflösung PM	0,001	0,001
Stellenzahl	4	4
Nennmaß	1,000	1,000
oberes Abmaß	0,250	0,250
unteres Abmaß	-0,250	-0,250
OGW	1,250	1,250
UGW	0,750	0,750

TQU Verlag, Magirus-Deutz-Straße 18, 89077 Ulm Deutschland, Telefon 0731/14660200, verlag@tqu-group.com, www.tqu-verlag.com

Maß1	Maß2
1,068	1,074
1,083	1,047
1,092	1,017
1,169	1,097
1,109	1,081
1,033	1,000
1,023	1,070
1,008	1,008
1,026	1,033
1,014	1,013
1,031	1,048
1,029	0,999
1,055	1,316
1,112	1,032
1,072	1,080
1,002	1,025
0,970	0,972
1,027	1,007
1,008	1,034
0,967	0,964
1,148	1,083
1,076	1,040
0,998	0,991
0,974	0,952
1,013	1,015
1,025	1,038
1,076	1,092
1,157	1,098
1,057	1,047
1,097	1,063
1,275	1,122
1,082	1,030
1,098	1,056
1,100	1,098
1,036	1,043
1,073	1,021
1,057	1,035
1,144	1,058
1,094	1,031
1,101	1,093
0,977	1,020
1,029	1,016
1,021	0,954
1,148	1,062
1,059	1,045

Hier die Daten eingeben oder einkopieren
bis n=2000 vorbereitet

QUALITY APPs im TQU VERLAG
www.tqu-verlag.de

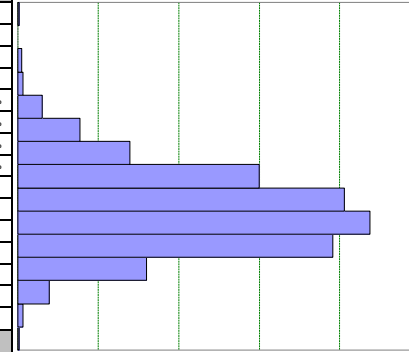
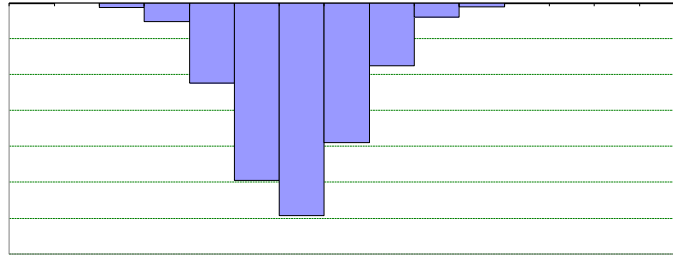
TQU Verlag, Magirus-Deutz-Straße 18, 89077 Ulm Deutschland,
Telefon 0731/14660200, verlag@tqu-group.com, www.tqu-verlag.com

Kennwerte		Maß1	Maß2
Mittelwerte	\bar{x}_j	1,0534	1,0274
Min		0,8570	0,7940
Max		1,3440	1,3410
Range	R_j	0,487	0,547
Varianzen	s_j^2	0,003246568	0,002564237
Standardabweichungen	s_j	1,0534	1,0274
Kovarianzmatrix \mathbf{A}		0,003246568	0,002299766
Korrelationskoeffizient	r	0,797548774	
Anzahl Wertepaare	n	1640	
www.tqu-verlag.de			
Kovarianzmatrix \mathbf{A}	Maß1	0,003246568	0,002299766
	Maß2	0,002299766	0,002564237
inverse Kovarianzmatrix \mathbf{A}^{-1}		844,5977224	-757,4874501
		-757,4874501	1069,34111
Korrelationsmatrix	Maß1	1	0,797548774
	Maß2	0,797548774	1
Determinante der Kovarianzmatrix		3,03605E-06	

