



TQU VERLAG

**Shapiro-Wilk
Test**

Testen und Anwenden

Shapiro-Wilk Test

[Autor: Dr. Konrad Reuter](#)

Der Shapiro-Wilk-Test ist ein statistischer Signifikanztest, der die Hypothese überprüft, dass die zugrunde liegende Grundgesamtheit einer Stichprobe normalverteilt ist. Das Testverfahren wurde 1965 von dem Amerikaner Samuel Shapiro und dem Kanadier Martin Wilk veröffentlicht und ist das Ergebnis ihrer ursprünglichen Idee, die graphischen Informationen der Analyse auf Normalverteilung mittels Normalwahrscheinlichkeitsplot in einer Kennzahl zusammenzufassen.

Die Nullhypothese H_0 nimmt an, dass eine Normalverteilung der Grundgesamtheit vorliegt. Demgegenüber unterstellt die Alternativhypothese H_1 , dass keine Normalverteilung gegeben ist. Wenn der Wert der Teststatistik W größer ist als der kritische Wert W_{kritisch} , wird die Nullhypothese nicht abgelehnt und es wird angenommen, dass eine Normalverteilung vorliegt. Wird alternativ der p-Wert des Tests ermittelt, so wird die Nullhypothese in der Regel nicht abgelehnt, wenn der p-Wert größer ist als das festgelegte Signifikanzniveau α .

Der Shapiro-Wilk Test dient zur Testung von Daten aus kleinen Stichprobenumfängen auf Normalverteilung. Der Test kann zum Überprüfen von univariaten Stichproben eingesetzt werden. Der Test sind in gängigen CAQ Systemen und statistischen Softwarepaketen meist enthalten.

Die Anwendung in EXCEL stützt sich auf die entsprechenden Tabellen zum Verfahren bis $n = 50$.

Ein weiterer empfehlenswerter Test auf Normalverteilung ist der Epps-Pulley Test, der auf der SixSigma CD des Verlages verfügbar ist.

Dieses QUALITY APP liefert dem Statistiker wertvolle Unterstützung beim Test auf Normalverteilung.

Die QUALITY Applikation ist im Excel-Format (ab 2013) 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 "Shapiro-Wilk Test" 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 entsprechenden Vereinbarung in Verbindung. Unsere Produkte werden 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

QUALITY APPS Applikationen für das Qualitätsmanagement

Anwendungshinweise

1. Statistik

Blatt

Anpassungstests auf Verteilungsmodelle gehören zu den Standardverfahren statistischer Analysen.

Insbesondere zur Anpassung von Daten an die Normalverteilung sind eine Reihe von Tests entwickelt worden. Die Eigenschaftend der Tests sind in der Spezialliteratur beschrieben.

Die Hypothese H_0 behauptet: die Daten stammen aus einer Normalverteilung.

Die Alternativhypothese H_1 behauptet: die Daten stammen nicht aus einer Normalverteilung.

Einfache Verfahren für eine schnelle Übersicht sind mit der Verfügbarkeit von Software eher in den Hintergrund getreten.

Die Anwendbarkeit einzelner Tests richtet sich auch nach dem Stichprobenumfang. Bis etwa $n = 50$ werden Einzelwerte verwendet. Darüber können auch klassierte Daten getestet werden.

Der Shapiro-Wilk Test gehört zur Klasse der Regressions- und Korrelationstests.

Für die Durchführung müssen die Daten aufsteigend geordnet werden.

Berechnung

Die Teststatistik W benötigt auch die Berechnung der Summe der quadratischen Abweichungen SS .

Weiterhin wird ein Faktor b berechnet.
$$b = \sum_{i=1}^m a_i (x_{n+1-i} - x_i)$$

Die m Differenzen in diesem Ausdruck erhält man aus den jeweils paarweise $n/2$ absteigend und aufsteigend angeordneten Daten.

Die Faktoren a_i sind den Tabellen der DIN ISO 5479 entnommen (ausgeblendet).

$$W = b^2 / SS$$

Die Zuordnung der Wahrscheinlichkeit p und damit der Bewertung der Signifikanz erfordert eine weitere Tabelle (ebenfalls ausgeblendet).

