HEXAGON

Q-DAS qs-STAT

Testverfahren

Erläuterung und Handhabung

FAQ
16 August 2023
Created with Version 14.0.3.2

HxGN Quality Assurance

Information about this document

Parts of this document may be automatically translated.

Document History

Version	Date	Author(s)	Modifications / Remarks
v-0.14	16.08.2023	MBR/SJ	Translation

INHALT

1 Testverfahren

Wir verwenden statistische Hypothesentests, um Informationen über signifikante Eigenschaften einer Messreihe zu gewinnen. Ein typisches Anwendungsschema lässt sich wie folgt zusammenfassen.

1. Formulierung der Nullhypothese H_0
2. Formulierung der Alternativhypothese H_1
3. Planung des Tests (α -Risiko,β -Risiko und erforderliche Mindeststichprobengröße)
4. Erfassung der erforderlichen Daten
5. Durchführung der testspezifischen Berechnung der jeweiligen Prüfgröße
6. Vergleich der Prüfgröße des Tests mit den testspezifischen kritischen Werten für die Signifikanzniveaus α = 5 %, 1 % und - falls verfügbar - 0,1 %.
7. Ergebnis des Tests: Falls die Prüfgröße des Tests einen kritischen Wert unter- oder überschreitet, wird die Nullhypothese H_0 zugunsten der Alternativhypothese ausgeschieden H_1.

Unsere Software verfügt über eine große Zahl an Testverfahren. Für einen schnellen Überblick bietet unsere Software ein Übersichtsfenster, das die Ergebnisse mehrerer Testverfahren listenartig darstellt. Über die *Registerkarte <Ergebnisse>*| <*Testverfahren*> | <*Testverfahren Übersicht*> in der Gruppe "Analyse Merkmalsbezogen" öffnet man das Fenster "Testverfahren Übersicht". Die meisten anderen Befehle öffnen ein Fenster, in dem die Detail-Ergebnisse eines einzelnen Tests dargestellt sind. Die Auswahl eines bestimmten Testverfahrens ist über das Dropdown-Menü direkt unter der Schaltfläche <Testverfahren> möglich.

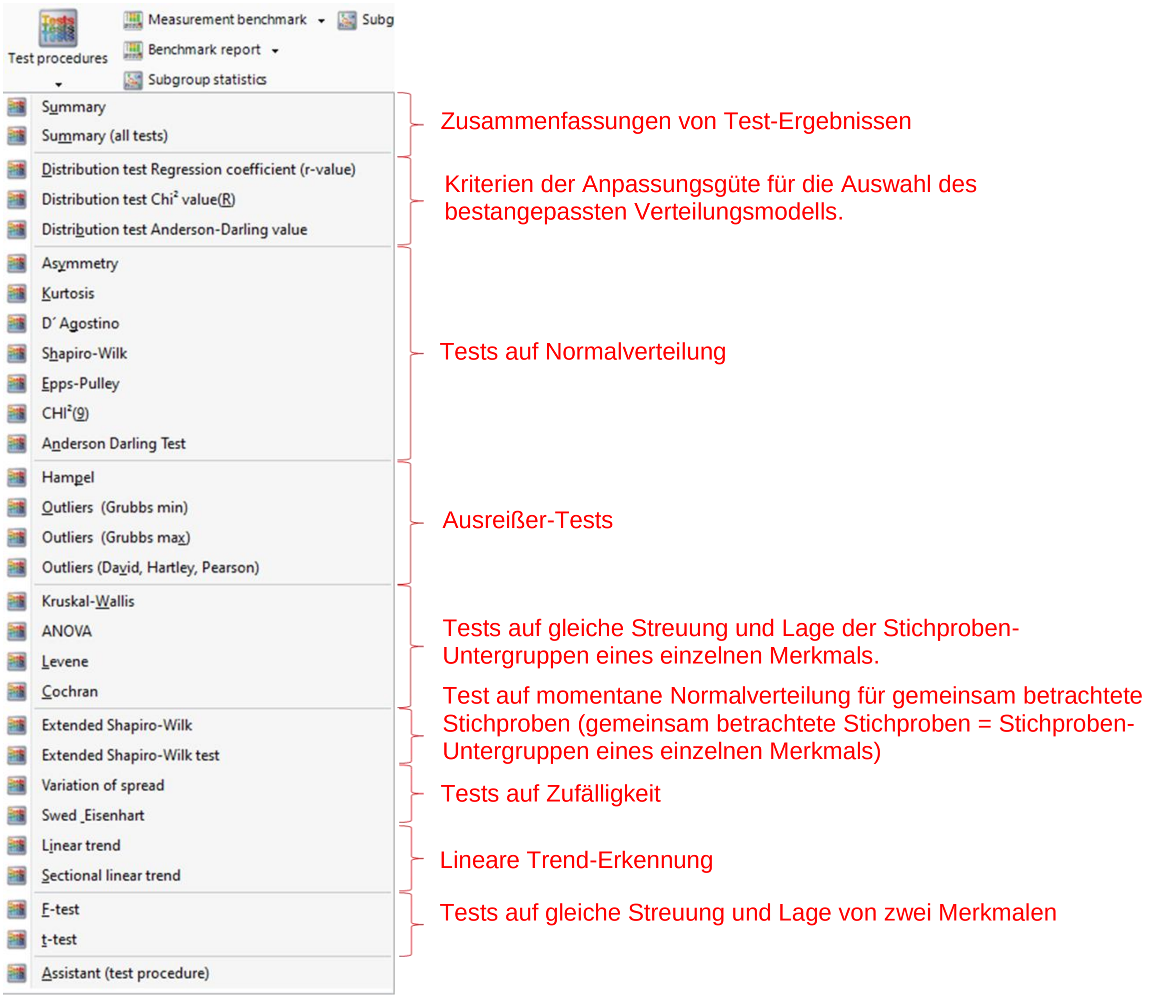


Abbildung 1: Multifunktionsleiste Unterbefehle der Registerkarte <Ergebnisse> | <Testverfahren>

Diese Vielfalt an Testverfahren umfasst verschiedene Typen von Tests.

- Zusammenfassung der Ergebnisse des Tests
  Eine listenartige Grafik, in der die Ergebnisse mehrerer Tests zusammenfassend darstellt sind.
- Normalitätstests
  Viele statistischen Auswerteverfahren setzen voraus, dass die Stichprobendaten aus einer normalverteilten Grundgesamtheit stammen. Mit einem Test auf Normalverteilung kann die Gültigkeit der Anwendungsvoraussetzung geprüft werden, bevor das entsprechende Verfahren ausgeführt wird.
- Ausreißer-Tests
  Manchmal sehen wir aufgrund unseres Fachwissen oder aufgrund langjähriger Erfahrung einen oder mehrere extreme Werte, von denen wir bezweifeln, dass sie aus derselben Grundgesamtheit stammen wie der Großteil der Daten. In einem solchen Fall kann ein Ausreißertest die Entscheidung erleichtern, den betreffenden Wert zu löschen. Wir müssen uns jedoch darüber im Klaren sein, dass die meisten Ausreißertestverfahren davon ausgehen, dass die Grundgesamtheit normalverteilt ist oder zumindest eine symmetrische Verteilung aufweist.
- Vergleich von Standort und Streuung
  Immer dann, wenn wir Messwerte aus verschiedenen Chargen und Losen erfassen oder die gleichen Artikel auf verschiedenen Maschinen oder zu unterschiedlichen Zeitpunkten produzieren, stellt sich die Frage, ob sich die Prozessstreuung (σ) oder der Prozessstandort (μ) aufgrund der unterschiedlichen Bedingungen verändert hat. Aus diesem Grund gibt es Testverfahren, die untersuchen, ob die Varianzen mehrerer Stichproben gleich sind. Auf ähnliche Weise kann man herausfinden, ob die Mittelwerte mehrerer Stichproben gleich sind.
- Tests auf Zufälligkeit
  Einige Testverfahren setzen voraus, dass die Daten eine Zufallsstichprobe aus einer Grundgesamtheit sind. Bestehen Zweifel daran, dass eine Stichprobe in einer zufälligen Reihenfolge vorliegt, sollten die Stichprobenwerte in der Grafik „Werteverlauf Einzelwerte“ betrachtet und zusätzlich ein Test auf Zufälligkeit durchgeführt werden.
- Lineare Trend-Erkennung
  Prozesse ändern sich oft systematisch, z. B. unterliegen sie dem Werkzeugverschleiß. Dies kann entweder zu einem langfristigen linearen Trend oder, wenn schnell Gegenmaßnahmen ergriffen werden, zu kurzfristigen linearen Trends führen.

Alle Testergebnisse enthalten eine Nullhypothese H0, eine Alternativhypothese H1, die berechnete Prüfgröße und das Testergebnis. Die Ergebnis-Fenster enthält im Kopfbereich gewöhnlich die Teile- und Merkmalnummer sowie den Namen des entsprechenden Testverfahrens. Speziell bei Einzeltest-Ergebnissen werden in der Grafik die kritischen Werte für das jeweilige Signifikanzniveau angezeigt.

1.1 Allgemein

1.1.1 Allgemeiner Ansatz

Ein Unternehmen kontrolliert die Qualität seiner Produkte, indem es z.B. alle 30 Minuten eine Stichprobe nimmt. Die wichtigen Merkmale (z. B. Länge, Durchmesser, Härte usw.) werden gemessen und eine statistische Auswertung damit erstellt. Berechnete Stichprobenkenngrößen sind jedoch zufälligen Einflüssen und damit Veränderungen unterworfen, die durch Unregelmäßigkeiten im Werkstoff, in den Maschinen und bei den Prüfern verursacht werden.

Wenn beispielsweise die Länge produzierter Teile $l = 20\ mm$ betragen soll, so könnte die Nullhypothese H0 für einen Einstichproben t-Test beispielsweise $\mu = 20$ mm lauten. Falls der Mittelwert $\bar{x}$ der Stichprobe nicht "zu weit" von diesem Vorgabewert $\mu = 20$ mm abweicht, wird die Nullhypothese nicht verworfen und die Produktion kann wie gewohnt fortgesetzt werden. Wenn die Abweichung jedoch "zu weit" ist, wird die Nullhypothese H0 verworfen und die Produktion wird gestoppt. Anschließend wird nach den Gründen für diese starke Abweichung gesucht.

Doch wo soll man die Grenze ziehen zwischen einer zufälligen Abweichung und einer zu großen Abweichung, die nicht allein durch den Zufall erklärt werden kann? Letzteres wird als **signifikante Abweichung bezeichnet**. Um eine objektive Aussage über diese Abweichungen zu erhalten, wurden die Verfahren der statistischen Hypothesentests entwickelt. Alle Verfahren haben einen gemeinsamen Ansatz.

Die Testverfahren beruhen auf zwei hypothetischen Annahmen -

- H_0 : Nullhypothese
- H_1 : Alternativhypothese.

Wenn ein Test auf Normalverteilung durchgeführt wird und die Nullhypothese H0 - die Stichprobe stammt aus einer normalverteilten Grundgesamtheit - NICHT abgelehnt wird, wird man in der Regel den Schluss ziehen, dass die Normalverteilung zumindest eine gute Annäherung an das Verteilungsmodell der Grundgesamtheit darstellt. Alle Tests auf Normalverteilung in unseren Softwareprodukten testen NICHT gegen ein spezifisches alternatives Verteilungsmodell (= spezifische Alternativhypothese H1). Wenn die Nullhypothese H0 durch den Test verworfen wird, gilt die folgende unspezifische Alternativhypothese H1: Die Stichprobendaten stammen nicht aus einer normalverteilten Grundgesamtheit.

Der Einstichproben t-Test ist ein Beispiel für ein Testverfahren, bei dem es möglich ist, sowohl für die Nullhypothese als auch für die Alternativhypothese eine spezifische Aussage zu formulieren. Im Falle des Einstichproben t-Tests werden unterschiedliche Aussagen über die Lage des Parameters µ für die Null- und die Alternativhypothese gewählt und es ist bei diesem Test möglich, die Risiken einer falschen Entscheidung zu bestimmen, noch bevor der Test durchgeführt wird:

		REALITÄT	
		H_0 ist richtig.	H_1 ist richtig
Ergebnis des Tests	H_0	Richtige Entscheidung! (Vertrauensniveau = $1 - \alpha$)	Falsche Entscheidung! Das Risiko für diese Fehlentscheidung zweiter Art (Typ II oder ß-Risiko) ist die Wahrscheinlichkeit β.
	H_1	Falsche Entscheidung! Das Risiko für diese Fehlentscheidung erster Art (Typ I oder α-Risiko) ist die Wahrscheinlichkeit α.	Richtige Entscheidung! (Power = $1 - \beta$)

Ein weiteres Risiko ist die Wahl eines falschen Verfahrens: Man wählt für seine Analyse ein Verfahren, bei dem die Annahmen und Voraussetzungen für dieses Verfahren nicht mit der Realität übereinstimmen. Ein Beispiel ist das Anwenden eines Ausreißertests bei asymmetrisch verteilten Daten, wenn der Test jedoch zwingend voraussetzt, dass dieser nur bei symmetrisch verteilten Daten verwendet werden darf.

Da das Testen von Hypothesen in der Regel verwendet wird, wenn die reale Situation oft weitgehend unbekannt ist, gibt es keine definitiv sichere Annahme. Jede Entscheidung birgt das Risiko einer Fehlentscheidung. Für eine erste Orientierung sollten immer zuerst die Grafiken „Werteverlauf Einzelwerte“ und „Histogramm“ der Einzelwerte genutzt werden.

Der Nicht-Verwerfen der Nullhypothese bedeutet nicht, dass die Aussage der Nullhypothese richtig ist. Die Wahl eines niedrigen Signifikanzniveaus verringert nur das Risiko eines Fehlers erster Art (Typ I oder α-Risiko), erhöht aber das Risiko eines Fehlers zweiter Art (Typ II oder ß-Risiko) unter denselben Testbedingungen. Ziel ist es, die Wahrscheinlichkeit beider Risiken so klein wie möglich zu wählen. Bei vielen statistischen Testverfahren kann man zwar leicht das Signifikanzniveau α wählen, aber es ist oft sehr schwierig bis unmöglich, die Wahrscheinlichkeit für den Fehler zweiter Art (Typ II oder ß-Risiko) zu bestimmen.

Wenn ein Testverfahren die Null-Hypothese nicht verworfen hat, sollten wir nicht sagen, dass die Null-Hypothese "richtig" ist, sondern dass sie "nicht verworfen" wurde. Insbesondere dann, wenn die verwendete Stichprobengröße für ein statistisches Testverfahren (zu) klein war: Unter einer solchen Bedingung ist die schlichte Wahrheit, dass die Stichprobengröße einfach nicht groß genug war, um überhaupt eine nennenswerte Abweichung von der H0-Bedingung feststellen zu können.

Woher leiten wir unsere Hypothesen ab?

- Von Vorgaben und Anforderungen, die erfüllt sein sollten/müssen.
- Der jeweilige Wert ist aus langjähriger Erfahrung bekannt.
- Von einer Theorie, die überprüft werden soll.
- Von Beobachtungen abgeleitet (Grafiken!).

Die folgende Vorgehensweise für das Fällen einer Testentscheidung ist im technischen Bereich üblich.

Es werden drei verschiedene Signifikanzniveaus definiert. Die Schwellenwerte des höchsten Signifikanzniveaus dienen als Kriterium für die Nichtverwerfung der Nullhypothese, während die Schwellenwerte des niedrigsten Niveaus die Grundlage für die Verwerfung der Nullhypothese bilden.

Eine typische Darstellung eines Ergebnisses in unseren Softwareprodukten sieht wie folgt aus.

Name des Tests			
Nullhypothese: < Aussage > Alternativhypothese: < Aussage >			
Die Niveaus der Bedeutung	Kritische Werte unter	obere	Statistischer Test
α = 5%	xxxxx	xxxxx	
α = 1%	xxxxx	xxxxx	xxxxx < *,**,*** >
α = 0.1%	xxxxx	xxxxx	

1.1.2 Interpretation des Testergebnisses (Prüfgröße)

Jede Tabelle mit Testergebnissen enthält eine Zelle der Prüfgröße des Tests. Die Hintergrundfarbe dieser Zelle hängt vom Ergebnis des Tests ab. Außerdem wird je nach Ergebnis ein Signifikanzcode (die kleinen Sterne) am Ende der Statistik hinzugefügt:

Tabelle 1: Die Hintergrundfarbe einer Prüfgröße und die Anzahl der Sterne, die dem Prüfgrößenwert in Abhängigkeit vom Testergebnis hinzugefügt werden.

Farbe	Anzahl der Sterne	Hintergrundfarbe der Prüfgröße	Bedeutung
grün	Keine	0.064098	Die Nullhypothese wird nicht verworfen (α > 5 %). Das Risiko einer Fehlentscheidung (H0 verwerfen, obwohl richtig) ist uns zu groß.
gelb	*	4.02001*	Die Nullhypothese wird auf dem Signifikanzniveau $\alpha \leq$ 5 % (genauer: 1 % < $\alpha \leq$ 5 %) verworfen. Das Fehlentscheidungsrisiko (H0 verwerfen, obwohl richtig) liegt also bei 1:20 oder weniger.
orange	**	7.96546**	Die Nullhypothese H0 wird auf dem Signifikanzniveau $\alpha \leq$ 1 % (genauer: 0,1 % <$\alpha \leq$ 1 %) abgelehnt. Das Fehlenscheidungsrisiko (H0 verwerfen, obwohl richtig) liegt also bei 1:100 oder weniger.
rot	***	4.36865***	Die Nullhypothese H0 wird auf dem Signifikanzniveau $\alpha \leq$ 0,1 % abgelehnt. Das Fehletnscheidungsrisiko (H0 verwerfen, obwohl richtig) liegt also bei 1:1000 oder weniger.

1.1.3 Bekannte Bedingungen, unter denen kein Testergebnis verfügbar ist

1.1.3.1 Keine kritischen Werte verfügbar (Stichprobengröße zu groß oder zu klein)

Bei einigen Tests ist die *theoretisch Verteilung der Prüfgröße* entweder unbekannt oder die praktische Anwendung ist numerisch zu aufwändig. In diesem Fall haben die Autoren der Testverfahren ihre Tabellen der kritischen Werten mit Hilfe von Monte-Carlo-Simulationen ermittelt. Und das bringt Einschränkungen mit sich: Es gibt nur einen bestimmten Wertebereich für die Stichprobengrößen, für den kritische Werte ermittelt wurden. Wenn der tatsächliche Stichprobenumfang der vorliegenden Daten außerhalb dieses testspezifischen Intervalls der verfügbaren Stichprobengrößen liegt (zu klein oder zu groß), gibt es keine kritischen Werte und der Test kann nicht durchgeführt werden.

1.1.3.2 Die kritischen Werte des Signifikanzniveaus α = 0,1 % fehlen

Eine weitere Einschränkung gilt für *einige* Testverfahren, bei denen die kritischen Werte aus Monte-Carlo-Simulationen gewonnen wurden: Die kritischen Werte für das Signifikanzniveau α = 0,1 % sind nicht verfügbar, weil die Autoren des Testverfahrens diese nicht ermittelt haben (vermutlich, weil die gewünschte numerische Genauigkeit nicht erreicht werden konnte). Dementsprechend zeigt das Programm qs-STAT keine kritischen Werte für das Signifikanzniveau $\alpha = 0{,}1$ % bei den davon betroffenen Testverfahren an.

1.1.3.3 Der Test wurde in der Auswertungsstrategie nicht selektiert

Ob ein Ergebnis verfügbar ist, hängt von den Options-Einstellungen in der Auswertungsstrategie ab: Jeder Test kann auf "an" oder "aus" gesetzt werden. Im Fenster "Testverfahren Übersicht" werden die Ergebnisse derjenigen Tests aufgelistet, die in der Auswertungsstrategie auf "an" gesetzt wurden. Wenn der gewünschte Test dort nicht aufgelistet ist, so werden auch im entsprechenden Einzeltestfenster keine Ergebnisse angezeigt.