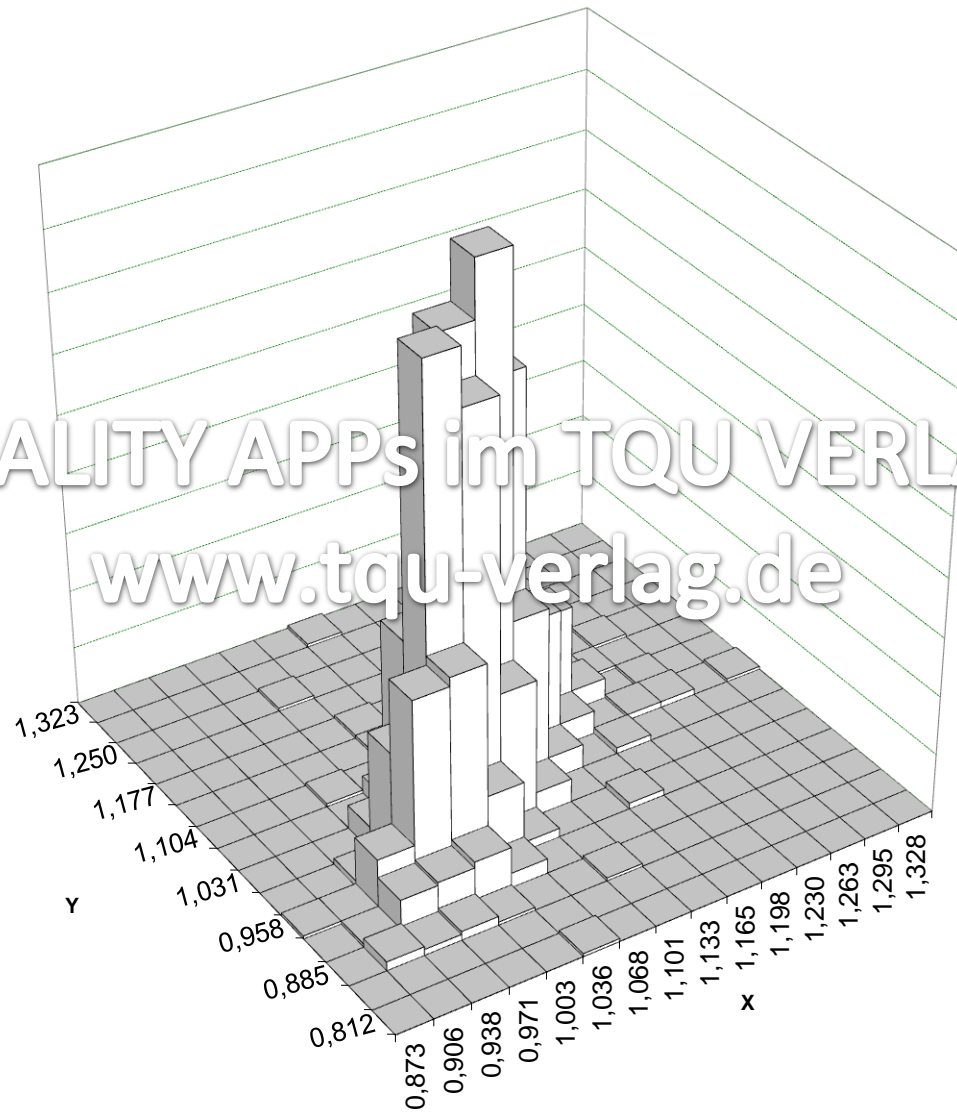
H_X	H_Y
7	8
7	7
8	7
10	9
8	8
6	6
7	8
5	6
6	7
6	6
7	15
8	7
7	8
5	7
4	5
6	6
5	7
9	8
7	7
5	6
4	5
5	7
6	7
7	9
10	9
7	7
8	8
13	9
7	7
8	8
8	9
4	7
6	7
6	5
9	8
7	7
9	8