Die Ablesung erfordert in der Regel eine (lineare) Interpolation aus den Tabellenwerten.

Die vorliegenden Tabellen erlauben einen Stichprobenumfang von max, $n = 50$.

Die Berechnung gibt einen p Wert zurück, der mit den üblichen Signifikanzschwellen verglichen werden kann.

2. Anwendung

Im Blatt Merkmal sind gewünschten Angaben einzutragen.
In das Blatt Daten sind die Daten einzutragen oder aus Anwendungen zu übernehmen.
Das Ergebnis wird als Wahrscheinlichkeit p ausgegeben und die Signifikanz bewertet.
Überprüfen Sie unbedingt den Datensatz auf Ausreißer oder andere Datenfehler!!!
Weitere Blätter enthalten Zwischenrechnungen oder dienen der Information.

Merkmal
Daten

Daten
Berechnung

3. EXCEL

Die EXCEL-Lösung stützt sich auf folgende Prinzipien:
Funktionelle Aufteilung auf verschiedene Blätter.
Optische Hervorhebung von Zellen in Abhängigkeit von ihrer Funktion.

Vergaben von Namen für Variable
Bezug auf Zellen mit Funktionen

Kommentierung wesentlicher Zellen
Reagieren auf Bedingungen /Verzweigungen

Ausblenden von Zellinhalten, die nicht zutreffend sind
Erzwingung einer geeigneten Zahlenformatierung im Ergebnis mit vorgegebener Stellenzahl.

Blattschutz gegen versehentliches Überschreiben (firmenintern anpassen).

Bereitstellung von Testdaten zu Überprüfung der Funktion der Datei.
Ergebnis bitte manuell eintragen.
Die Pflege einer Logdatei für die Änderungen ist sehr zu empfehlen.

Merkmal

| |
|------------------|
| Eingabe Daten |
| errechnete Werte |
| Bezeichnungen |

Namen

| |
|----------------|
| VERGLEICH(;;0) |
| INDEX(;;) |
| SVERWEIS(;;0) |
| ISTLEER() |
| ISTZAHL() |
| WENN(;;) |
| FEST(;) |
| |

4. Quellen

Rinne: Taschenbuch der Statistik, Verlag Harry Deutsch, Frankfurt am Main 2003
Sachs, L.: Taschenbuch der Statistik, Springer Verlag
ISO 5479 Tests auf Abweichung von der Normalverteilung

STATGRAPHICS Centurion XVI

Validierung

Änderungen

| | | | |
|------------|----------------|------------|----|
| Allgemein | Datum | 18.08.16 | |
| | Verantwortlich | Reuter | |
| | Abt./Kst. | QM | |
| | Prüfort | Leipzig | |
| Prüfmittel | Bezeichnung | Messlatte | |
| | Ident-Nummer | 0815 | |
| | Messbereich | 0 - 100 | cm |
| | Auflösung | 1,000 | cm |
| Merkmal | Bezeichnung | Population | |
| | Ident-Nummer | 4711 | |
| | Merkmal | Größe | |
| | Maßeinheit | cm | |
| | Sollmaß | 50,000 | cm |
| | OSG | 90,000 | cm |
| | USG | 20,000 | cm |
| | Toleranz | 70,000 | cm |