1.2 Übersichten zu den Testverfahren

1.2.1 Testverfahren Übersicht

Multifunktionsleiste:
Registerkarte <Ergebnisse> | <Testverfahren> | <Testverfahren Übersicht>

In diesem Fenster wird eine Zusammenfassung der verschiedenen Tests auf einen Blick angezeigt. Das Fenster enthält eine Tabelle, in der die Ergebnisse derjenigen Testverfahren aufgelistet sind, die *in der gewählten Auswertungsstrategie aktiviert* wurden. Mit anderen Worten: die Länge der aufgelisteten Ergebnisse im Fenster "Testverfahren Übersicht" hängt von der Anzahl der aktivierten Testverfahren in der verwendeten Auswertungsstrategie ab.

Summary

Part no.	1	Part descr.	Assembly #1
Char.No.	1	Char.Descr.	Test 1

Test		Test hypothesis	Test statistic
Swed & Eisenhart	H_0	Random data sequence	4.36865***
	H_1	Non random data sequence	
D` Agostino	H_0	The log transformed sample was taken from a normal distribution. (→)	-411.034**
	H_1	The log transformed sample was NOT taken from a normal distribution. (→)	
CHI² test	H_0	The log transformed sample was taken from the assumed distribution. (→)	28.3156***
	H_1	The log transformed sample was NOT taken from the assumed distribution. (→)	
Outliers David, Hartley, Pearson	H_0	Neither x_{min} nor x_{max} is an outlier	7.96546**
	H_1	Either x_{min} or x_{max} is an outlier	
Outlier Grubbs maximum value	H_0	Maximum value x_{max} is not an outlier	4.02001*
	H_1	Maximum value x_{max} is an outlier	
Outlier Grubbs minimum value	H_0	Minimum value x_{min} is not an outlier	3.94545*
	H_1	Minimum value x_{min} is an outlier	
Asymmetry	H_0	The log transformed values were taken from a symmetric distribution. (→)	0.037891
	H_1	The log transformed values were NOT taken from a symmetric distribution. (→)	
Kurtosis	H_0	The log transformed values were taken from a distribution with kurtosis = 3. (→)	5.00912**
	H_1	The log transformed values were NOT taken from a distribution with kurtosis = 3. (→)	
Successive spread variation	H_0	Successive spread variations are uncorrelated	1.40934***
	H_1	Successive spread variations are correlated	
Extended Shapiro-Wilk	H_0	Subgroup is derived from a Normal distribution	-2.26934*
	H_1	Subgroup is NOT derived from a Normal distribution	
Hampel test	H_0	The Hampel test has established no outliers	4.81786**
	H_1	The Hampel test has established outliers	
Outlier Grubbs Minimum and maximum value	H_0	The minimum value x_{min} and the maximum value x_{max} is no outlier	4.02001*
	H_1	The minimum value x_{min} or the maximum value x_{max} is an outlier	
Test for sectional linear trend	H_0	Sectional linear trend exists	0.00000*
	H_1	No sectional linear trend exists	
Anderson Darling Test	H_0	The log transformed sample was taken from a normal distribution. (→)	3.36128***
	H_1	The log transformed sample was NOT taken from a normal distribution. (→)	

Delete outliers temporarily? Delete outliers?

Abbildung 2: Das Fenster Zusammenfassung *enthält die Ergebnisse der in der verwendeten Auswertungsstrategie aktivierten Testverfahren.*

Spaltenüberschrift "Test"
Hier steht der Name des Prüfverfahrens.

Spaltenüberschrift "Testhypothese"
Diese Spalte enthält die Aussagen der Null-Hypothese H0 und der Alternativ-Hypothese H1 für jeden aufgeführten Test.

Spaltenüberschrift "Prüfgröße"
Diese Spalte enthält die Werte der Prüfgrößen. Die Hintergrundfarbe jeder Zelle und auch die Anzahl der zum Wert der Prüfgröße hinzugefügten Sterne hängt vom Ergebnis des Tests ab (Farbcode und Anzahl der Sterne: siehe Tabelle 1 auf Seite 8).

Schaltflächen "Ausreißer temporär entfernen?" und "Ausreißer entfernen?"
Wenn ein Ausreißertest aktiv ist und dieser Ausreißertest Ausreißer für das ausgewählte Merkmal festgestellt hat, werden unten im Fenster „Testverfahren Übersicht“ zwei Schaltflächen: "*Ausreißer temporär entfernen?"* und "*Ausreißer entfernen?"*. Durch Klicken auf "*Ausreißer entfernen?"* werden Ausreißer mit Hilfe des Attributs 191 gelöscht. Wenn die gleichen Daten gespeichert und wieder geöffnet werden, sind die Ausreißer bereits entfernt. Mit der Schaltfläche "*Ausreißer temporär entfernen?"* wird ein Ausreißer nur für die aktuelle Sitzung gelöscht. Die Messwerte werden nicht mit dem Attribut 191 gespeichert und gehen beim nächsten Laden oder Öffnen des betreffenden Datensatzes wieder in die Berechnung ein.

1.2.1.1 Testverfahren Übersicht (alle Tests)

Multifunktionsleiste:
Registerkarte <Ergebnisse> | <Testverfahren> | <Testverfahren Übersicht (alle Tests)>

Die Verfügbarkeit der Ergebnisse im Fenster "*Testverfahren Übersicht (alle Tests)*" ist an eine Vorbedingung gebunden: Die Option für diese Funktionalität muss in der verwendeten Auswertungsstrategie aktiv sein:

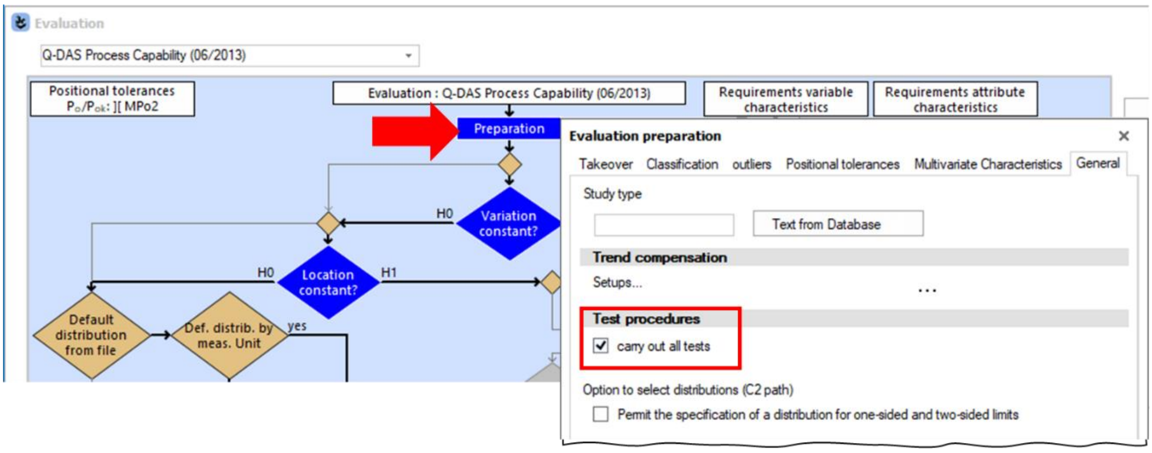


Abbildung 3: Auszug aus der Auswertungsstrategie (Modul Prozessfähigkeit) mit dem Eigenschaftsfenster des Symbols "Vorbereitung". In diesem Fenster muss die Option "Testverfahren: alle Tests durchführen" aktiv sein, um Ergebnisse im Fenster "Testverfahren Übersicht (alle Tests)" zu erhalten.

Die im Fenster "*Testverfahren Übersicht (alle Tests)*" aufgeführten Testergebnisse unterscheiden sich in einigen Punkten von denjenigen Inhalten, die im Fenster "*Testverfahren Übersicht*" aufgelistet sind:

1. *Alle* Testverfahren werden unabhängig von anderen testspezifischen Optionseinstellungen in der Auswertungsstrategie durchgeführt.
2. Es gibt KEINE Fenster für die Einzeltest-Ergebnisse, in dem die Details zu einem einzelnen Testverfahren angezeigt werden können.
3. Tests auf Normalverteilung werden mit den Originaldaten durchgeführt (und nicht mit den transformierten Daten).
4. Die Beschränkungen bezüglich der oberen Grenze des Stichprobenumfangs für die kritischen Werte werden bei einigen der Testverfahren ignoriert.

Zu beachten ist auch, dass die Option "alle Tests durchführen" zusätzliche Rechenressourcen verbraucht.

Summary (all tests)

Test		Test hypothesis	Test statistic
Swed & Eisenhart	H_0	Random data sequence	4.36865***
	H_1	Non random data sequence	
D` Agostino	H_0	Subgroup is derived from a Normal distribution	-9.68420**
	H_1	Subgroup is NOT derived from a Normal distribution	
CHI² test	H_0	Subgroup is derived from assumed distribution	34.2343***
	H_1	Subgroup is not derived from assumed distribution	
Outliers David, Hartley, Pearson	H_0	Neither x_{min} nor x_{max} is an outlier	7.96546**
	H_1	Either x_{min} or x_{max} is an outlier	
Outlier Grubbs maximum value	H_0	Maximum value x_{max} is not an outlier	4.02001*
	H_1	Maximum value x_{max} is an outlier	
Outlier Grubbs minimum value	H_0	Minimum value x_{min} is not an outlier	3.94545*
	H_1	Minimum value x_{min} is an outlier	
Asymmetry	H_0	Values are derived from a symmetric distribution (ND)	0.064098
	H_1	Values are not derived from a symmetric distribution (ND)	
Kurtosis	H_0	Values are derived from a distribution with Kurtosis = 3 (ND)	5.00195**
	H_1	Values are not derived from a distribution with Kurtosis = 3 (ND)	
Successive spread variation	H_0	Successive spread variations are uncorrelated	1.40934***
	H_1	Successive spread variations are correlated	
Epps-Pulley	H_0	Subgroup is derived from a Normal distribution	2.31249**
	H_1	Subgroup is NOT derived from a Normal distribution	
Hampel test	H_0	The Hampel test has established no outliers	4.81786**
	H_1	The Hampel test has established outliers	
Outlier Grubbs Minimum and maximum value	H_0	The minimum value x_{min} and the maximum value x_{max} is no outlier	4.02001*
	H_1	The minimum value x_{min} or the maximum value x_{max} is an outlier	
Test for sectional linear trend	H_0	Sectional linear trend exists	0.00000*
	H_1	No sectional linear trend exists	
Anderson Darling Test	H_0	Subgroup is derived from a Normal distribution	3.38954***
	H_1	Subgroup is NOT derived from a Normal distribution	

Abbildung 4: Das Fenster Testverfahren Übersicht (alle Tests) *enthält eine Zusammenfassung aller Tests, unabhängig von den Einstellungen der testspezifischen Auswertungsstrategie.*

1.2.1.2 Verteilungsauswahl mit dem Verteilungstest Regressionskoeffizienten (r-Wert)

Multifunktionsleiste:
Registerkarte <Ergebnisse> | <Testverfahren> | <Verteilungstest Regressionskoeffizient (r...)>

Der Titel dieses Fensters ist etwas irreführend: In diesem Fenster werden keine Ergebnisse eines statistischen Hypothesentests angezeigt.

Dieses Fenster wird geöffnet, wenn der *Regressionskoeffizient* in der Auswertungsstrategie als Kriterium für die Ermittlung des bestgeeigneten Verteilungsmodells ausgewählt wurde.

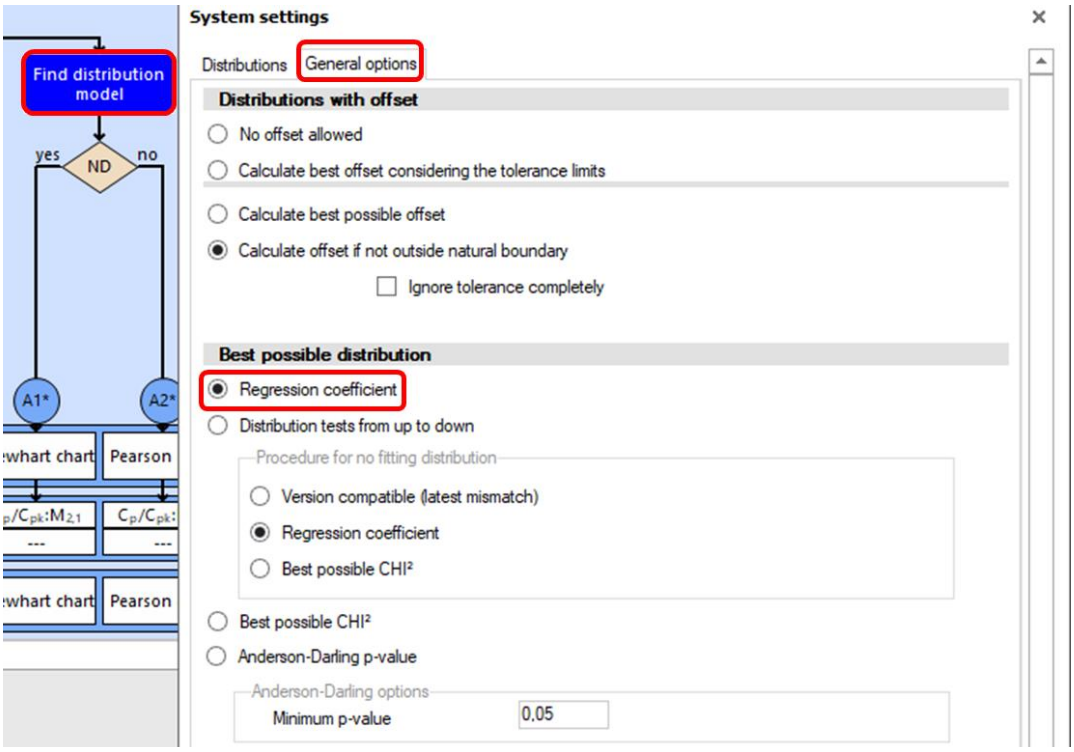


Abbildung 5: Auszug aus der Auswertungsstrategie - In diesem Beispiel wird die Option "Regressionskoeffizient" als Kriterium für die Suche nach der bestgeeigneten Verteilung ausgewählt.

Wird das - ursprünglich in der Auswertungsstrategie festgelegte - Standard-Verteilungsmodell durch einen Anpassungsgütetest abgelehnt, wählt das automatische Verfahren zur Selektion der Verteilung das Verteilungsmodell mit der höchsten *Summe der Regressionskoeffizienten* (➔ r-Wert). Der r-Wert ist eine Summe[1] von *zwei* Regressionskoeffizienten $r_{100\,\%} + r_{25\,\%}$. Daher ist der r-Wert in der Regel größer als eins. Das Fenster zeigt uns die r-Werte für die Menge der Verteilungsmodelle an, die in der Auswertungsstrategie als mögliche Kandidaten für die Suche nach der bestgeeigneten Verteilung vorausgewählt wurden.

Ein Regressionskoeffizient drückt den Grad der Anpassungsgüte zwischen der empirischen Verteilungsfunktion (Stichprobendaten) und der Verteilungsfunktion des angepassten Verteilungsmodells auf einer numerischen Skala von 0 bis 1 aus. Je höher sein Wert ist, desto besser ist die Anpassung zwischen dem Verteilungsmodell und den Stichprobendaten. Der Regressionskoeffizient $r_{100\,\%}$ wird für alle Stichprobendaten berechnet. Ein Indikator für eine gute Anpassung ist ein Regressionskoeffizient $r_{100\,\%}$ der sehr nahe bei 1,0 liegt.

[1] Wenn Spezifikationsgrenzen existieren. Wenn KEINE Spezifikationsgrenzen vorhanden sind, wird nur $r_{100\,\%}$ verwendet.

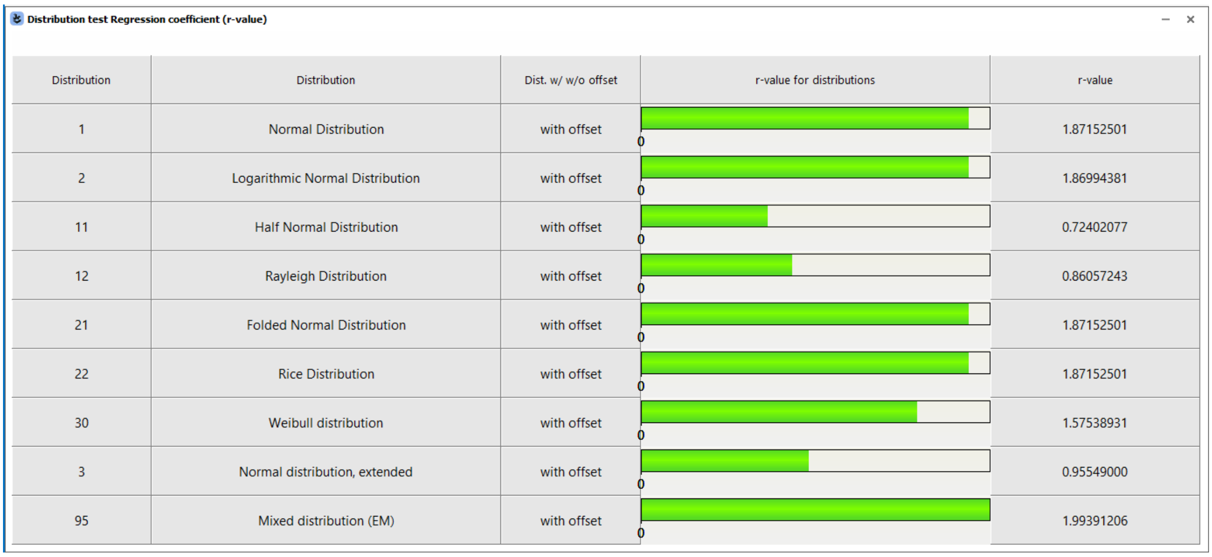
Ein zweiter Regressionskoeffizient, $r_{25\,\%}$wird berechnet, um ein zusätzliches Urteil über die Anpassungsgüte in dem "kritischen Randbereich der Verteilung" zu erhalten. Was bedeutet ein "kritischer Randbereich der Verteilung"? Werfen wir einen Blick auf die Definition des kritischen Fähigkeitsindexes C_{pk}der in den meisten Auswertungsstrategien verwendet wird:

$$C_{pk} = min\{C_{pk.upper}; C_{pk.lower}\} = min\left\{\frac{USL - X_{50\,\%}}{X_{99.865\,\%} - X_{50\,\%}}; \frac{X_{50\,\%} - LSL}{X_{50\,\%} - X_{0.135\,\%}}\right\}$$

Das heißt, entweder die untere Spezifikationsgrenze oder die obere Spezifikationsgrenze ist entscheidend für C_{pk}:

- Wenn die untere Grenze die kritische Seite für C_{pk}ist, verwendet die Software die ersten 25 % der Stichprobendaten zur Berechnung des Regressionskoeffizienten $r_{25\,\%}$ als Maß für die Anpassungsgüte im unteren (linken) Randbereich der Verteilung.
- Wenn die obere Grenze die kritische Seite für C_{pk}ist, verwendet die Software die letzten 25 % der Stichprobendaten zur Berechnung des Regressionskoeffizienten $r_{25\,\%}$ als Maß für die Anpassungsgüte im oberen (rechten) Randbereich der Verteilung.
- Die Summe der beiden Koeffizienten, $r_{100\,\%} + r_{25\,\%}$wird in der Spalte "r-Wert" aufgeführt-. Ein "r-Wert" in der Nähe von 2,0 zeigt eine gute Anpassung an.

Wenn das anfänglich ausgewählte Standardmodell gut passt (= H0 nicht abgelehnt), wird das Verfahren zur Selektion des Verteilungsmodells überhaupt nicht angewendet: In diesem Fall zeigt das Fenster keine Ergebnisse an.


Distribution test Regression coefficient (r-value)

Distribution	Distribution	Dist. w/ w/o offset	r-value for distributions	r-value
1	Normal Distribution	with offset	0	1.87152501
2	Logarithmic Normal Distribution	with offset	0	1.86994381
11	Half Normal Distribution	with offset	0	0.72402077
12	Rayleigh Distribution	with offset	0	0.86057243
21	Folded Normal Distribution	with offset	0	1.87152501
22	Rice Distribution	with offset	0	1.87152501
30	Weibull distribution	with offset	0	1.57538931
3	Normal distribution, extended	with offset	0	0.95549000
95	Mixed distribution (EM)	with offset	0	1.99391206

Abbildung 6: Das Fenster "Verteilungstest Regressionskoeffizient (r-Wert)" zeigt das Ergebnis der Summe der beiden Koeffizienten $r_{100\,\%}$ und $r_{25\,\%}$, die für die Suche nach dem bestgeeigneten Verteilungsmodell verwendet wurden.