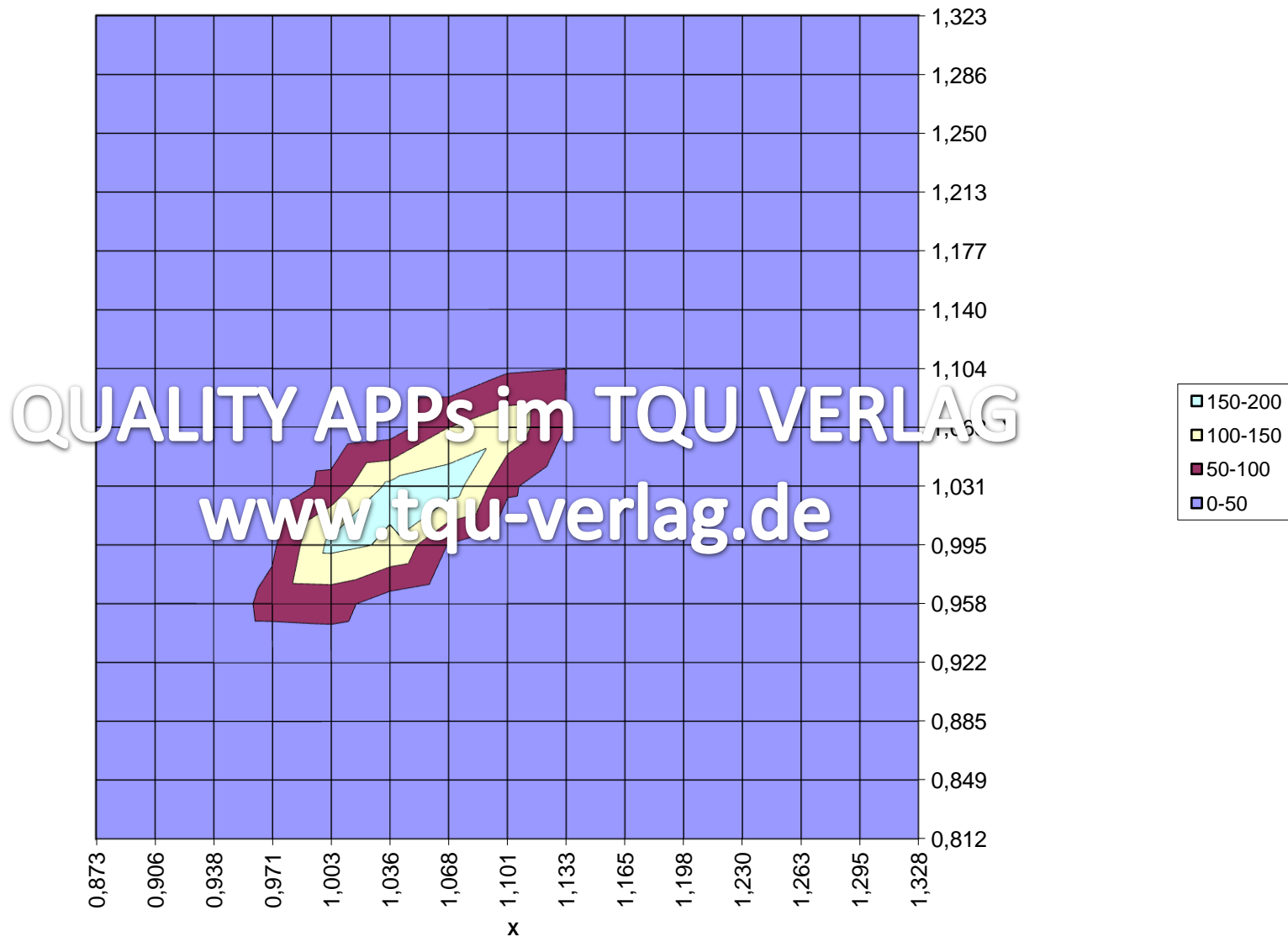
Nr	KL_X	KL_Y	Klassenmiten	
			KL_X	KL_Y
1	0,8570	0,7940		
2	0,8895	0,8305	0,873	0,812
3	0,9219	0,8669	0,906	0,849
4	0,9544	0,9034	0,938	0,885
5	0,9869	0,9399	0,971	0,922
6	1,0193	0,9763	1,003	0,958
7	1,0518	1,0128	1,036	0,995
8	1,0843	1,0493	1,068	1,031
9	1,1168	1,0857	1,100	1,068
10	1,1493	1,1222	1,133	1,104
11	1,1817	1,1587	1,165	1,140
12	1,2142	1,1951	1,200	1,176
13	1,2466	1,2316	1,230	1,211
14	1,2791	1,2681	1,263	1,250
15	1,3115	1,3045	1,295	1,286
16	1,3440	1,3410	1,328	1,323
Anzahl Klassen <=15				
15	15			