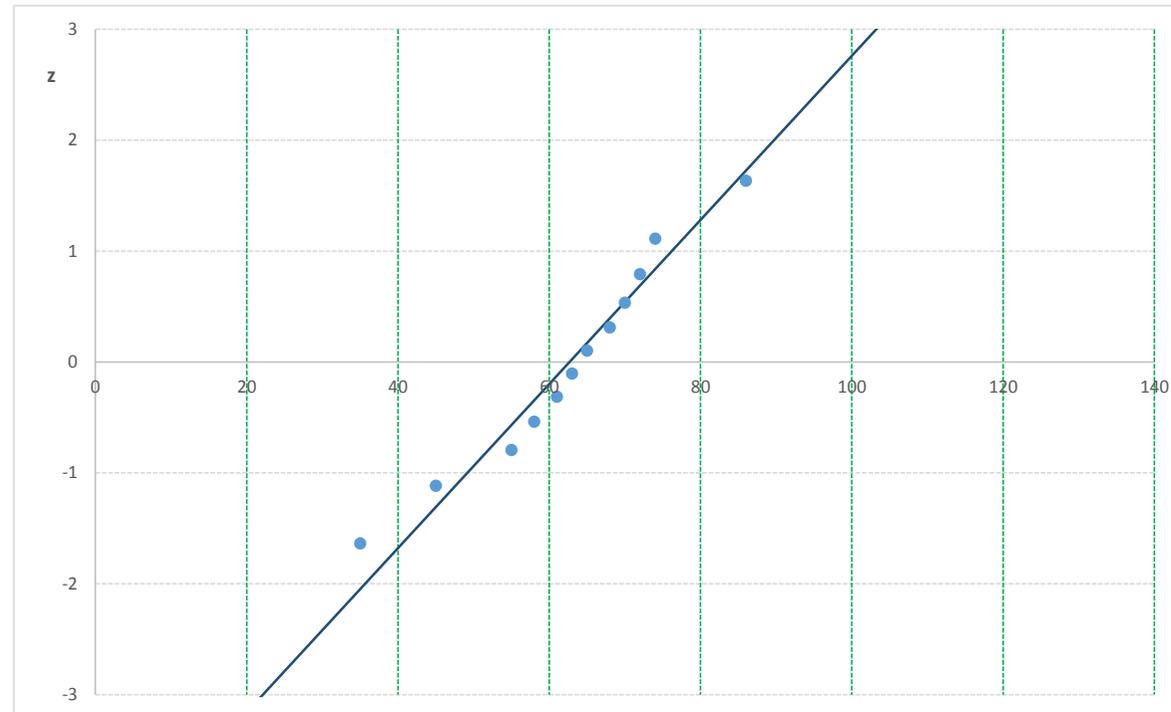
Stellen 0

Hier Daten eingeben oder einkopieren

| | Daten |
|----|-------|
| 1 | 65 |
| 2 | 61 |
| 3 | 63 |
| 4 | 86 |
| 5 | 70 |
| 6 | 55 |
| 7 | 74 |
| 8 | 35 |
| 9 | 72 |
| 10 | 68 |
| 11 | 45 |
| 12 | 58 |
| 13 | |
| 14 | |
| 15 | |
| 16 | |
| 17 | |
| 18 | |
| 19 | |
| 20 | |
| 21 | |
| 22 | |
| 23 | |
| 24 | |
| 25 | |
| 26 | |
| 27 | |
| 28 | |
| 29 | |
| 30 | |
| 31 | |
| 32 | |
| 33 | |
| 34 | |
| 35 | |
| 36 | |
| 37 | |
| 38 | |
| 39 | |

| | | | |
|--------------------|-------|--------|-------|
| Mittelwert | 62,67 | Median | 64,00 |
| Standardabweichung | 13,51 | Max | 86,00 |
| Schiefe | -0,52 | Min | 35,00 |
| Kurtosis | 0,75 | R | 51,00 |

| | | |
|--------------------|-------------------|--------|
| Ergebnis | <i>n</i> | 12 |
| Wahrscheinlichkeit | <i>p</i> | 0,8737 |
| Signifikanz | Nicht signifikant | |



| | | |
|---|---|---|
| XYZ GmbH | Shapiro-Wilk Test | FB ##-## Freigabe ## |
| Datum 18. Aug. 16 | Bearbeiter Reuter | Abt./Kst. QM |
| Prüfmittel Bezeichnung Messlatte Ident-Nummer 0815 Messbereich 0 - 100 cm Auflösung 1 cm rel.Auflösung 1,4% | Teil Bezeichnung Population Ident-Nummer 4711 Anzahl n 12 | Merkmal Bezeichnung Größe Sollmaß 50 cm OSG 90 cm USG 20 cm T 70 cm |