1.2.1.3 Verteilungsauswahl mit dem Chi^2-Wert (G)

Multifunktionsleiste:
Registerkarte <Ergebnisse> | <Testverfahren> | <Verteilungstest Chi^2-Wert(G)>

Dieses Fenster wird geöffnet, wenn der p-Wert des χ^2- Anpassungstest das ausgewählte Kriterium für die Suche nach dem bestangepassten Verteilungsmodell ist.

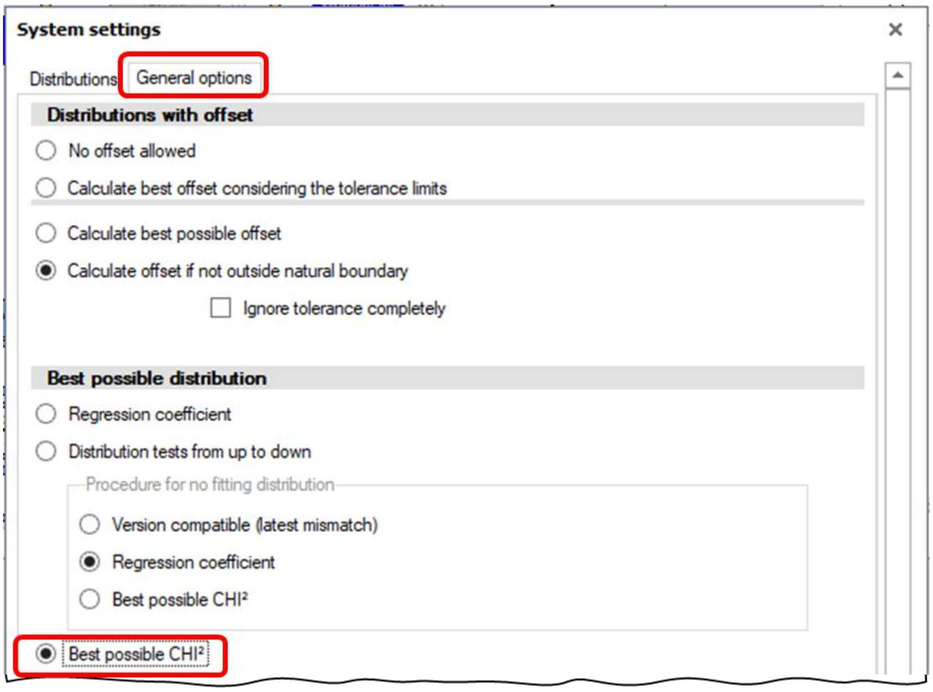
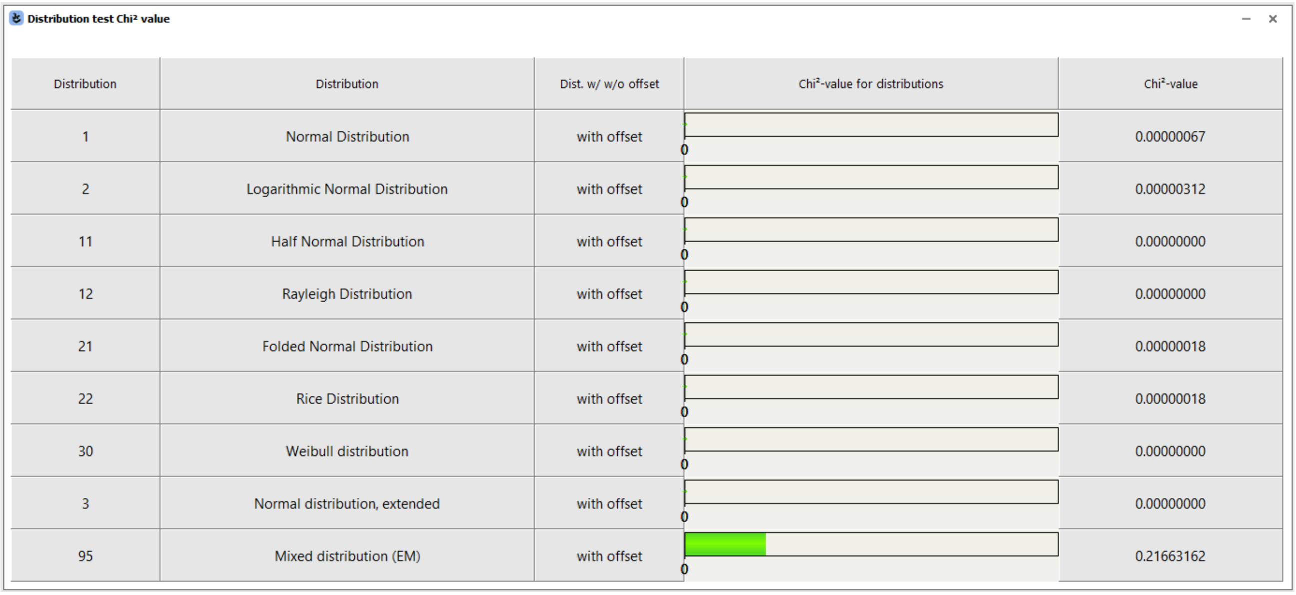


Abbildung 7: Auszug aus einer Auswertungsstrategie - In diesem Beispiel ist die Option "Bestes CHI^2" das gewählte Kriterium für die Suche nach dem am besten passenden Verteilungsmodell.

Wenn das Standard-Verteilungsmodell - das ursprünglich in der Auswertungsstrategie festgelegt wurde - durch einen Test auf Güte der Verteilungsanpassung abgelehnt wird, führt die automatische Selektion der Verteilungen für jedes Verteilungsmodell den χ^2Anpassungstest durch und wählt als bestangepasstes Modell das Verteilungsmodell mit dem *höchsten p-Wert* aus. Der Name der Spaltenüberschrift "Chi^2-Wert" ist etwas irreführend, der in dieser Spalte angegebene Wert ist der *p-Wert* des Chi^2-Anpassungstests und nicht die Prüfgröße.


Distribution test Chi² value

Distribution	Distribution	Dist. w/ w/o offset	Chi²-value for distributions	Chi²-value
1	Normal Distribution	with offset	0	0.00000067
2	Logarithmic Normal Distribution	with offset	0	0.00000312
11	Half Normal Distribution	with offset	0	0.00000000
12	Rayleigh Distribution	with offset	0	0.00000000
21	Folded Normal Distribution	with offset	0	0.00000018
22	Rice Distribution	with offset	0	0.00000018
30	Weibull distribution	with offset	0	0.00000000
3	Normal distribution, extended	with offset	0	0.00000000
95	Mixed distribution (EM)	with offset	0	0.21663162

Abbildung 8: Das Fenster "Verteilungstest Chi²-Wert" listet in der Spalte "Chi².Wert" die p-Werte der χ^2 –Anpassungstests für alle in der automatischen Verteilungssuche vorselektierten Modellverteilungen. Das Modell, für das der höchste p-Wert berechnet wurde, wird als das am besten passende Verteilungsmodell ausgewählt.

Wenn das anfänglich in der Auswertestrategie vorausgewählte Default-Verteilungsmodell gut passt (= H0 wurde nicht abgelehnt), wird das Verfahren zur Auswahl des bestangepassten Verteilungsmodells nicht angewendet: In diesem Fall zeigt das Fenster keine Ergebnisse an.

1.2.1.4 Verteilungsauswahl mit dem Anderson-Darling Wert

Multifunktionsleiste:
Registerkarte <Ergebnisse> | <Testverfahren> | <Verteilungstest Anderson-Darling-Wert>

Dieses Fenster wird geöffnet, wenn der p-Wert des Anderson Darling-Tests für die Anpassungsgüte das ausgewählte Kriterium für die Suche nach dem bestangepassten Verteilungsmodell ist:

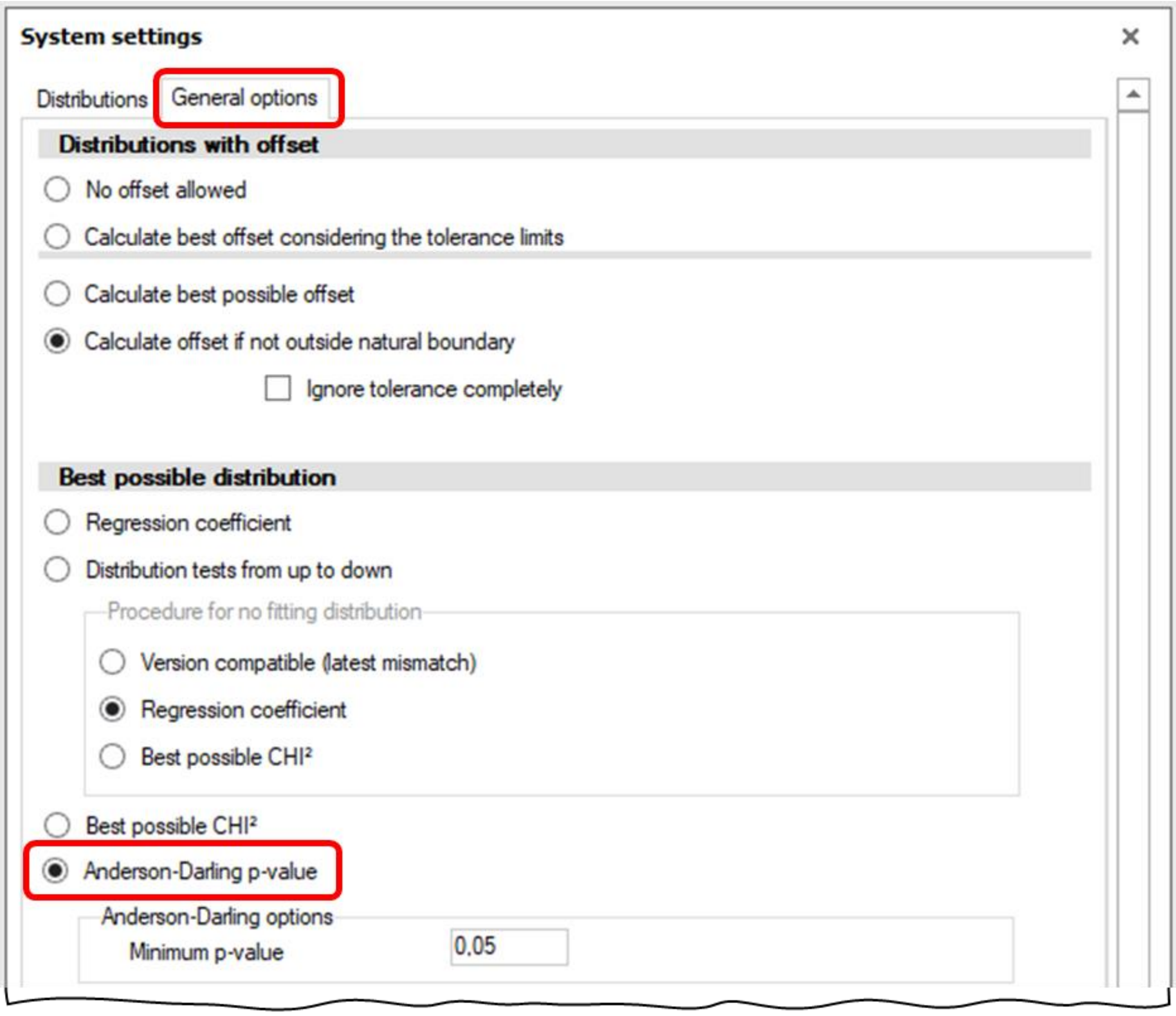
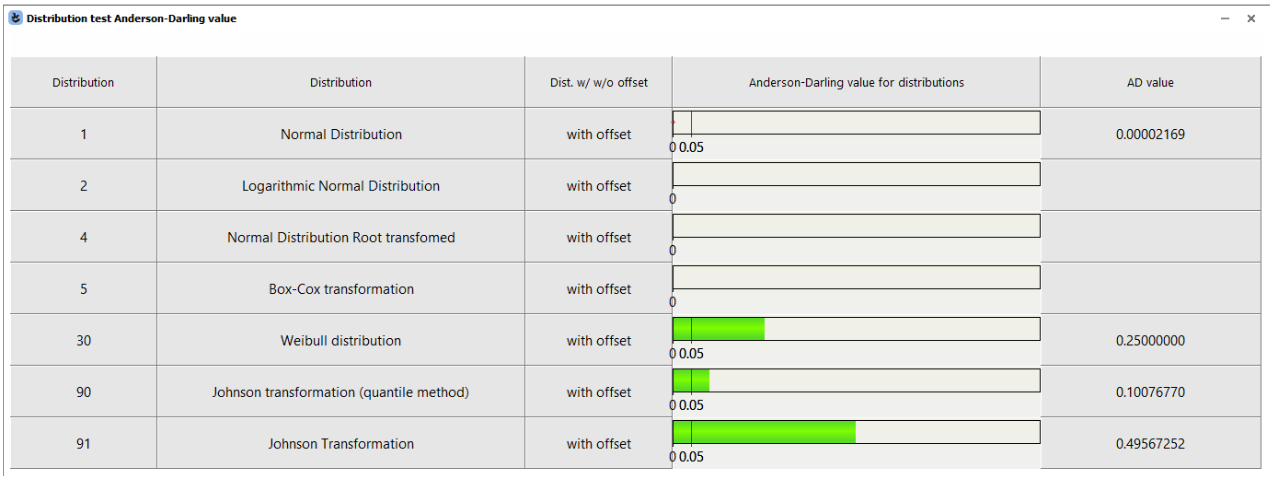


Abbildung 9: Auszug aus der Auswertungsstrategie - In diesem Beispiel ist die Option "Anderson-Darling p-Wert" das gewählte Kriterium für die Suche nach dem bestgeeigneten Verteilungsmodell

Wenn das in der Auswertestrategie vorausgewählte Default-Verteilungsmodell durch einen Test auf Güte der Verteilungsanpassung abgelehnt wird, so führt das Programm die automatische Suche nach dem bestangepassten Verteilungsmodell aus. Dabei werden für alle vorausgewählten Verteilungsmodell-Kandidaten (eingestellt in der Auswertestrategie) die P-Werte des Anderson-Darling-Test berechnet. *Das Verteilungsmodell, für dass der höchste P-Wert ermittelt wurde, wird final ausgewählt*. Die Spaltenüberschrift "AD-Wert" ist etwas irreführend, denn der in dieser Spalte angegebene Wert ist der P-*Wert* des Anderson-Darling-Tests und nicht die Prüfgröße.



Distribution test Anderson-Darling value

Distribution	Distribution	Dist. w/ w/o offset	Anderson-Darling value for distributions	AD value
1	Normal Distribution	with offset	0 0.05	0.00002169
2	Logarithmic Normal Distribution	with offset	0	
4	Normal Distribution Root transfomed	with offset	0	
5	Box-Cox transformation	with offset	0	
30	Weibull distribution	with offset	0 0.05	0.25000000
90	Johnson transformation (quantile method)	with offset	0 0.05	0.10076770
91	Johnson Transformation	with offset	0 0.05	0.49567252

Abbildung 10: Im Fenster "Verteilungstest Anderson-Darling-Wert" sind in der Spalte "AD-Wert" die p-Werte des Tests der gespeicherten Verteilung nach CHI²-Anpassung aufgeführt. Das Modell, für das der höchste p-Wert berechnet wurde, wird als bestangepasstes Verteilungsmodell ausgewählt.

Hinweis:

Der p-Wert des Anderson-Darling-Tests ist nicht für alle Verteilungsmodelle verfügbar. Die folgende **Teilmenge** an Verteilungsmodellen kann für die Suche nach dem bestangepassten Modell verwendet werden, wenn die Option "Anderson-Darling p-Wert" verwendet wird:

- Normalverteilung
- Logarithmische Normalverteilung
- Normalverteilung Wurzel transformiert
- Box-Cox Transformation
- Weibullverteilung
- Johnson-Transformation

1.3 Fenster für Test-Ergebnisse einzelner Tests

Nicht *jeder* Test ist in einer Auswertungsstrategie aktiviert. Um einen schnellen Überblick zu erhalten, welche Testverfahren in der verwendeten Auswertungsstrategie aktiviert sind oder nicht, ist zunächst das Fenster "Übersicht Testverfahren" zu öffnen (Multifunktionsleiste: Registerkarte <Ergebnis> **|** <Testverfahren> **|** <Testverfahren Übersicht>).

Ist ein bestimmtes Testverfahren im Fenster "Übersicht Testverfahren" aufgelistet, enthält das Fenster der Einzeltest-Ergebnisse die Details zu den Testergebnissen. Ist ein bestimmter Test nicht in dem Fenster "Testverfahren Übersicht" aufgeführt, so werden im Fenster "Einzelergebnisse" des jeweiligen Tests keine Ergebnisse angezeigt. Stattdessen wird die Bemerkung "Testverfahren nicht in der Auswertungsstrategie selektiert" eingeblendet.

1.3.1 Tests auf Anpassungsgüte für Verteilungsmodelle

1.3.1.1 Asymmetrie Test

Multifunktionsleiste:
Registerkarte <Ergebnisse> **|** <Testverfahren> **|** <Asymmetrie>

Dieser Test auf Normalverteilung wird gemäß der internationalen Norm ISO 5479: 1997 durchgeführt. Die Prüfgröße ist der *absolute Wert der Schiefe* der Stichprobe.

$$\left|\sqrt{b_1}\right| = \left|\frac{m_3}{\sqrt{m_2^3}}\right|$$

wobei m_2 und m_3 die *zentralen Momente* zweiter und dritter Ordnung der Stichprobendaten sind:

$$m_r = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^r \,; r = 2, 3$$

Eine Schiefe von Null weist auf eine symmetrische Verteilung hin. Eine Normalverteilung hat eine Schiefe von Null, da es sich um eine symmetrische Verteilung handelt. Ist der *absolute* Wert der Schiefe der Stichprobe deutlich größer als Null, deutet dies darauf hin, dass die Stichprobendaten aus einer Grundgesamtheit mit einer asymmetrischen Verteilung stammen.

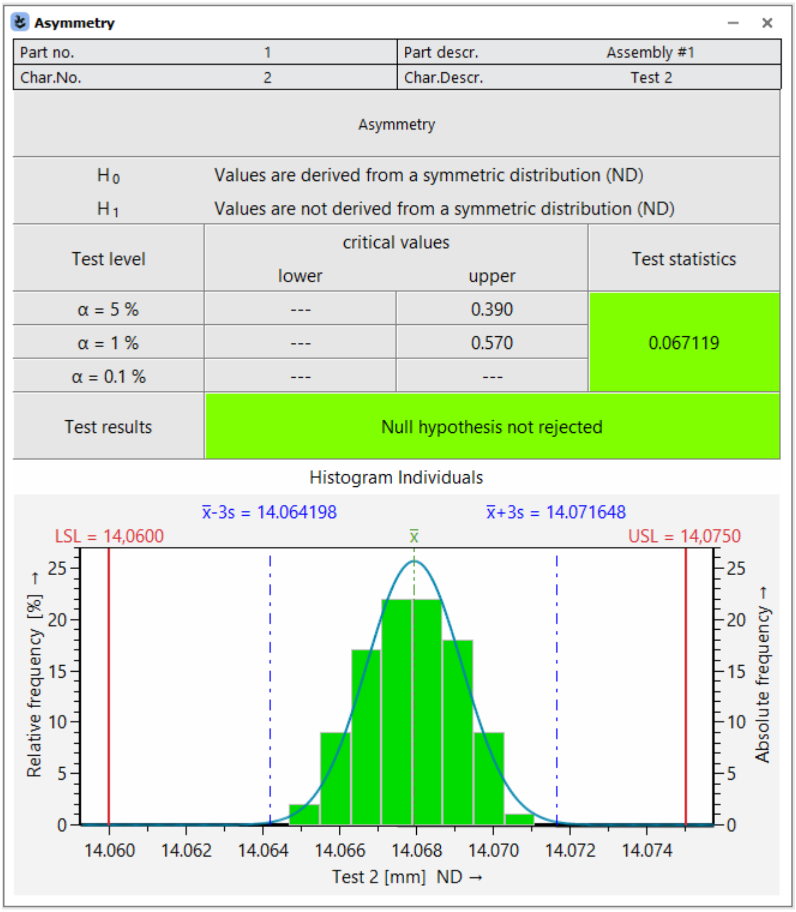


Abbildung 11: Fenster "Asymmetrie" - Die Prüfgröße ist der absolute Wert der Schiefe der Stichprobe $|\sqrt{b_1}|$.