Klassenmiten	0,812	0,849	0,885	0,922	0,958	0,995	1,031	1,068	1,104	1,140	1,177	1,213	1,250	1,286	1,323	Summe	h rel X
0,873	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0,1%
0,906	2	0	0	3	1	1	0	0	0	0	0	0	0	0	0	5	0,3%
0,938	3	0	0	2	11	17	2	0	0	0	0	0	0	0	0	32	2,0%
0,971	4	0	0	3	12	66	41	8	0	1	0	0	0	0	0	131	8,0%
1,003	5	0	0	1	14	69	163	67	6	0	0	0	0	1	0	321	19,6%
1,036	6	1	0	0	5	24	144	161	20	2	1	1	0	0	0	359	21,9%
1,068	7	0	0	0	0	3	47	180	99	3	0	0	0	0	1	333	20,3%
1,100	8	0	0	1	0	0	8	60	133	42	2	0	0	0	0	246	15,0%
1,133	9	0	0	0	0	0	0	8	48	49	8	1	0	0	0	114	7,0%
1,165	10	0	0	0	0	2	0	0	8	38	15	0	0	0	0	63	3,8%
1,198	11	0	0	0	0	0	0	2	1	8	6	6	0	0	1	25	1,5%
1,230	12	0	0	0	0	0	0	0	2	0	1	0	2	0	0	5	0,3%
1,263	13	0	0	0	0	0	0	0	2	1	0	1	0	0	0	4	0,2%
1,295	14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0%
1,328	15	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0,1%
Summe	1	0	10	43	183	405	486	320	144	33	9	2	1	1	2	1640	
h rel Y	0,1%	0,0%	0,6%	2,6%	11,2%	24,7%	29,6%	19,5%	8,8%	2,0%	0,5%	0,1%	0,1%	0,1%	0,1%	100,0%	



QUALITY APPs im TQU VERLAG
www.tqu-verlag.de





Maß1	Maß2
1,068	1,074
1,083	1,047
1,092	1,017
1,169	1,097
1,109	1,081
1,033	1,006
1,063	1,073
1,008	1,008
1,026	1,033
1,014	1,013
1,031	1,048
1,029	0,999
1,055	1,316
1,112	1,032
1,072	1,08
1,002	1,025
0,97	0,972
1,027	1,007
1,008	1,034
0,967	0,964
1,148	1,083
1,076	1,04
0,998	0,991
0,974	0,952
1,013	1,015
1,025	1,038
1,076	1,092
1,157	1,098
1,057	1,047
1,097	1,063
1,275	1,122
1,082	1,03
1,098	1,056
1,1	1,098

Statistische Werte Soll	Ist	Vergleich
Var1 0,003246568	0,003246568	1
Var2 0,002564237	0,002564237	1
KOVAR 0,002299766	0,002299766	1

i.O.

Datum	Tester	Ergebnis
14.04.2011	Reuter	i.O.

Die Testdaten kopieren und in Blatt "Daten" manuell als "Wert" einkopieren.

TQU Verlag, Magirus-Deutz-Straße 18, 89077 Ulm Deutschland, Telefon 0731/14660200, verlag@tqu-group.com, www.tqu-verlag.com