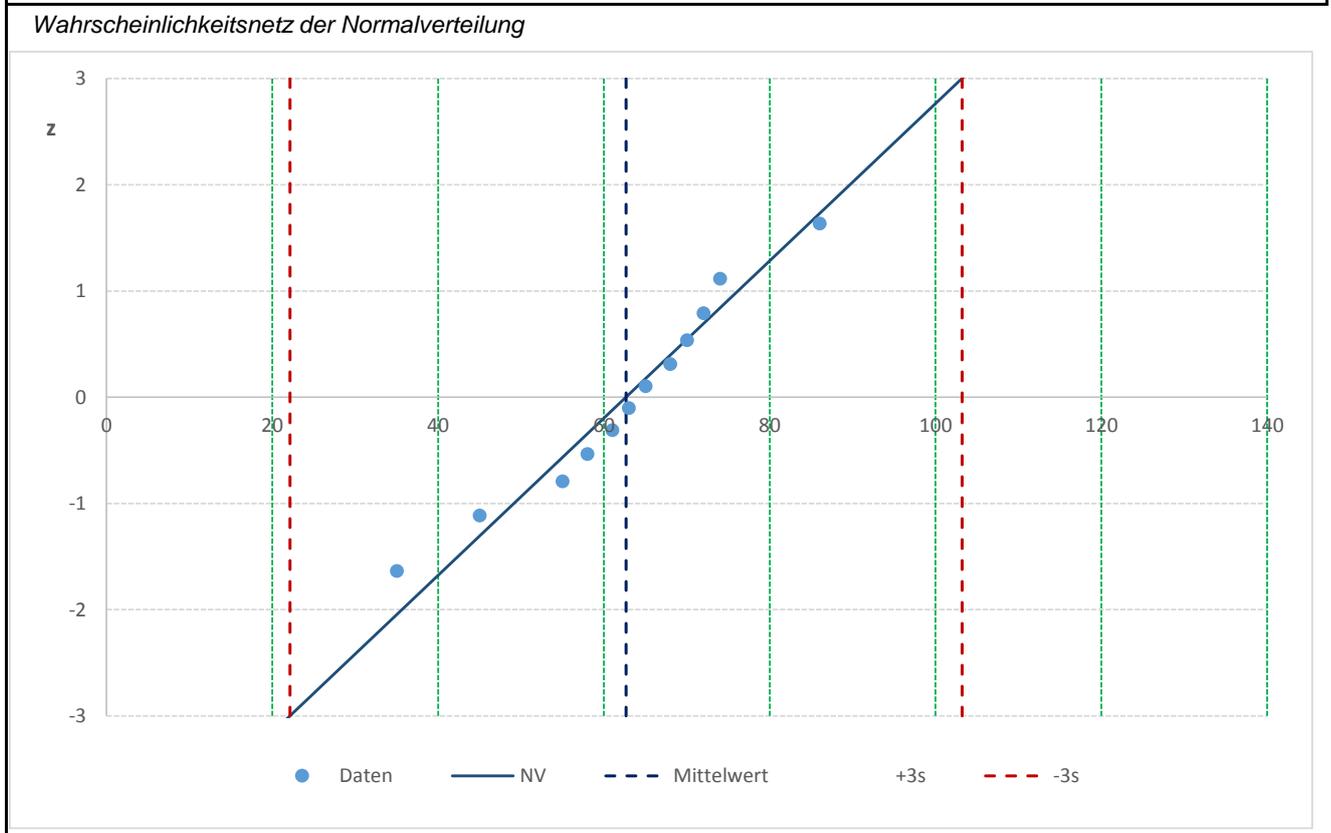
Daten

| | | | | | | | | | | |
|-------------------|----|----|----|----|----|----|----|----|----|----|
| $X_1 - X_{10}$ | 65 | 61 | 63 | 86 | 70 | 55 | 74 | 35 | 72 | 68 |
| $X_{11} - X_{20}$ | 45 | 58 | | | | | | | | |
| $X_{21} - X_{30}$ | | | | | | | | | | |
| $X_{31} - X_{40}$ | | | | | | | | | | |
| $X_{41} - X_{50}$ | | | | | | | | | | |