Kritische Werte für das Testverfahren sind für Stichproben verfügbar, deren Gesamtstichprobenumfänge im Wertebereich von $n = 8$ bis $n = 5000$ liegen.

Referenz:

ISO 5479: 1997, Kapitel 6.2 und Tabelle 8 - Test auf Schiefe auf Seite 24.

Der Asymmetrie-Test und der Kurtosis-Test sollten stets gemeinsam ausgeführt werden.

1.3.1.2 Kurtosis

Multifunktionsleiste:
Registerkarte <Ergebnis> | <Testverfahren> | <Kurtosis>

Dieser Normalitätstest wird gemäß der internationalen Norm ISO 5479: 1997 durchgeführt. Die Kurtosis einer Stichprobe oder Verteilung gibt uns Aufschluss darüber, ob diese Verteilung eine hohe und steile Spitze oder eine flache und breite Spitze aufweist. Die verwendete Prüfgröße des Tests ist die Kurtosis der Stichprobendaten.

$$b_2 = \frac{m_4}{m_2^2}$$

wobei m_2 und m_4 die *zentralen Momente* zweiter und vierter Ordnung der Stichprobendaten sind:

$$m_r = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^r \,; r = 2, 4$$

Das Modell Normalverteilung hat eine Kurtosis von drei. Weicht die Prüfgröße des Tests (Wölbung der Stichprobe b_2) erheblich von dem Wert drei ab, so deutet das darauf hin, dass die betreffende Stichprobe nicht aus einer normalverteilten Grundgesamtheit stammt.

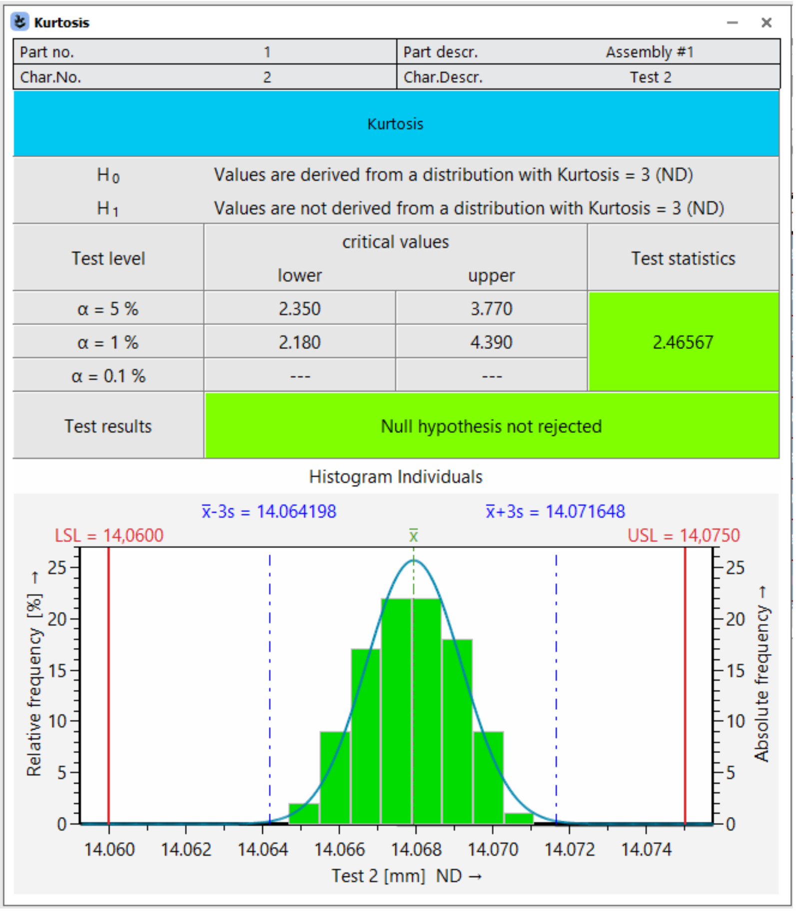


Abbildung 12: Fenster "Kurtosis" - Die Prüfgröße ist die Stichproben-Kurtosis b_2.

In der Software sind die kritischen Werte für das Testverfahren vertafelt. Kritische Werte stehen für Stichproben zur Verfügung, deren Gesamtumfang im Wertebereich von $n = 8$ bis $n = 5000$ liegt.

Referenz:

ISO 5479: 1997, Kapitel 6.3 und Tabelle 9 Test auf Kurtosis auf Seite 25.

Der Kurtosis-Test und der Asymmetrie-Test sollten gemeinsam ausgeführt werden.

1.3.1.3 D'Agostino's D-Test

Multifunktionsleiste:
Registerkarte <Ergebnisse> | <Testverfahren> | <D'Agostino>

Dieser Test auf Normalverteilung war ursprünglich als einfacher zu berechnendes alternatives Testverfahren für den Shapiro-Wilk W-Test gedacht. Die Prüfgröße *D* wird auf der Grundlage sortierter Stichprobendaten berechnet x (aufsteigende Reihenfolge):

$$D = \frac{\sum_{i=1}^{n}\left(i - \frac{(n+1)}{2}\right)x_i}{n^2\sqrt{m_2}}$$

wobei n der Gesamtumfang der Stichprobe ist und m_2 das zentrale Moment zweiter Ordnung der Stichprobendaten ist.

Schließlich wird die Prüfgröße D in eine annähernd standardisierte Prüfgröße transformiert. Der Wert der standardisierten Prüfgröße y wird in der Zelle "Prüfgröße" angezeigt:

$$y = \frac{\sqrt{n}(D - 0.28209479)}{0.02998598}$$

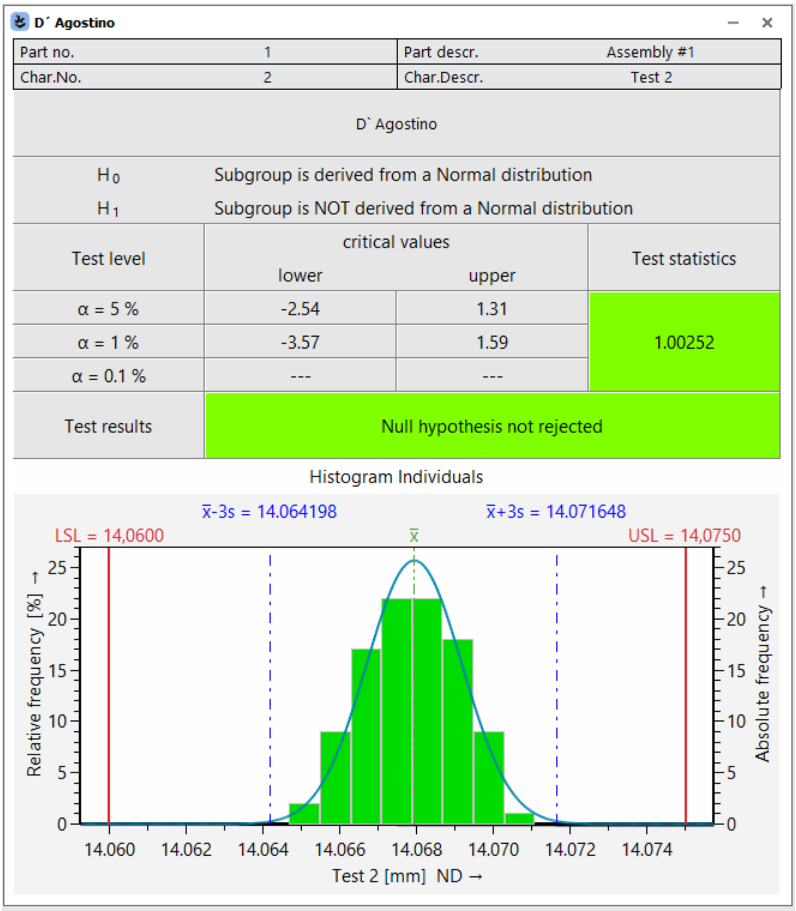


Abbildung 13: Fenster "D' Agostino" (D' Agostino's D-Test)

In der Software sind die kritischen Werte für diesen Test vertafelt. Kritische Werte stehen für Stichproben mit einem Gesamtumfang im Wertebereich von $n = 50$ bis $n = 1000$ zur Verfügung.

Referenz:

Ralph B. D'AGOSTINO, Michael A. STEPHENS
GOODNESS-OF-FIT-TECHNIQUES
CRC-Press (Taylor & Francis Group) 1986
ISBN 978-0-8247-7487-6
Kapitel 9.3.4.2 D' Agostino's D-Test;
Kritische Werte der Prüfgröße Y: Tabelle 9.7.

1.3.1.4 Shapiro-Wilk

Multifunktionsleiste:
Registerkarte <Ergebnisse> | <Testverfahren> | <Shapiro-Wilk>

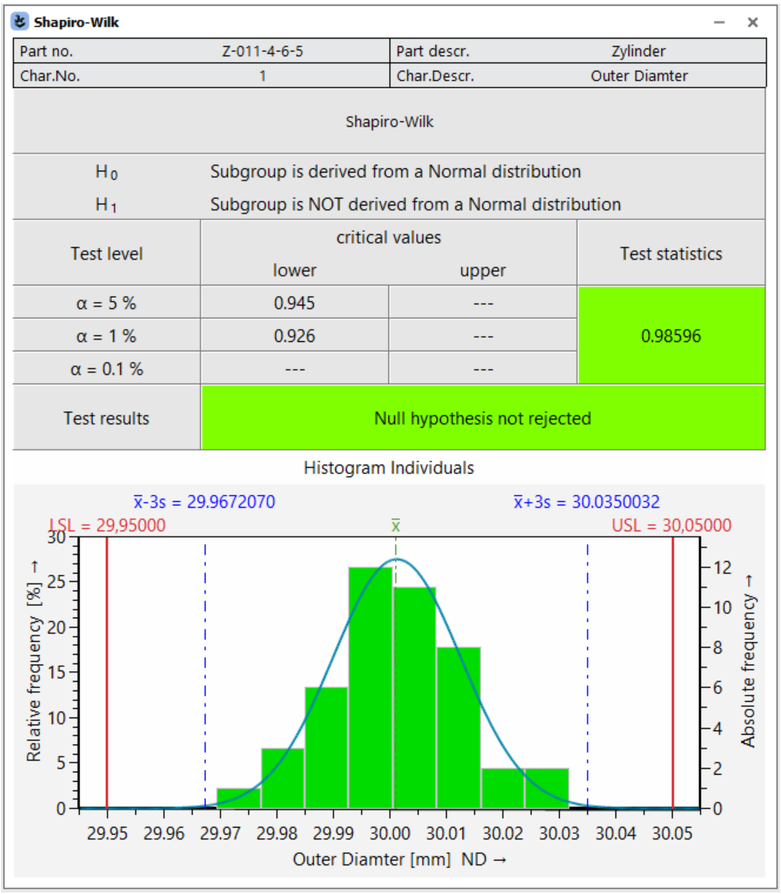
Der Shapiro-Wilk-Test auf Normalverteilung wird gemäß der internationalen Norm ISO 5479: 1997 durchgeführt. Dieses Verfahren basiert auf der Regression der Rangwerte mit ihren Erwartungswerten. Es handelt sich um einen Test vom Typ Varianzanalyse für eine vollständige Stichprobe. Die Prüfgröße W des Shapiro-Wilk-Tests wird wie folgt berechnet:

$$W = \frac{S^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

wobei

$$S = \sum_{i=1}^{k} a_k\left[x_{(n+1-k)} - x_k\right];\, k = \begin{cases} \frac{n}{2} & n \text{ is even} \\ \frac{n-1}{2} & n \text{ is odd} \end{cases}$$

Werte von a_k: siehe ISO 5479: 1997, Tabelle 10 - Shapiro-Wilk-Testkoeffizienten a_k zur Berechnung der Statistik des Tests W Seite 28.



Shapiro-Wilk

Part no.	Z-011-4-6-5	Part descr.	Zylinder
Char.No.	1	Char.Descr.	Outer Diamter

Shapiro-Wilk

H_0	Subgroup is derived from a Normal distribution
H_1	Subgroup is NOT derived from a Normal distribution

Test level	critical values lower	critical values upper	Test statistics
α = 5 %	0.945	---	0.98596
α = 1 %	0.926	---	
α = 0.1 %	---	---	
Test results	Null hypothesis not rejected		

Abbildung 14: Fenster "Shapiro-Wilk"

In der Software sind die kritischen Werte für das Testverfahren vertafelt. Es stehen kritische Werte für Stichproben mit einem Gesamtumfang im Bereich von $n = 3$ bis $n = 50$ zur Verfügung.

Referenz:

ISO 5479: 1997, Kapitel 8.2 und Tabelle 11 auf Seite 40.

1.3.1.5 Epps-Pulley

Multifunktionsleiste:
Registerkarte <Ergebnisse> | <Testverfahren> | <Epps-Pulley>

Der Epps-Pulley-Test auf Normalverteilung wird gemäß der internationalen Norm ISO 5479: 1997 durchgeführt. Es handelt sich um einen Omnibus-Test, der eine hohe Aussagekraft gegenüber vielen alternativen Hypothesen hat.

Die Berechnung der Prüfgröße T_{EP} des Epps-Pulley-Tests erfordert keine sortierten Stichprobendaten x aber die gewählte Reihenfolge der Daten muss während der gesamten Berechnung unverändert bleiben:

$$T_{EP} = 1 + \frac{n}{\sqrt{3}} + \frac{2}{n}\sum_{k=2}^{n}\sum_{j=1}^{k-1} exp^{\left\{\frac{-(x_j - x_k)^2}{2m_2}\right\}} - \sqrt{2}\sum_{j=1}^{n} exp^{\left\{\frac{-(x_j - \bar{x})^2}{4m_2}\right\}}$$

Hierbei ist m_2 das zentrale Moment zweiter Ordnung und n der Gesamtumfang der Stichprobe.

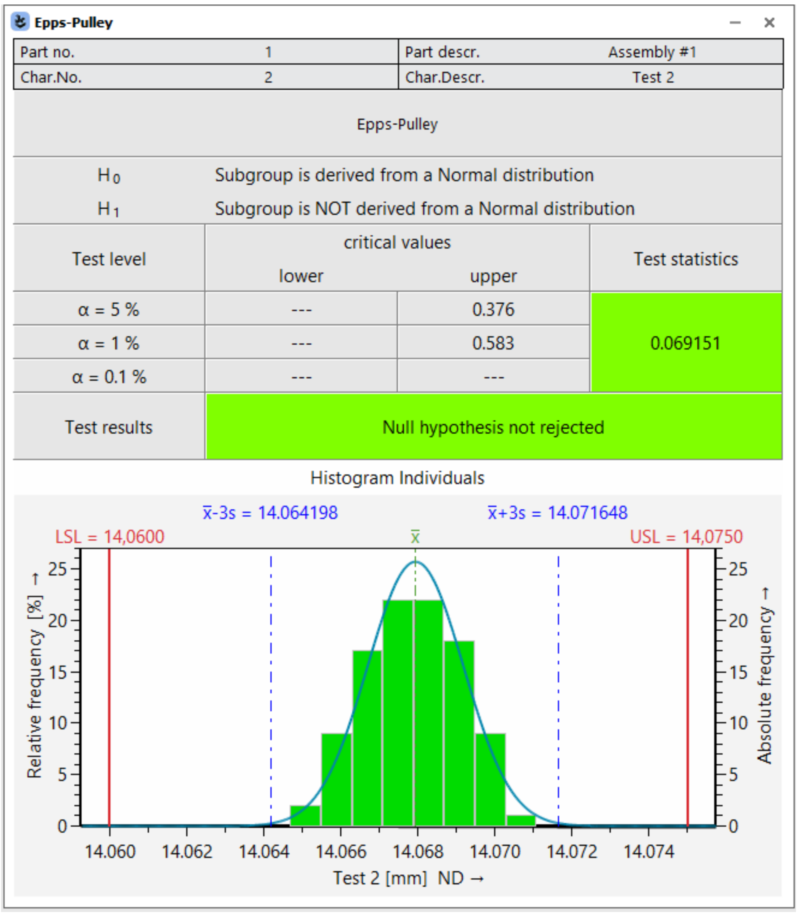


Abbildung 15: Fenster "Epps-Pulley"

Die kritischen Werte für den Test sind in der Software vertafelt. Diese kritischen Werte sind für Stichproben mit einem Gesamtumfang im Bereich von $n = 8$ bis $n = 200$ verfügbar.

Referenz:

ISO 5479: 1997, Kapitel 8.3 und Tabelle 12 auf Seite 30.

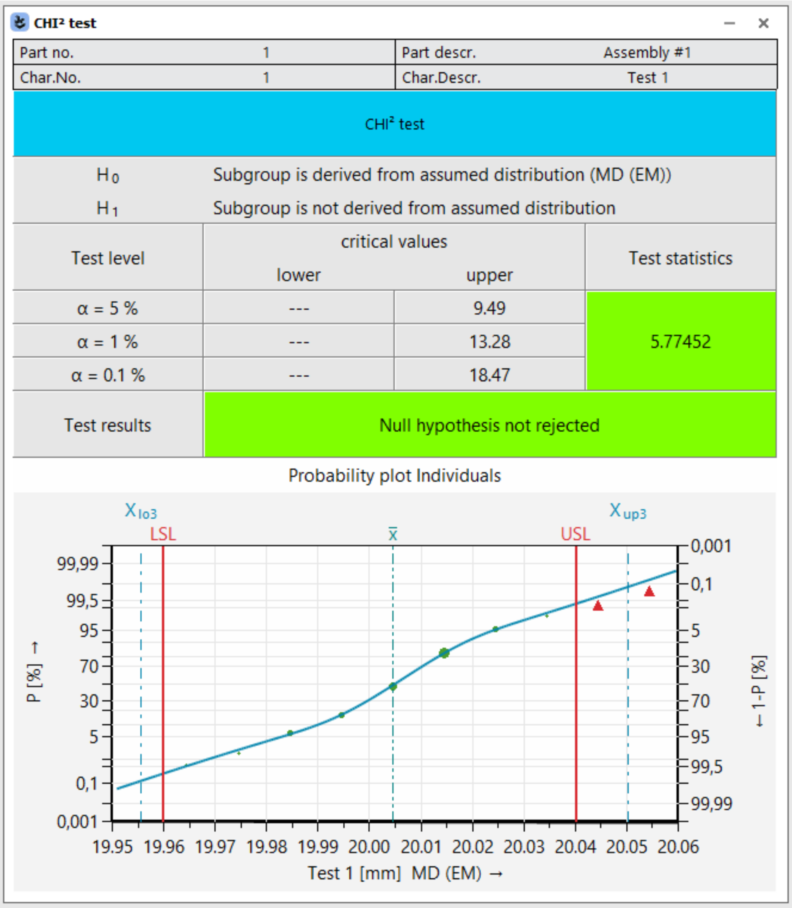
1.3.1.6 Chi²(C)

Multifunktionsleiste:
Registerkarte <Ergebnisse> | <Testverfahren> | <CHI²(C)>

Dieser Test auf Anpassungsgüte dient dazu, ein zuvor ausgewähltes *nicht-normalverteiltes* Verteilungsmodell (z. B. eine Weibullverteilung) zu testen. Der erste Schritt besteht darin, den Wertebereich der Stichprobendaten in Klassen (=Intervalle gleicher Breite) zu unterteilen. Anschließend wird gezählt, wie viele Werte in jede Klasse vorkommen. Die Klassen werden als aufeinanderfolgende, sich nicht überschneidende Intervalle bestimmt. Anschließend wird für jede Klasse die beobachtete Häufigkeit und die theoretisch erwartete Häufigkeit bestimmt. Die Berechnung der erwarteten Häufigkeit basiert auf dem ausgewählten (nicht normalverteilten) Verteilungsmodell. Die Prüfgröße lautet:

$$\chi^2 = \sum_{i=1}^{k} \frac{(o_i - e_i)^2}{e_i}$$

Dabei ist o_i die beobachtete Häufigkeit an Werten in der $i$ten Klasse und e_i die erwartete Häufigkeit an Werten in derselben Klasse. Die Variable k ist die Anzahl der Klassen.