Statistik

| | | | | | | | | |
|--------------------|-------|----|--------|------|----|----------------|--------|----|
| Mittelwert | 62,7 | cm | Median | 64,0 | cm | Mittelwert +3s | 103,21 | cm |
| Standardabweichung | 13,51 | cm | Max | 86 | cm | Mittelwert -3s | 22,13 | cm |
| Schiefe | -0,5 | | Min | 35 | cm | | | |
| Kurtosis | 0,7 | | Range | 51 | cm | | | |

Testergebnis Nullhypothese H_0 : Daten stammen aus einer Normalverteilung.
 Alternativhypothese H_1 : Daten stammen nicht aus einer Normalverteilung.
 Testwahrscheinlichkeit p : 87,37%
 Testentscheidung H_1 : Nicht signifikant



 Datum _____ Unterschrift _____ Abteilung _____

| | | |
|----|------------|------|
| n | 12 | WAHR |
| SS | 2008,66667 | |
| m | 6 | |
| b | 44,1641 | |
| W | 0,97103 | |

| | | |
|----------------|-----|------|
| rel. Auflösung | RE% | 1,4% |
|----------------|-----|------|

| | | |
|--------------------|--------|-------------------|
| Auslesen Tabelle 2 | | W |
| p _{low} | 0,5 | 0,943 |
| p _{up} | 0,9 | 0,973 |
| diff | 0,4 | 0,03 |
| Interpolation | | 0,9342 |
| p | 0,8737 | Nicht signifikant |

| |
|---------|
| Spalten |
| 5 |
| 6 |

| | | | | | | | | | |
|-------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| Grafikdaten | | | | | | | | | |
| auslesen W | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| p | 0,01 | 0,02 | 0,05 | 0,1 | 0,5 | 0,9 | 0,95 | 0,98 | 0,99 |
| 12 | 0,805 | 0,828 | 0,859 | 0,883 | 0,943 | 0,973 | 0,979 | 0,984 | 0,986 |

Modell NV

| | | |
|----|------------|-------|
| OP | 116,719391 | 4,00 |
| UP | 8,61394269 | -4,00 |

| | | |
|------------|------------|---|
| Mittel | 62,6666667 | |
| stabw | 13,513181 | |
| Schiefe | -0,517998 | -0,732559811 <small>stand. Schiefe nach Sachs</small> |
| Kurt | 0,74659711 | 0,527923876 <small>stand. Kurtosis nach Sachs</small> |
| Mittel +3s | 103,20621 | |
| Mittel -3s | 22,1271237 | |
| Median | 64 | |
| Max | 86 | |
| Min | 35 | |
| Range | 51 | |

| | |
|------------|----|
| 62,6666667 | 3 |
| 62,6666667 | -3 |

| | |
|-------------|----|
| 103,2062097 | 3 |
| 103,2062097 | -3 |
| 22,12712368 | 3 |
| 22,12712368 | -3 |

| Daten |
|-------|
| 65 |
| 61 |
| 63 |
| 86 |
| 70 |
| 55 |
| 74 |
| 35 |
| 72 |
| 68 |
| 45 |
| 58 |

Diese Daten manuell einkopieren

| Ergebnis Testdaten | | | aktuell | |
|--------------------|-------------|-------------------|-------------|-------------------|
| plow | 0,5 | 0,943 | 0,5 | 0,943 |
| pup | 0,9 | 0,973 | 0,9 | 0,973 |
| diff | 0,4 | 0,03 | 0,4 | 0,03 |
| Interpolation | | 0,934202823 | | 0,934202823 |
| p | 0,873681129 | Nicht signifikant | 0,873681129 | Nicht signifikant |

Es sollten diese Ergebnisse erscheinen.