CHI² test

Part no.	1	Part descr.	Assembly #1
Char.No.	1	Char.Descr.	Test 1

CHI² test

H_0	Subgroup is derived from assumed distribution (MD (EM))
H_1	Subgroup is not derived from assumed distribution

Test level	critical values lower	critical values upper	Test statistics
α = 5 %	---	9.49	5.77452
α = 1 %	---	13.28	
α = 0.1 %	---	18.47	
Test results	Null hypothesis not rejected		

Abbildung 16: Fenster "CHI²-Test" mit dem Ergebnis des Tests für den Datensatz "test_01.dfq"

Die kritischen Werte werden anhand der inversen Verteilungsfunktion (Quantilfunktion) der χ^2 −Verteilung mit $df = k - a - 1$ Freiheitsgraden (k = Anzahl der Klassen, a = Anzahl der geschätzten Verteilungsparameter).

Der CHI²-Test auf Anpassungsgüte wird durchgeführt, wenn *alle* erwarteten Häufigkeiten größer als 1 sind und höchstens 20% der erwarteten Klassenhäufigkeiten kleiner als 5 sind. Um diese Anforderung zu erfüllen, fasst die Software "Randklassen" zu einer einzigen (breiteren) Klasse zusammen.

Achtung: Das Ergebnis dieses Tests ist von der gewählten Klassierung abhängig. Mit Hilfe eines Histogramms kann zunächst überprüft werden, ob der Test für den jeweiligen Datensatz grundsätzlich geeignet ist. Das Ergebnis ist unbrauchbar, wenn einige Klassen nicht besetzt sind. Dieser Test sollte nur für Stichproben mit einem Mindestgesamtumfang von $n = 100$ Werten (oder gerne mehr) verwendet werden.

Referenz:

Ulrich GRAF, Hans-Joachim HENNING, Kurt STANGE, Peter-Theodor WILRICH
Formeln und Tabellen der angewandten mathematischen Statistik
Zweiter Nachdruck der dritten Auflage
Springer 1998
ISBN 3-540-16901-6
Seiten 131-132

1.3.1.7 Anderson Darling-Test

Multifunktionsleiste:
Registerkarte <Ergebnisse> | <Testverfahren> | <Anderson Darling-Test>

Der Anderson Darling-Test basiert auf der Empirischen Verteilungsfunktion. Die Software berechnet die Prüfgröße wie folgt:

- Die Beispieldaten werden in aufsteigender Reihenfolge angeordnet.
- Die sortierten Stichprobendaten werden in standardisierte Daten transformiert:
  $z_i = \frac{x_i - \bar{x}}{s}$ für $i = 1, \ldots, n$
- Berechnet wird der Wert der Verteilungsfunktion der Standardisierten Normalverteilung für jeden der transformierten Werte:
  $P_i = G(z_i)$ für $i = 1, \ldots, n$
- Berechnung der Anderson-Darling-Prüfgröße:
  $$A^2 = \sum_{i=1}^{n} [(2i-1)\{\ln(P_i) + \ln(1 - P_{n+1-i})\}/n] - n$$
- Berechnung der modifizierten Prüfgröße:
  $$A^* = A^2 \left(1.0 + \frac{3}{4n} + \frac{9}{4n^2}\right)$$

Hierbei ist:

n Gesamtumfang der Stichprobe.

x_i die Stichprobenwerte (in aufsteigender Reihenfolge!)

$\bar{x}$ Mittelwert der Gesamtstichprobe.

s Standardabweichung der Gesamtstichprobe.

$G()$: Verteilungsfunktion der standardisierten Normalverteilung.

Berechnung der kritischen Werte und des P-Wertes

Signifikanz-Niveau	Kritische Werte für die Prüfgröße A*
5 %	0.752
1 %	1.035
0.1 %	1.445

Der P-Wert für die modifizierte Prüfgröße A^*:

1. Wenn $A^* < 0.2$, dann:
   $\log\_q = -13.436 + 101.14 \times A^* - 223.73 \times (A^*)^2$
   $p = 1 - \exp(\log\_q)$
2. Wenn $0.2 \leq A^* < 0.34$, dann:
   $\log\_q = -8.318 + 42.796 \times A^* - 59.938 \times (A^*)^2$
   $p = 1 - \exp(\log\_q)$
3. Wenn $0.34 \leq A^* < 0.6$, dann:
   $\log\_p = 0.9177 - 4.279 \times A^* - 1.38 \times (A^*)^2$
   $p = exp(\log\_p)$
4. Wenn $0.6 \leq A^*$, dann:
   $\log\_p = 1.2937 - 5.709 \times A^* + 0.0186 \times (A^*)^2$
   $p = exp(\log\_p)$

Referenz:
Ralph B. D'AGOSTINO, Michael A. STEPHENS
GOODNEESS-OF-FIT TECHNIQUES
CRC-Press (Taylor & Francis Group, LLC) 1986
ISBN 978-0-8247-7487-6
Kapitel 4.8.2 und 9.3.2.2

1.3.2 Ausreißer-Tests

1.3.2.1 Verhalten der Software beim Löschen von Werten (Modul Prozessanalyse):

In der Auswertungsstrategie des Moduls "Prozessfähigkeit" gibt es drei verschiedene Optionen für die Behandlung der *verbleibenden Werte einer Stichprobe*, wenn ein oder mehrere Werte gelöscht werden. Mit einem Klick auf das Symbol "Vorbereitung" in der Auswertestrategie, wird das Fenster "Vorbereitung der Auswertung" geöffnet. Anschließend auf die Registerkarte "Ausreißer" klicken. Die entsprechenden Behandlungsregeln sind im unteren Teil des folgenden Bildes dargestellt:

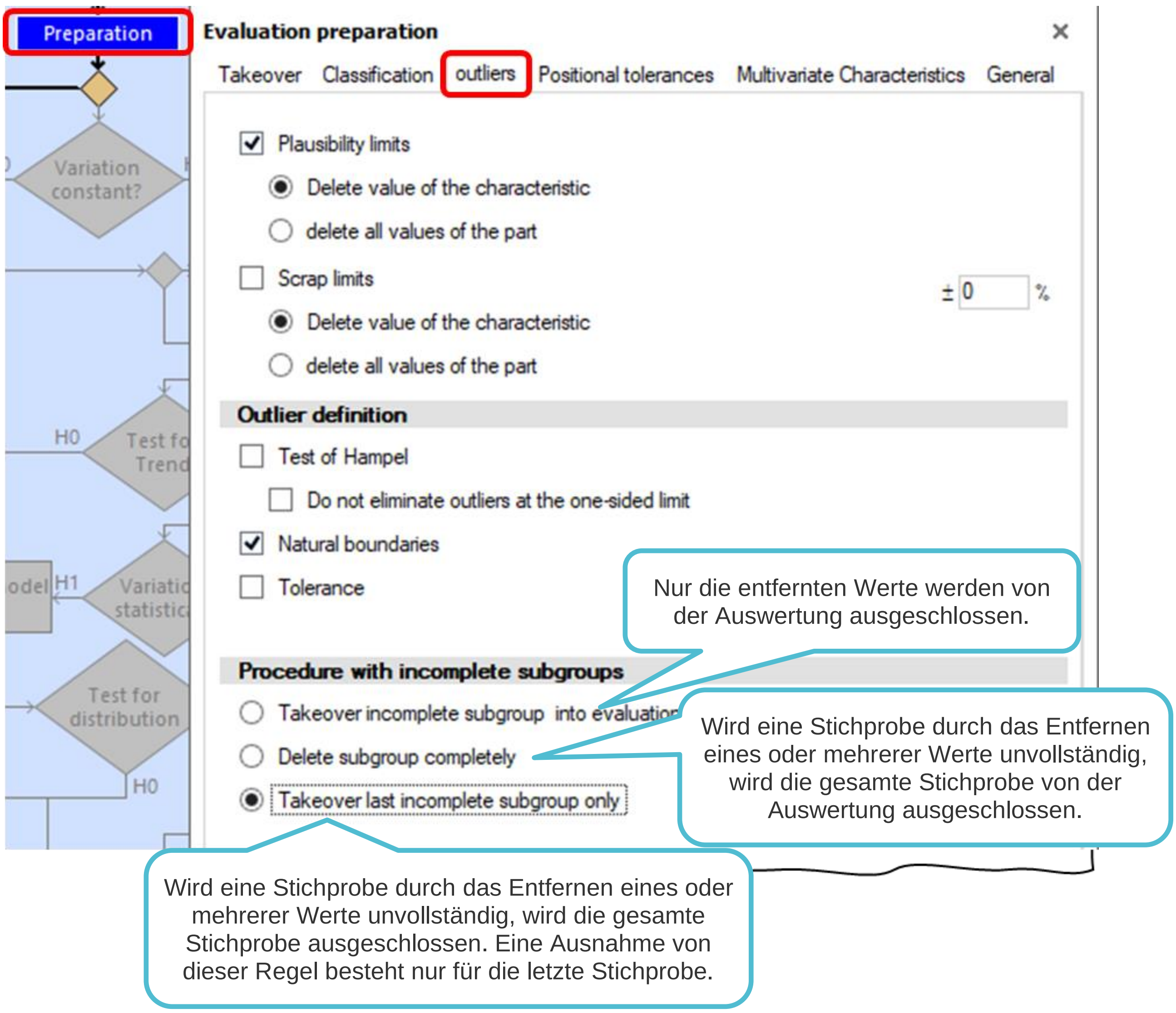
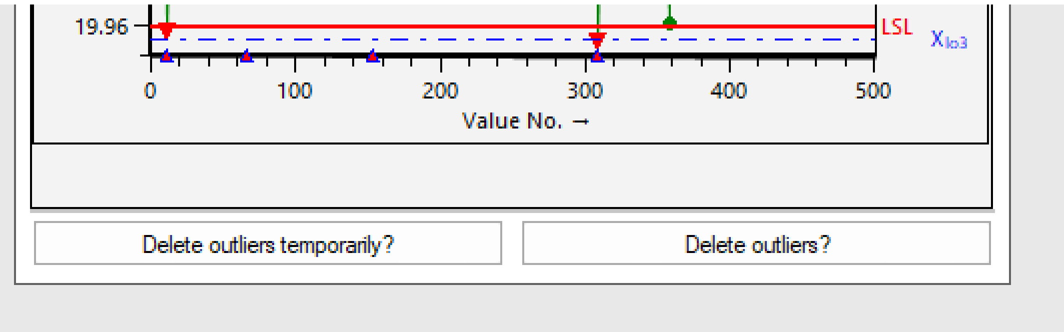


Abbildung 17: Auszug aus der Auswertungsstrategie (Modul Prozessfähigkeit) mit den Optionseinstellungen für die Behandlung der verbleibenden Werte einer Stichprobe, wenn wir mindestens einen Wert der jeweiligen Stichprobe entfernen. Die Behandlung der verbleibenden Werte einer Untergruppe ist auch dann relevant, wenn die Software einen Ausreißer automatisch entfernt (Ausreißertest, Schrottgrenze(n), natürliche Grenze, Plausibilitätsgrenze(n)).

1.3.2.2 Allgemeines Verfahren zur Beseitigung von Ausreißern

Wird ein Ausreißertest durchgeführt und werden Ausreißer festgestellt, stehen 2 Schaltflächen zur Verfügung: "Ausreißer temporär entfernen?" und "Ausreißer entfernen?".



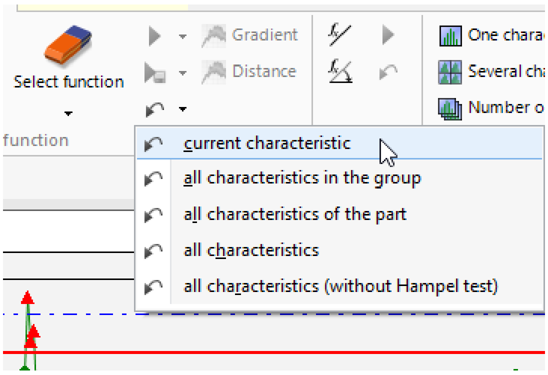
Die englische Übersetzung ist ein wenig irreführend. "Löschen" bedeutet in diesem Fall keinen echten "Löschvorgang", sondern die betreffenden Werte werden bei Auswertungen nicht berücksichtigt.

Ausreißer temporär entfernen

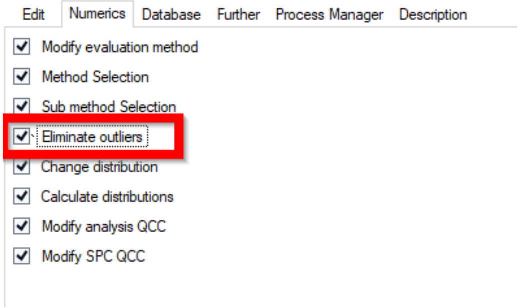
Die betreffenden Werte werden im Rahmen der Sitzung bei Auswertungen nicht berücksichtigt. Wird der Datensatz erneut geöffnet, werden die zuvor temporär entfernten Werte wieder in die Berechnung mit einbezogen.

Ausreißer entfernen

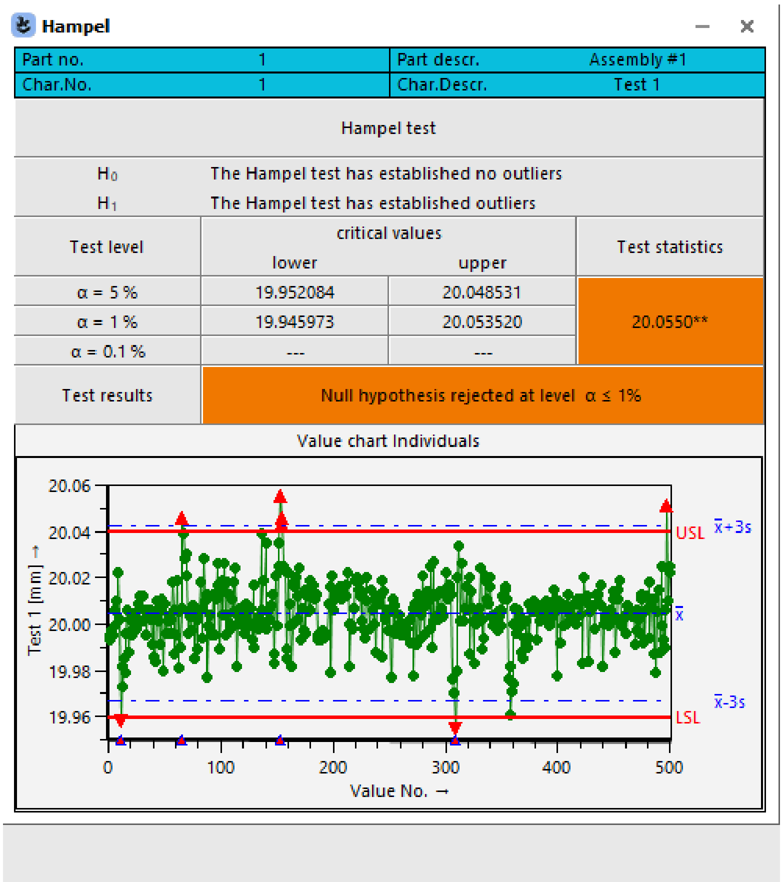
Die betreffenden Werte sind mit einem Attribut gekennzeichnet (K0002 = 191). Sollen bereits gespeicherte Ausreißer wieder in die Berechnung eingebracht werden, ist dies bei geöffneter Grafik „Werteverlauf Einzelwerte“ und anschließender Auswahl des Befehls „Selekt-Funktionen rückgängig“ (Register <Teil/Merkmal> | <Selekt-Funktion rückgängig>) möglich:



Das Recht zum Entfernen von Werten mit einem Ausreißertest haben nur solche Benutzer, die zu einer Benutzergruppe gehören, für die das Recht "Ausreißer entfernen" aktiviert wurde (Das Fenster „Benutzerverwaltung“ ist erreichbar über: Register <Datei> | <Konfigurationen> | <Benutzerverwaltung>).



Ohne dieses Recht werden in einem Fenster eines Ausreißertests die (oder der) Ausreißer zwar angezeigt, aber der Anwender kann diese Ausreißer nicht entfernen.



1.3.2.3 Ausreißer Grubbs (min)

Multifunktionsleiste:
Registerkarte <Ergebnisse> | <Testverfahren> | <Ausreißer (Grubbs min)>

Der Grubbs-Test (min) prüft für normalverteilte Stichproben, ob der Kleinstwert x_{min} ein Ausreißer ist.

Die Prüfgröße ist die Differenz aus Mittelwert und Kleinstwert geteilt durch die Standardabweichung (Mittelwert und Standardabweichung der Gesamtstichprobe):

$$G_{min} = \frac{(\bar{x} - x_{min})}{s}$$

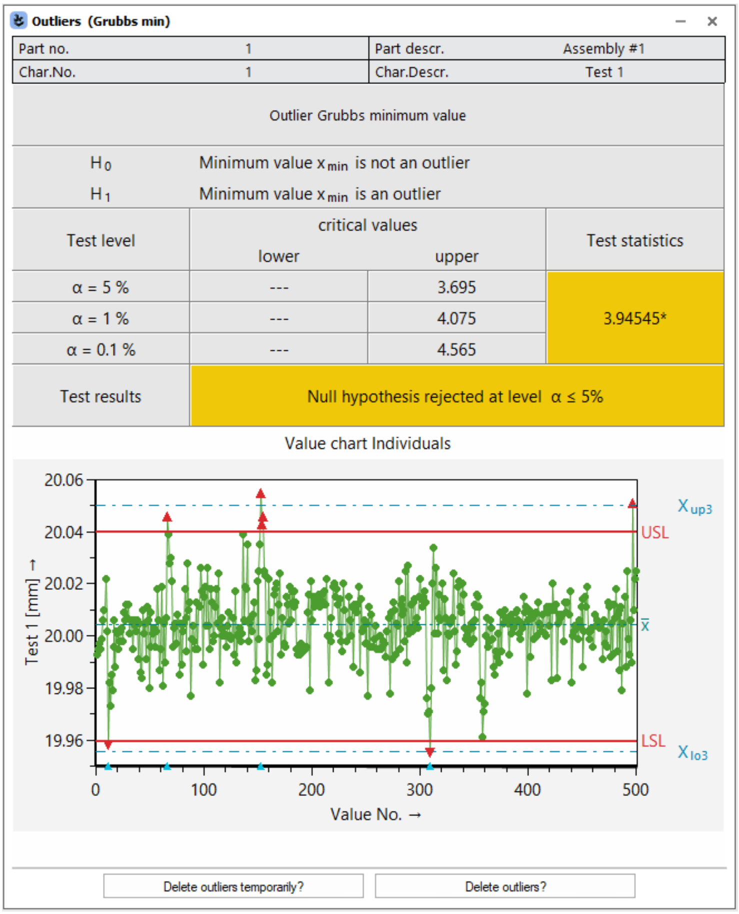


Abbildung 18: Fenster "Ausreißer (Grubbs min)" mit dem Ergebnis für den Datensatz "Test_01.dfq"

In der Software werden die kritische Werte (einseitige obere Grenzen der Prüfgröße) für die Signifikanzniveaus $\alpha = 0.05,\ 0.01$ und 0.001 berechnet:

$$T_{n,\alpha} = \frac{(n-1)t_{1-\frac{\alpha}{n};n-2}}{\sqrt{n\left(n-2+\left(t_{1-\frac{\alpha}{n};n-2}\right)^2\right)}}$$

Hierbei ist:

n : Gesamtumfang der Stichprobe.

$t_{1-\frac{\alpha}{n};n-2}$: Das $(1-\alpha)$-Quantil der Studenten-Verteilung mit $df = n-2$ Freiheitsgraden.

Referenz:

Lothar SACHS, Juergen HEDDERICH
Angewandte Statistik
17. Auflage (2020)
Springer Spektrum
ISBN 978-3-662-62293-3
Abschnitt 7.2.9.1 Grubbs-Test für Ausreißer
Formel 7.37 auf der Seite 502

Bemerkung: Die in unserer Software implementierte Formel berechnet die *einseitig* obere Grenze der Prüfgröße (die Formel im Buch berechnet die *zweiseitig* obere Grenze).

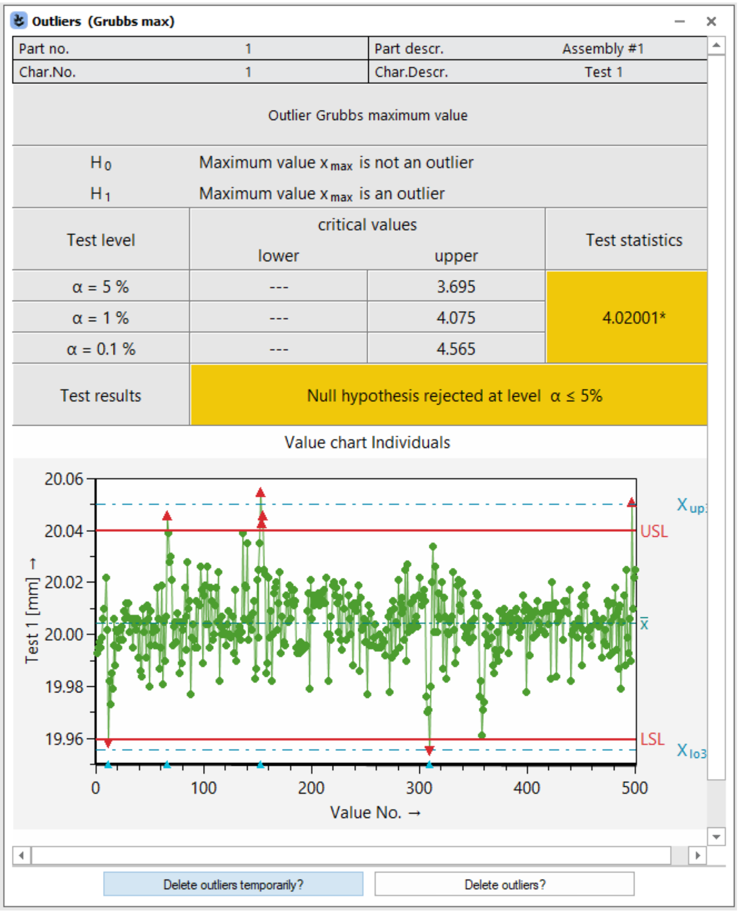
1.3.2.4 Ausreißer Grubbs (max)

Multifunktionsleiste:
Registerkarte <Ergebnisse> | <Testverfahren> | <Ausreißer (Grubbs max)>

Der Grubbs-Test (max) prüft für (normalverteilte) Stichproben, ob der Größtwert x_{max} ein Ausreißer ist.

Die Prüfgröße ist die Differenz aus dem Größtwert und dem Mittelwert geteilt durch die Standardabweichung (Mittelwert und Standardabweichung der Gesamtstichprobe):

$$G_{max} = \frac{(x_{max} - \bar{x})}{s}$$


Outliers (Grubbs max)

Part no.	1	Part descr.	Assembly #1
Char.No.	1	Char.Descr.	Test 1

Outlier Grubbs maximum value

H_0	Maximum value x_{max} is not an outlier
H_1	Maximum value x_{max} is an outlier

Test level	critical values lower	critical values upper	Test statistics
α = 5 %	---	3.695	4.02001*
α = 1 %	---	4.075	
α = 0.1 %	---	4.565	
Test results	Null hypothesis rejected at level α ≤ 5%		

Abbildung 19: Das Fenster "Ausreißer (Grubbs max)" zeigt das Ergebnis für den Datensatz "test_01.dfq".

Kritische Werte: dieselben wie beim Test "Ausreißer (Grubbs min)"

1.3.2.5 Ausreißer (David, Hartley, Pearson)

Multifunktionsleiste:
Registerkarte <Ergebnisse> | <Testverfahren> | <Ausreißer (David, Hartley, Pearson)>

Dieser Test prüft, ob *entweder* der Größtwert x_{max} *oder* der Kleinstwert x_{min} ein Ausreißer ist. Wird die Nullhypothese verworfen, identifiziert der Test nur den Extremwert mit dem größten Abstand zum arithmetischen Mittelwert als Ausreißer. Im Sonderfall, bei dem *sowohl* der Kleinstwert *als auch* der Größtwert beide den *gleichen* Abstand zum Mittelwert haben, identifiziert der Test bei einmaliger Durchführung *beide* Extremwerte als Ausreißer.

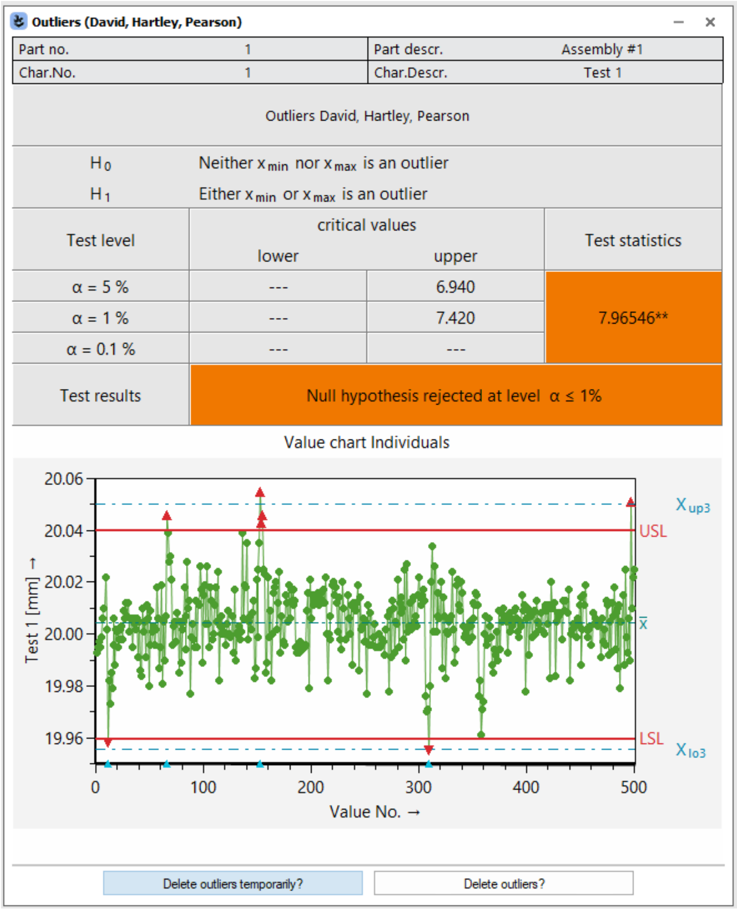


Abbildung 20: Das Fenster "Ausreißer (David, Hartley, Pearson)" zeigt das Ergebnis für den Datensatz Test_01.dfq.

Die Prüfgröße lautet:

$$q = \frac{R}{s}$$

Hierbei ist.

R : Spannweite der Gesamtstichprobe.

s : Standardabweichung der Gesamtstichprobe.

In der Software sind die kritischen Werte für diesen Test vertafelt. Kritische Werte stehen für Stichproben mit einem Gesamtumfang im Bereich von $n = 3$ bis $n = 1000$ zur Verfügung.

Referenz:

Ulrich GRAF, Hans-Joachim HENNING, Kurt STANGE, Peter-Theodor WILRICH
Formeln und Tabellen der angewandten mathematischen Statistik
3. Auflage (1998),
Springer Verlag
ISBN: 3-540-16901-6
Tabelle 6.12 - Kritische Werte $q_{n,1-\alpha}$ für den David-Hartley-Pearson-Test
Seite 144

1.3.2.6 Hampel-Test

Multifunktionsleiste:
Registerkarte <Ergebnisse> | <Testverfahren> | <Hampel>

Der Hampel-Test identifiziert mehrere Ausreißer gleichzeitig. In unserer Software gibt es zwei Varianten dieses Testverfahrens:

(1) Hampel-Test (Standard): Der klassische Hampel-Ausreißer-Test

(2) Hampel-Kölling-Test: Werner KOELLING hat[2] eine Heuristik vorgeschlagen, um den Hampel-Test bei schief verteilten Stichprobendaten anwenden zu können.

Im Fenster "Hampel-Test" wird *entweder* das Ergebnis des Standard-Hampel-Tests *oder* das Ergebnis der modifizierten Variante (Hampel-Kölling-Test) angezeigt. Welche der beiden Varianten verwendet wird, hängt von den entsprechenden Optionseinstellungen in der Auswertungsstrategie ab:

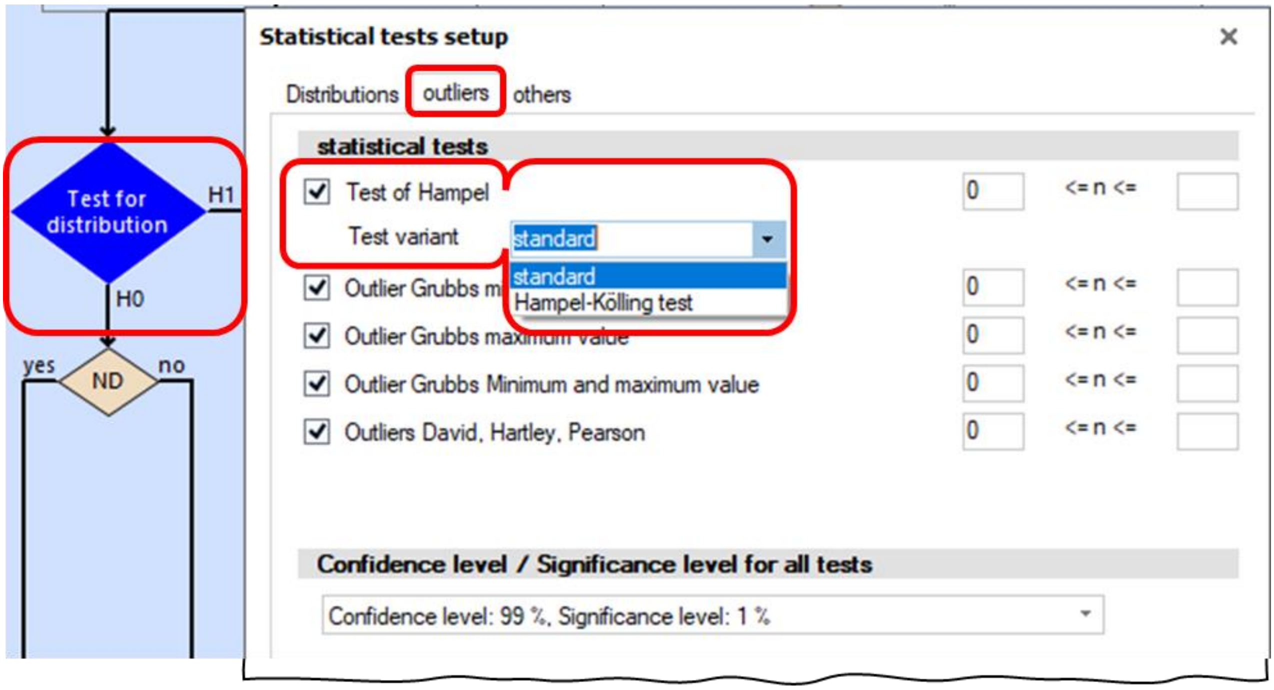


Abbildung 21: Auszug aus der Auswertungsstrategie: Einstellungen des Hampel-Ausreißertests.

1.3.2.6.1 Hampel-Ausreißertest (Standard)

Als Vorbedingung setzt dieser Test eine *symmetrische* Verteilung voraus. Um robust gegenüber dem Einfluss von Ausreißern zu sein, wird der Streuungsparameter mit Hilfe der absoluten Abweichung vom Median (MAD) geschätzt.

$$MAD = \frac{MED(|x_1 - \tilde{x}|, \dots, |x_i - \tilde{x}|, \dots, |x_n - \tilde{x}|)}{0.6745}$$

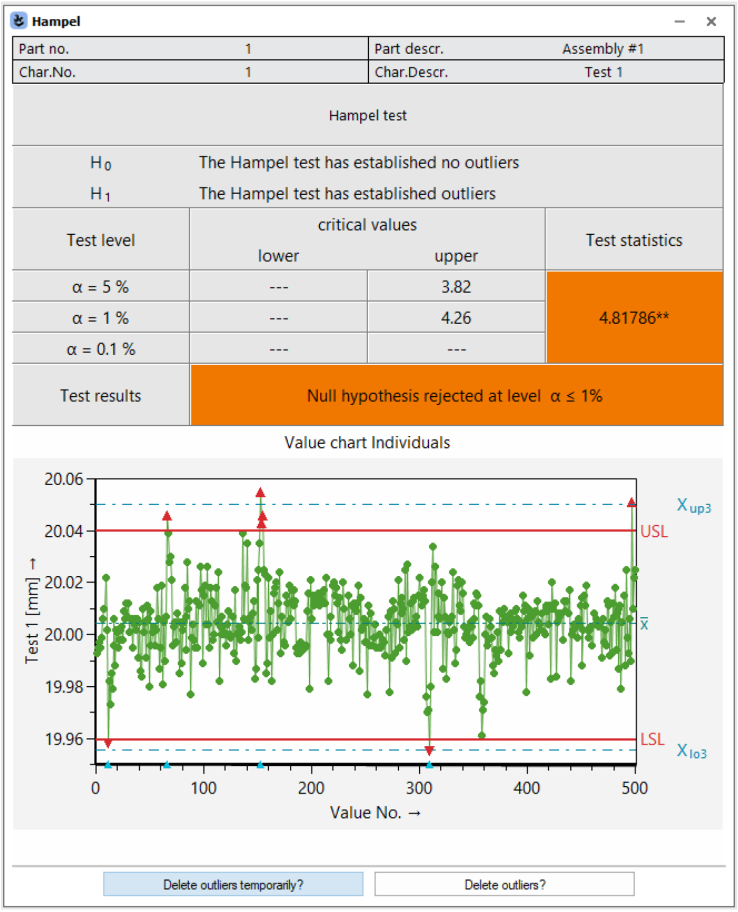
wobei $\tilde{x}$ der Median aller Stichprobendaten ist und MED() eine Funktion zur Berechnung des Medians ist.

Für jeden Wert der Stichprobe x_i wird die folgende Statistik berechnet:

[2] Werner Kölling: Aus der Reihe getanzt, Qualität und Zuverlässigkeit 46 (2001), S. 315-319

$$y_i = \frac{|x_i - \tilde{x}|}{MAD}$$

Bei diesem Verfahren werden mehrere Ausreißer gleichzeitig ermittelt (falls vorhanden). In unserer Software wird nur die Prüfgröße für den extremsten Wert angezeigt, also entweder die Prüfgröße des Größtwertes oder des Kleinstwertes. Im Hintergrund (für uns unsichtbar) werden alle Ausreißer ermittelt.



Hampel

Part no.	1	Part descr.	Assembly #1
Char.No.	1	Char.Descr.	Test 1

Hampel test

H_0	The Hampel test has established no outliers
H_1	The Hampel test has established outliers

Test level	critical values lower	critical values upper	Test statistics
α = 5 %	---	3.82	4.81786**
α = 1 %	---	4.26	
α = 0.1 %	---	---	
Test results	Null hypothesis rejected at level α ≤ 1%		

Abbildung 22: Das Fenster "Hampel" zeigt das Ergebnis des Hampel-Ausreißertests (Standard) für den Datensatz "Test-01.dfq".

In der Software sind die kritischen Werte für diesen Test vertafel. Kritische Werte stehen für Stichproben mit einem Gesamtumfang im Bereich von $n = 10$ bis $n = 200$ zur Verfügung. Die folgende Grafik (nicht in unserer Software verfügbar) zeigt die Prüfgrößenwerte für alle Werte des Test-Datensatzes "Test_01.dfq".

Mit dem Klick auf die Schaltfläche "Ausreißer entfernen?", werden *alle* identifizierten Ausreißer entfernt. Das gilt für den Datensatz Test_01.dfq:

- Die Software identifiziert vier Ausreißer gleichzeitig, wenn in der Auswertungsstrategie das Signifikanzniveau von $\alpha = 1$ % für alle Ausreißer-Tests verwendet wird.
- Die Software identifiziert sieben Ausreißer gleichzeitig, wenn in der Auswertungsstrategie das Signifikanzniveau von $\alpha = 5$ % für alle Ausreißer-Tests verwendet wird.

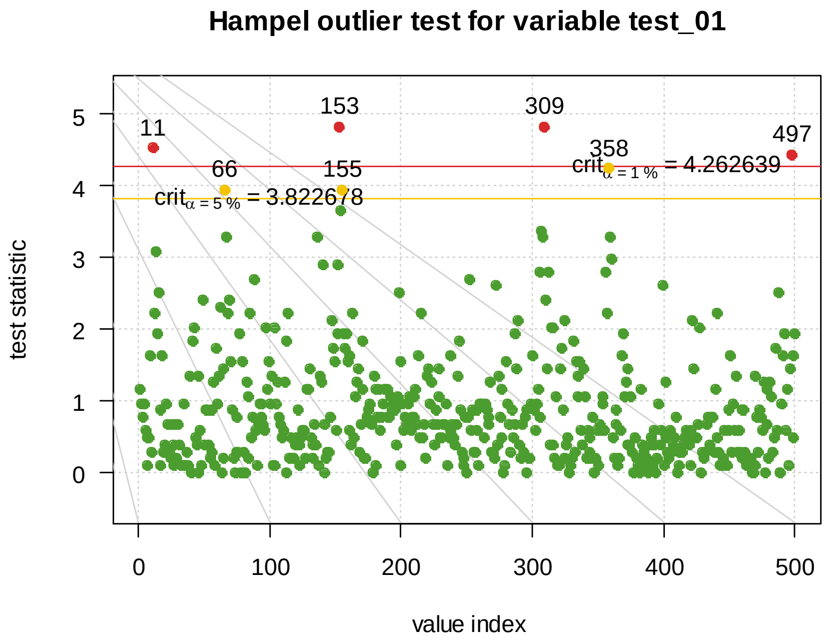


Abbildung 23: Visualisierung der gleichzeitigen Erkennung von multiplen Ausreißern des Hampel-Standardtests; Testdatensatz "test_01.dfq". Diese Grafik ist in unserer Software nicht verfügbar.

Mit einem Klick auf die Schaltfläche "*Ausreißer entfernen?" oder auf die Schaltfläche "Ausreißer temporär entfernen?*", wertet das Programm die Daten erneut aus und aktualisiert damit die berechneten Werte des Tests.

Wenn zunächst die Schaltfläche *"Ausreißer löschen*" angeklickt und der Datensatz anschließend gespeichert wird, sind die Ausreißer mit dem Attribut 191 im Datensatz gekennzeichnet. Beim erneuten Öffnen des Datensatzes sind die Ausreißer aufgrund dieses Attributs bereits entfernt.

Referenz:

Christian KELLER, Ausreißer - und dann?, Qualität und Zuverlässigkeit QZ(44) 1999, Tabelle 3 - Kritische Grenzen des zweiseitigen Hampel-Tests ..., . S. 93

DAVIES, L.; GATHER, U., The identification of multiple outliers, Journal of American Statistical Association 88 (1993) 423, S. 782-792

1.3.2.6.2 Hampel-Kölling-Test

Multifunktionsleiste:
Registerkarte <Ergebnisse> | <Testverfahren> | <Hampel>

Diese Variante des Hampel-Ausreißertests basiert auf einer von Werner Kölling vorgeschlagenen Heuristik und erweitert die Anwendung auf asymmetrische Verteilungen (J- oder L-förmig), bei denen eine Seite der Stichprobenverteilung einen längeren Auslauf hat als die andere Seite (am häufigsten beobachtet für Daten eines GD&T-Merkmales).
Die Lösung von Kölling beruht auf der Idee, die beiden Hälften der Stichprobenverteilung - die linke Seite und die rechte Seite vom Median - unterschiedlich zu behandeln:

Für den gegebenen Gesamtstichprobenumfang n und das gewählte Signifikanzniveau αwerden die kritischen Werte $T_{n,1-\alpha}$ der Standardversion des Hampel-Tests bestimmt. Anschließend werden die folgenden Formeln zur Berechnung der Ausreißer-Grenzen verwendet:

Obere Hampel-Kölling-Grenze für Ausreißer (rechts vom Median):

$$OHKG = \tilde{x} + \frac{\left|Q_{p.oben} - \tilde{x}\right|}{z_{p.oben}} \times T_{n,1-\alpha}$$

Untere Hampel-Kölling-Grenze für Ausreißer (links vom Median):

$$UHKG = \tilde{x} + \frac{\left|Q_{p.unten} - \tilde{x}\right|}{z_{p.unten}} \times T_{n,1-\alpha}$$

Hierbei ist:
$\tilde{x}$: Median der Gesamtstichprobe.
Q_p : Das empirische p-Quantil der Stichprobe.
z_p : Das p-Quantil der gespeicherten Verteilung der Standardisierten Normalverteilung.

Der Parameter (Wahrscheinlichkeit) p wird in Abhängigkeit von der Gesamtstichprobengröße aus der folgenden Tabelle entnommen n:

Umfang Stichprobe	$p.unten$	$p.oben$
10 ... 25	0.25	0.75
26 ... 40	0.20	0.80
41 ... 80	0.15	0.85
81 ... 120	0.10	0.90
≥ 120	0.05	0.95

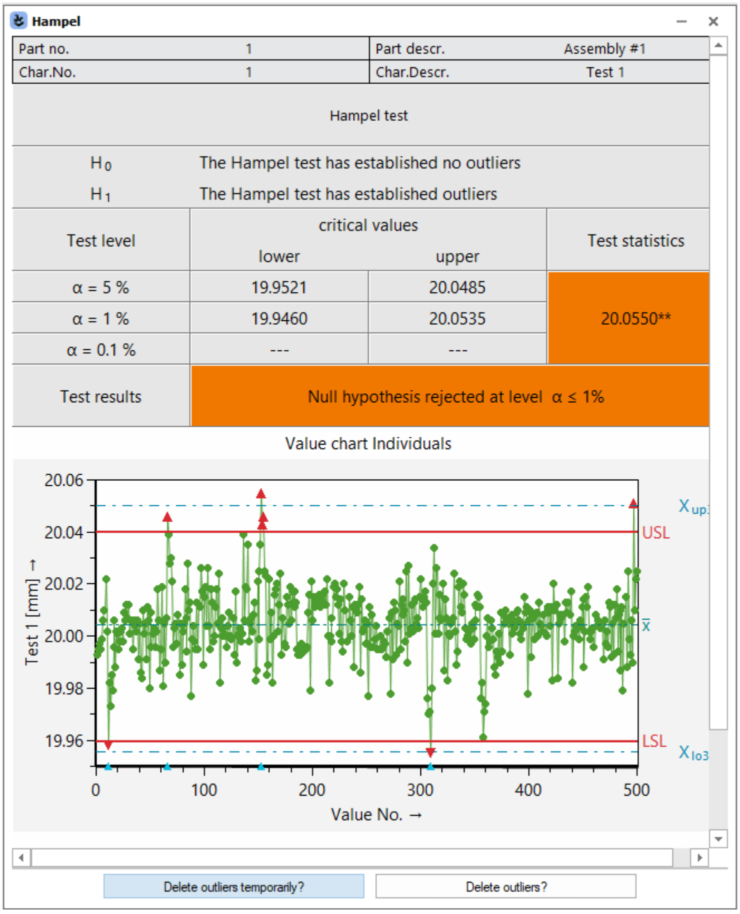


Abbildung 24: Das Fenster "Hampel" zeigt das Ergebnis des Hampel-Kölling-Tests für den Datensatz "test_01.dfq".

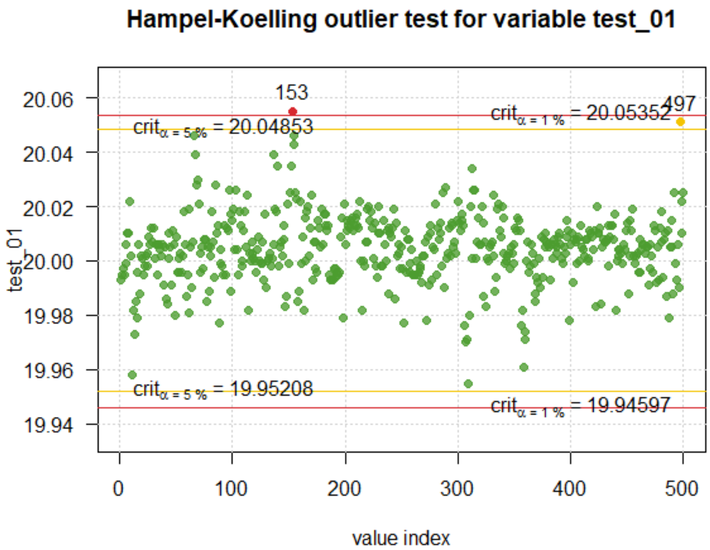


Abbildung 25: Visualisierung der Ausreißer-Grenzen des Hampel-Kölling-Ausreißertests für den Datensatz "test_01.dfq". Diese Grafik ist in unserer Software nicht verfügbar.

Referenz:
Werner KOELLING,
Aus der Reihe getanzt
Qualität und Zuverlässigkeit (QZ) 46, 2001
Carl Hanser, S. 315-319

1.3.3 Tests auf homogene Lage

1.3.3.1 Kruskal-Wallis

Multifunktionsleiste:
Registerkarte <Ergebnisse> | <Testverfahren> | <Kruskal-Wallis>

Dieser Test ist nur im Softwaremodul "Prozessfähigkeit" verfügbar. Für ein einzelnes Merkmal prüft der Test, ob alle Stichproben aus einer Grundgesamtheit mit *demselben Lageparameter* stammen. Die Analyse basiert auf den Rängen r_i der Stichprobendaten, nicht auf den Stichprobendaten selbst. Enthält eine Stichprobe wiederholte Werte (dieselben Werte treten multiple Male auf), wird den jeweiligen Werten der Mittelwert der Ränge zugewiesen.

Die Kruskal-Wallis-Prüfgröße H ist wie folgt definiert:

$$H = \frac{1}{B}\left[\frac{12}{N(N-1)}\sum_{i=1}^{p}\frac{1}{n_i}r_i^2 - 3(N+1)\right]$$

Hierbei ist:

N : Gesamtumfang der Stichprobe.

n_i : Umfang der $i$ten Stichprobe.

p : Die Gesamtzahl der Stichproben. (Annahme: jede Stichprobe ist gleich groß)

B : Ein Korrekturblock für gebundene Werte:

$$B = 1 - \frac{1}{N^3 - N}\sum_{j=1}^{g}\left(t_j^3 - t_j\right)$$

Dabei ist g die Anzahl der Bindungs-Gruppen (Bindungen sind hier mehrfach vorkommende *gleiche* Werte). Der Wert t_j ist die *Anzahl der Werte in der j*ten Bindungs-Gruppe.

Kruskal-Wallis

Test of Kruskal and Wallis			
H_0	The Kruskal-Wallis test has established no location variations		
H_1	The Kruskal-Wallis test has established location variations		
Test level	critical values lower	critical values upper	Test statistics
α = 5 %	---	123.23	215.144***
α = 1 %	---	134.64	
α = 0.1 %	---	148.23	
Test results	Null hypothesis rejected at level α ≤ 0,1%		

Value chart Individuals

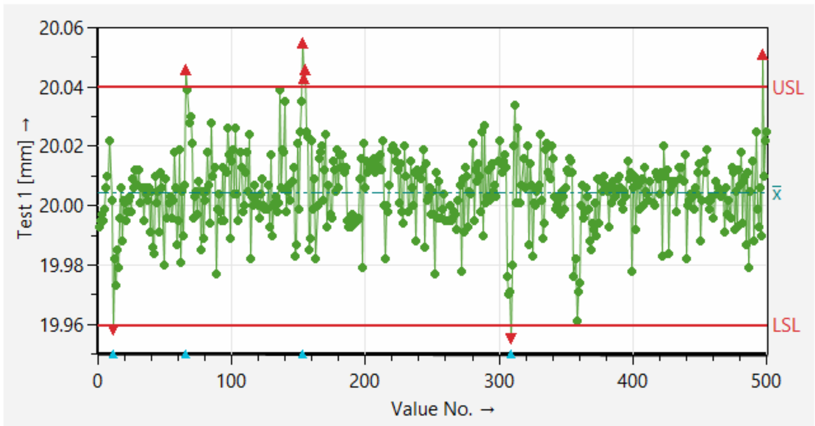


Abbildung 26: Fenster "Kruskal-Wallis" mit dem Ergebnis des Tests für den Datensatz "test_01.dfq.

Kritische Werte werden unter Verwendung der umgekehrten Verteilungsfunktion (Quantilfunktion) der Chi²-Verteilung für Signifikanzniveaus berechnet α=5 %, 1 % und 0.1 %:

$$H_{crit.\alpha} = \chi^2_{1-\alpha,\, p-1}$$

Referenz:

Lothar SACHS, Juergen HEDDERICH
Angewandte Statistik
17. Auflage (2020)
Springer Spektrum
ISBN 978-3-662-62293-3
Abschnitt 7.6.5 H-Test von Kruskal und Wallis
Formeln 7.237 und 7.240 auf der Seite 651 ff

1.3.3.2 ANOVA

Multifunktionsleiste:
Registerkarte <Ergebnisse> | <Testverfahren> | <ANOVA>

Mit diesem Testverfahren wird geprüft, ob alle Stichproben aus Grundgesamtheiten mit demselben Parameter µ stammen oder nicht. Was beim ersten Blick auf das Ergebnisfenster verwirrend sein kann: Warum sehen wir Varianzen und nicht Mittelwerte? Dieses Verfahren prüft, ob die Streuung *zwischen den Stichproben* - verursacht durch die unterschiedliche Lage der Subgruppenmittelwerte - größer ist, als die dafür (durch den Zufall) erwartete Streuung (➔ signifikant).

Der Einfachheit halber kann die gesamte ANOVA-Analyse übersprungen werden, und wir konzentrieren uns allein auf die Prüfgröße F. Diese wird wie folgt berechnet:

$$F = \frac{MS_{Factor}}{MS_{Error}}$$

wobei MS_{Error} und MS_{Factor} die berechneten mittleren Quadratsummen der einfachen ANOVA-Tabelle sind. Als Faktor-Stufenwerte dienen die Indizes der Stichproben-Untergruppen.

$$MS_{Factor} = \frac{SS_{Factor}}{df1} = \frac{n \times \sum_{i=1}^{p} (\bar{x}_i - \bar{\bar{x}})^2}{p - 1}$$

$$MS_{Error} = \frac{SS_{Error}}{df2} = \frac{\sum_{i=1}^{p} \sum_{k=1}^{n} (x_{i.k} - \bar{x}_i)^2}{p \times (n - 1)}$$

Hierbei sind:

p : Anzahl der Stichproben-Untergruppen.

n : Umfang der einzelnen Stichprobe-Untergruppe (*nicht* der Stichproben-Gesamtumfang!).
Nebenbei bemerkt: Es wird davon ausgegangen, dass alle Stichproben-Untergruppen den gleichen Umfang n haben (➔ balancierter Versuchsplan).

$x_{i,k}$: Der $k$te Wert der $i$ten Stichproben-Untergruppe.

$\bar{x}_i$: Mittelwert der $i$ten Stichproben-Untergruppe.

$\bar{\bar{x}}$: Mittelwert aller Stichproben-Mittelwerte.

ANOVA

Part no.	1	Part descr.	Assembly #1
Char.No.	1	Char.Descr.	Test 1

ANOVA

Variation within subgroups	=	s_I^2	0.000093687
Additional variation between subgroups	=	s_A^2	0.000064438
Proportion of additional variation between subgroups	=	s_A^2/s_{tot}^2	0.41

H_0 Variation between subgroups is zero

H_1 Variance between subgroups is NOT zero

Test level	critical values lower	critical values upper	Test statistics
α = 5 %	---	1.28	4.43899***
α = 1 %	---	1.42	
α = 0.1 %	---	1.59	
Test results	Null hypothesis rejected at level α ≤ 0,1%		

Abbildung 27: Das Fenster "ANOVA" zeigt das Ergebnis für den Datensatz "test_01.dfq".

Kritische Werte werden berechnet für Signifikanzniveaus $\alpha = 5\ \%,\ 1\ \%$ und $0.1\ \%$. Dafür wird die inversen Verteilungsfunktion (Quantilfunktion) der F-Verteilung mit den folgenden Freiheitsgraden verwendet:

$$F_{crit.\alpha} = F_{1-\alpha, df1,\ df2}$$

$$df1 = p - 1$$

$$df2 = p \times (n - 1)$$

Hierbei sind:

p : Anzahl der Stichproben-Untergruppe

n : Anzahl der Werte in einer einzelnen Stichproben-Untergruppe

Es wird davon ausgegangen, dass alle Stichproben-Untergruppen gleich groß sind (balancierter Versuchsplan).

Referenz:

NIST Handbuch der technischen Statistik (online)

https://www.itl.nist.gov/div898/handbook/ppc/section2/ppc231.htm

1.3.4 Tests auf homogene Streuung

1.3.4.1 Levene

Multifunktionsleiste:
Registerkarte <Ergebnisse> | <Testverfahren> | <Levene>

Dieser Test prüft für ein einziges Merkmal, ob die Stichproben-Untergruppen aus einer Grundgesamtheit mit demselben Streuungsparameter stammen oder nicht. Dieses Verfahren hängt nicht von der Voraussetzung ab, dass die Stichproben aus einer Normalverteilung entnommen werden, sondern aus kontinuierlichen, gleichförmigen Verteilungen.

In einem ersten Schritt wird für jede Stichproben-Untergruppe die *absolute Abweichung vom Mittelwert der Stichprobe* berechnet. Im nächsten Schritt werden den absoluten Abweichungen *Ränge* zugewiesen. Im letzten Schritt wird mit diesen Rängen der Kruskal-Wallis-Test durchgeführt.

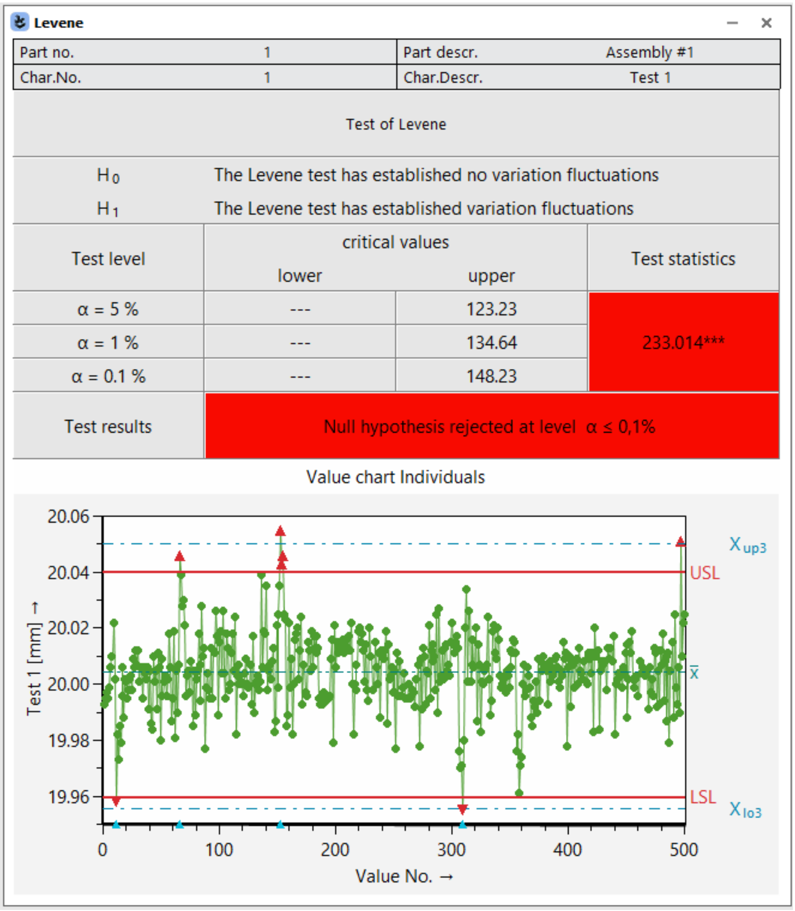


Abbildung 28: Das Fenster "Levene" zeigt das Ergebnis des Datensatzes "test_01.dfq".

Das Programm berechnet die kritische Werte unter Verwendung der inversen Verteilungsfunktion (Quantilfunktion) der Chi²-Verteilung mit den Signifikanzniveaus α=5 %, 1 % und 0.1 %.:

$$H_{crit.\alpha} = \chi^2_{1-\alpha,\,p-1}$$

1.3.4.2 Cochran

Multifunktionsleiste:
Registerkarte <Ergebnisse> | <Testverfahren> | <Cochran>

Der Cochran-Test dient der Überprüfung auf homogene Varianzen. Für ein einzelnes Merkmal wird dabei geprüft, ob alle Stichproben aus einer Grundgesamtheit mit gleicher Varianz stammen. Der Test ist nur im Modul „Prozessanalyse“ verfügbar.

Die Prüfgröße ist das Verhältnis der größten beobachteten Varianz der Stichprobe zur Summe aller Varianzen der Stichprobe:

$$C = \frac{s_{max}^2}{\sum_{i=1}^{p} s_i^2}$$

Dabei ist p die Anzahl der Stichproben-Untergruppen und s_i^2 die Varianz der $i$ten Stichproben-Untergruppe.

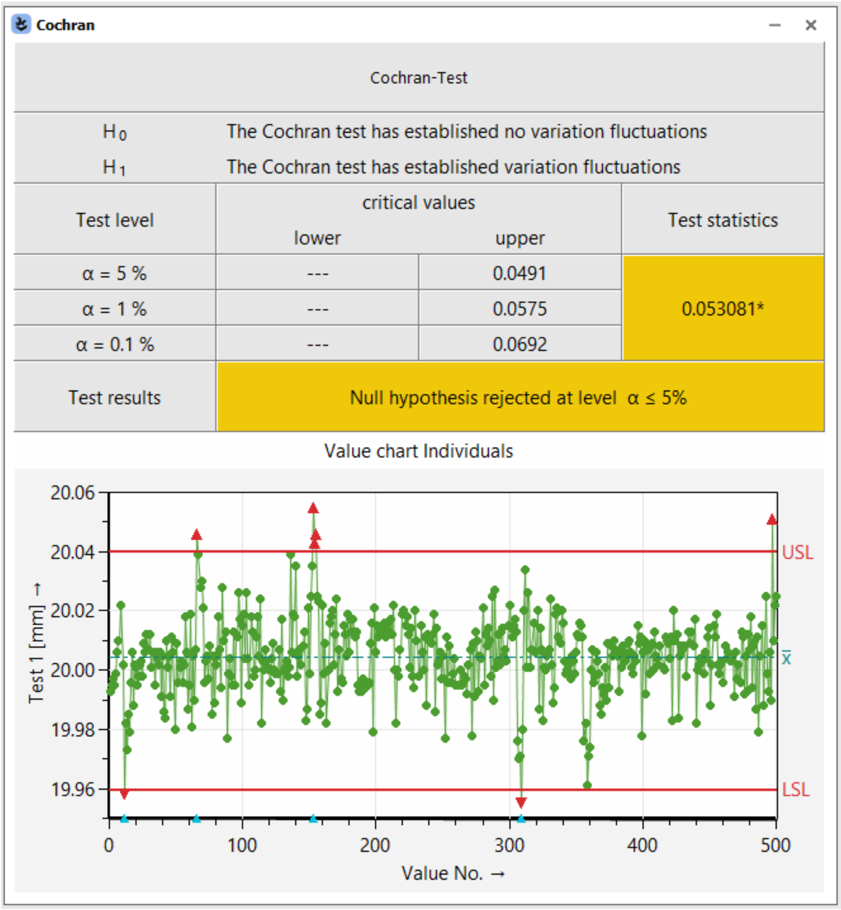


Abbildung 29: Das Fenster "Cochrane" zeigt das Ergebnis für den Datensatz "test_01.dfq".

Die ungefähren kritischen Werte des Cochran-Tests werden wie folgt berechnet:

$$c_{crit} = \frac{F_{df1,\,df2,\;1-\alpha/p}}{F_{df1,\,df2,\;1-\alpha/p} + (p-1)}$$

Dabei ist p die Anzahl der Stichproben-Untergruppen ist und n die Anzahl der Werte in einer einzelnen Stichproben-Untergruppe (nicht der Gesamtumfang der Stichprobe!). Es wird davon ausgegangen, dass alle Stichproben gleich groß sind (balancierter Versuchsplan). Die Freiheitsgrade bestimmt man mit:

$$df1 = n - 1$$

$$df2 = (n-1)(p-1)$$

$F_{df1,\,df2,\;1-\alpha/p}$ ist die inverse Verteilungsfunktion der F-Verteilung (Quantilfunktion)

Referenz:

Ulrich GRAF, Hans-Joachim HENNING, Kurt STANGE, Peter-Theodor WILRICH
Formeln und Tabellen der angewandten mathematischen Statistik
3. Auflage (1998)
Springer Verlag
ISBN 3-540-16901-6
Formel 6.8.17 auf Seite 179

1.3.5 Erweiterter Shapiro-Wilk Test

1.3.5.1 Erweiterter Shapiro-Wilk test für gemeinsam betrachtete Stichproben

Multifunktionsleiste:
Registerkarte <Ergebnisse> | <Testverfahren> | <Erweiterter Shapiro-Wilk>

Dieser Test ist nur im Softwaremodul "Prozessfähigkeit" verfügbar. Es handelt sich um einen Test auf momentane Normalverteilung für gemeinsam betrachtete Stichproben, wobei in unserer Software die "gemeinsam betrachteten Stichproben" die "gemeinsam betrachteten Stichproben-Untergruppen" eines einzelnen Merkmals sind.

In einem ersten Schritt wertet dieses Testverfahren für jede Stichproben-Untergruppe den Shapiro-Wilk-Test aus, um die Prüfgröße W_j für jede der p Stichproben-Untergruppen zu erhalten. Im nächsten Schritt werden die Prüfgrößen W_j in die Variablen G_j transformiert.

$$G_j = \gamma(n) + \delta(n)v_j$$

wobei

$$v_j = ln\left\{\frac{W_j - \varepsilon(n)}{1 - W_j}\right\}$$

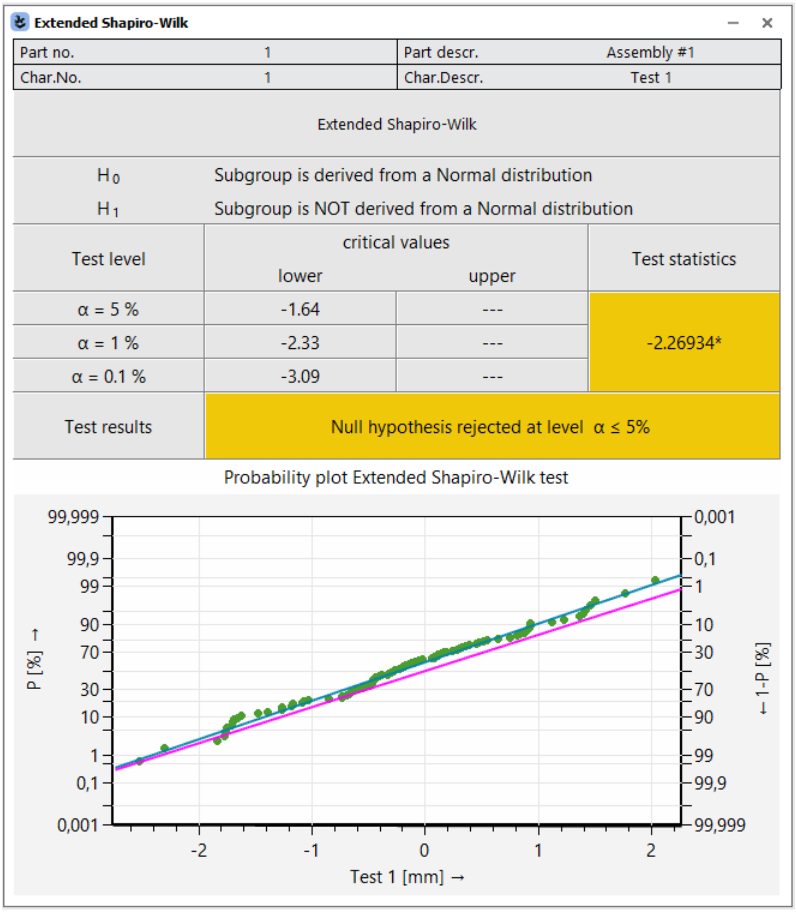
Die Koeffizienten $\gamma(n)$, $\delta(n)$ und $\varepsilon(n)$ für die Umrechnung der W_j in die Variable G_j werden aus einer Tabelle entnommen (z. B. Tabelle 13 auf Seite 31 der ISO 5479: 1997). Falls die den Stichprobendaten zugrunde liegende Verteilung eine Normalverteilung ist, folgt die Variable G_j annähernd der standardisierten Normalverteilung. Der Mittelwert der Variablen G_j ist:

$$\bar{G} = \frac{1}{p}\sum_{j=1}^{p} G_j$$

Schließlich lautet die Prüfgröße des Shapiro-Wilk-Tests für gemeinsam betrachtete Stichproben:

$$SWJ = \sqrt{p} \times \bar{G}$$

wobei p die Anzahl der Stichproben-Untergruppen ist.



Extended Shapiro-Wilk

Part no.	1	Part descr.	Assembly #1
Char.No.	1	Char.Descr.	Test 1

Extended Shapiro-Wilk

H_0	Subgroup is derived from a Normal distribution		
H_1	Subgroup is NOT derived from a Normal distribution		
Test level	critical values lower	critical values upper	Test statistics
α = 5 %	-1.64	---	-2.26934*
α = 1 %	-2.33	---	
α = 0.1 %	-3.09	---	
Test results	Null hypothesis rejected at level α ≤ 5%		

Abbildung 30: Fenster "Erweiterter Shapiro-Wilk" mit den Ergebnissen des Shapiro-Wilk-Tests für gemeinsam betrachtete Stichproben, berechnet für den Datensatz "test_01.dfq". Die Grafik zeigt die transformierten Werte G_j (einer für jede Stichprobe = 100 G-Werte).

Die kritischen Werte (untere Grenze) des Shapiro-Wilk-Tests für gemeinsam betrachtete Stichproben werden unter Verwendung der inversen Verteilungsfunktion (Quantilfunktion) der Standardnormalverteilung für die Signifikanzniveaus α = 5 %, 1 % und 0,1 % berechnet:

$$SWJ_{crit.\alpha} = z_\alpha$$

Referenz:

ISO 5479: 1997, Kapitel 9 (S. 22-23) und Tabelle 13 auf Seite 31

1.3.5.2 Erweiterter Shapiro-Wilk-Test (erweitert)

Multifunktionsleiste:
Registerkarte <Ergebnisse> | <Testverfahren> | <Erweiterter Shapiro-Wilk (W-Netz)>

Öffnet das Grafikfenster, das das Wahrscheinlichkeitsnetz mit den transformierten Prüfgrößenwerten G_j des Shapiro-Wilk-Tests für gemeinsam betrachtete Stichproben-Untergruppen enthält. Wenn die G_j einer Geraden folgen, ist dies ein Indikator dafür, dass die Stichproben-Untergruppen normalverteilt sind.

1.3.6 Tests auf linearen Trend und Zufälligkeit

1.3.6.1 Test auf Differenzenstreuung

Multifunktionsleiste:
Registerkarte <Ergebnisse> | <Testverfahren> | <Differenzenstreuung>

Wurden die Stichprobendaten in zeitlicher Folge erhoben, kann die mittlere quadrierte sukzessive Differenz zur Prüfung auf Zufälligkeit (der Reihenfolge der Daten) verwendet werden.

Die Prüfgröße ist das Verhältnis zwischen der *mittleren quadrierten sukzessiven Differenz* Δ^2 und der (gesamten) Varianz *der Stichprobe* s^2:

$$r = \frac{\Delta^2}{s^2}$$

wobei n die Gesamtstichprobengröße ist. Die *mittlere quadrierte suksessive Differenz* Δ^2 ist definiert als:

$$\Delta^2 = \frac{1}{n-1}\sum_{i=1}^{n-1}(x_{i+1} - x_i)^2$$

Wenn der Gesamtstichprobenumfang größer als $n = 20$ Werte beträgt und die Daten aus einer normalverteilten Grundgesamtheit stammen, dann ist die Größe

$$z = \frac{1 - \frac{r}{2}}{\sqrt{\frac{n-2}{(n-1)(n+1)}}}$$

annähernd standardnormalverteilt ist (diese Beziehung wird zur Berechnung der kritischen Werte verwendet).

Variation of spread

Part no.	1	Part descr.	Assembly #1
Char.No.	1	Char.Descr.	Test 1

Successive spread variation			
H_0	Successive spread variations are uncorrelated		
H_1	Successive spread variations are correlated		
Test level	critical values lower	critical values upper	Test statistics
α = 5 %	1.8532	---	1.40934***
α = 1 %	1.7924	---	
α = 0.1 %	1.7242	---	
Test results	Null hypothesis rejected at level α ≤ 0,1%		

Averages

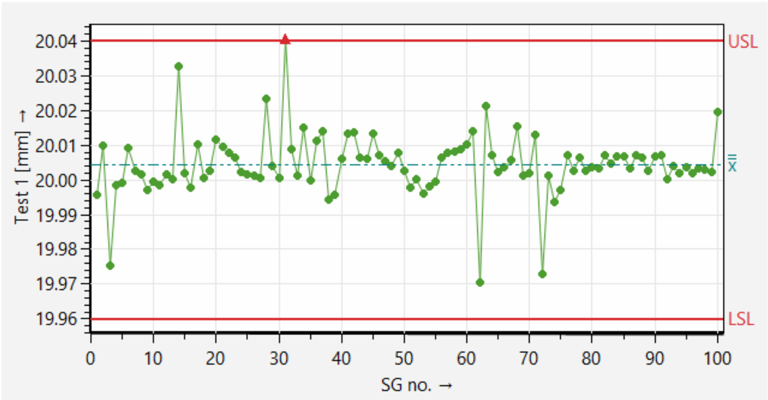


Abbildung 31: Das Fenster "Ausprägung der Streuung" zeigt die Ergebnisse für den Datensatz "test_01.dfq".

Näherungen für die kritische Werte (untere Grenze) werden unter Verwendung des Quantils der Standardnormalverteilung $z_{1-\alpha}$ berechnet (für die Signifikanzniveaus $\alpha = 5\,\%, 1\,\%$ und $0.1\,\%$):

$$r_{crit.\alpha} = 2 - 2 \times z_{1-\alpha} \times \sqrt{\frac{n-2}{(n-1)(n+1)}}$$

Referenz:

Wilfried J. DIXON, Frank J. MASSEY, Jr.
Einführung in die statistische Analyse
Vierte Auflage
McGraw-Hill 1985
ISBN 0-07-017073-8
Chapter 17-12 Test der Zufälligkeit, S. 406-407

1.3.6.2 Swed Eisenhart (Run-Test)

Multifunktionsleiste:
Registerkarte <Ergebnisse> | <Testverfahren> | <Swed Eisenhart>

Bei diesem Verfahren handelt es sich um einen Test auf die Zufälligkeit der Reihenfolge der Stichproben-Einzelwerte.

Swed & Eisenhart

Part no.	1	Part descr.	Assembly #1
Char.No.	1	Char.Descr.	Test 1

Swed & Eisenhart

H_0	Random data sequence		
H_1	Non random data sequence		
Test level	critical values lower	critical values upper	Test statistics
α = 5 %	---	1.960	4.36865***
α = 1 %	---	2.576	
α = 0.1 %	---	3.291	
Test results	Null hypothesis rejected at level α ≤ 0,1%		

Value chart Individuals

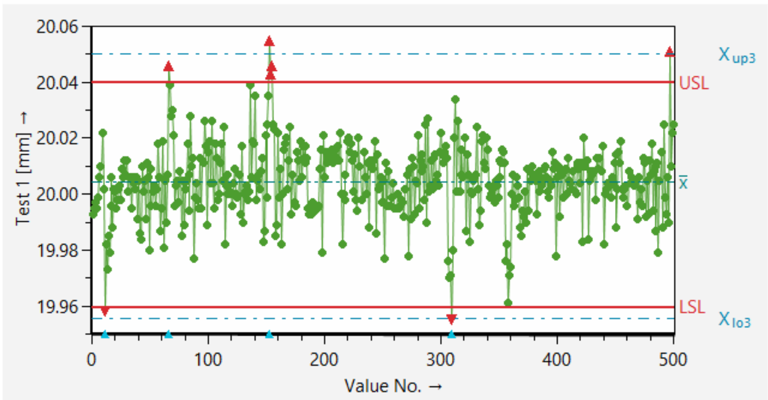


Abbildung 32: Das Fenster "Swed & Eisenhart" " mit dem Ergebnis des Tests für den Datensatz "Test_01.dfq"

Der Weg zur Prüfgröße ist wie folgt:

1) Bestimme den Median aller Werte (Gesamte Daten)
2) Zählen der Anzahl der Einzelwerte, die größer als der Median sind: $n1$.
3) Zählen der Anzahl der Einzelwerte, die kleiner oder gleich dem Median sind: $n2$.
4) Ein *Run* ist eine ununterbrochene Folge von Einzelwerten, die größer als der Median sind, oder eine ununterbrochene Folge von Einzelwerten, die kleiner oder gleich dem Median sind.
   Die Gesamtzahl dieser beobachteten *Runs* wird ausgezählt: r_{obs}.
5) Berechnung der zu erwarteten Anzahl der *Runs* und die damit verbundene Varianz:
   $$r_{exp} = \frac{2 \times n_1 \times n_2}{n_1 + n_2} + 1$$
   $$s_{r.exp}^2 = \frac{2 \times n_1 \times n_2 \times (2 \times n_1 \times n_2 - n_1 - n_2)}{(n_1 + n_2)^2 \times (n_1 + n_2 - 1)}$$
6) Berechnung der Prüfgröße (absoluter Wert):
   $$|z_r| = \left| \frac{r_{obs} - r_{exp}}{\sqrt{s_{r.exp}^2}} \right|$$

Die kritischen Werte werden mit Hilfe der umgekehrten Verteilungsfunktion (Quantilfunktion) der gespeicherten Normalverteilung berechnet; für die Signifikanzniveaus $\alpha = 5\,\%, 1\,\%$ und $0.1\,\%$.

$$z_{crit.\alpha} = z_{1-\frac{\alpha}{2}}$$

Referenz:

NIST Handbuch für technische Statistiken (online)
https://www.itl.nist.gov/div898/handbook/eda/section3/eda35d.htm

1.3.6.3 Linearer Trend

Multifunktionsleiste:
Registerkarte <Ergebnisse> | <Testverfahren> | <Linearen Trend>

Dieser Test kann ggf. aussagekräftige Ergebnisse berechnen, wenn das Verlaufsmuster des vorliegenden Datensatzes in der Grafik „Werteverlauf Einzelwerte“ insgesamte einen einzigen Aufwärtstrend oder aber einen einzigen Abwärtstrend bildet.

Die Prüfgröße des Tests auf linearen Trend wird wie folgt berechnet:

1) Berechnung einer Gewichtung für jede Stichproben-Untergruppe:

$$w_i = \frac{\sqrt{n_i}}{s_i^2 \times a_n}; i = 1, 2, \ldots, m$$

wobei
n_i: Der Umfang der $i$ten Stichproben-Untergruppe.
s_i^2: Die Varianz der $i$ten Stichproben-Untergruppe.
m: Die Anzahl der Stichproben-Untergruppen.

$$a_n = \sqrt{\frac{2}{n_i - 1}} \times \frac{\Gamma\left(\frac{n_i}{2}\right)}{\Gamma\left(\frac{n_i - 1}{2}\right)}$$

2) Berechnung der gewichteten Summen:

$$S1 = \sum_{i=1}^{m} w_i$$

$$S_x = \sum_{i=1}^{m} (i-1) \times w_i$$

$$S_{xx} = \sum_{i=1}^{m} (i-1)^2 \times w_i$$

$$S_y = \sum_{i=1}^{m} (\bar{x}_i - \bar{x}_1) \times w_i$$

$$S_{yy} = \sum_{i=1}^{m} (\bar{x}_i - \bar{x}_1)^2 \times w_i$$

$$S_{xy} = \sum_{i=1}^{m} (i-1) \times (\bar{x}_i - \bar{x}_1) \times w_i$$

hierbei ist $\bar{x}_i$ der Mittelwert der $i$ten Stichproben-Untergruppe

3) Auf der Grundlage der gewichteten Summen werden zusätzliche Variablen berechnet:

$$D_x = S1 \times S_{xx} - s_x^2$$
$$D_y = S1 \times S_{yy} - s_y^2$$
$$p_b = \frac{S1 \times S_{xy} - S_x \times S_y}{D_x}$$
$$p_{db} = \sqrt{\frac{S1}{D_x}}$$
$$p_a = \frac{S_y - S_x \times p_b}{S1} + \bar{x}_1 - p_b$$
$$p_{da} = \sqrt{\frac{S_{xx} + 2 \times S_x + S1}{D_x}}$$
$$p_r = p_b \times \sqrt{\frac{D_x}{D_y}}$$
$$test_{statistic} = (1 - p_r)^2 \times \frac{D_y}{S1}$$

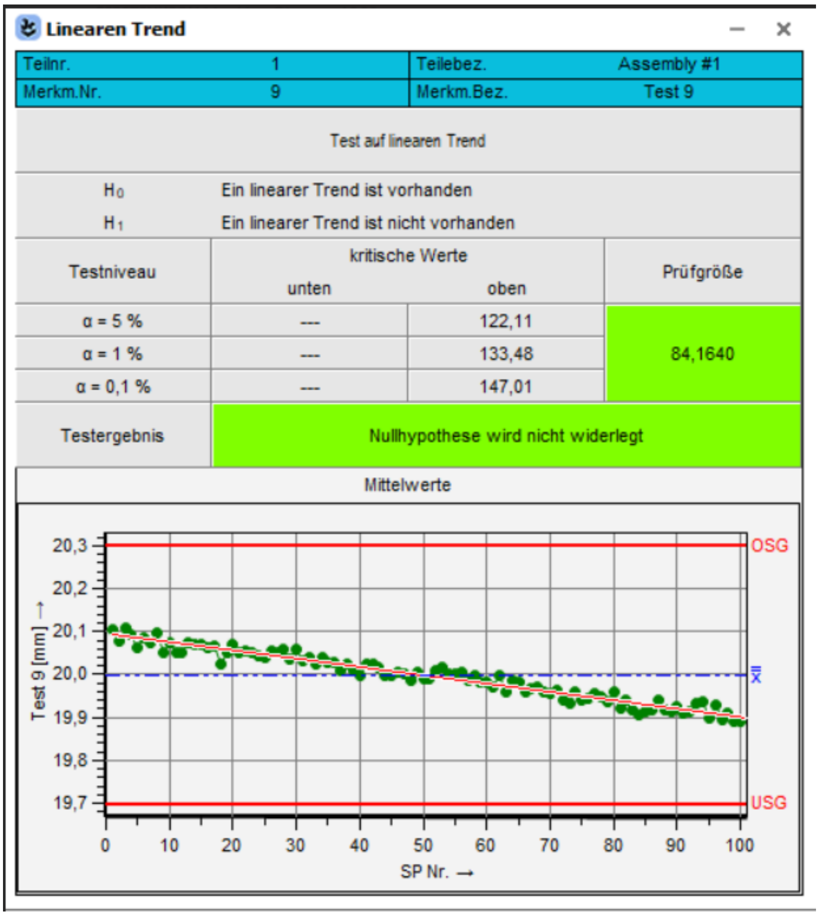


Abbildung 33: Fenster "Linearer Trend" mit dem Ergebnis des Tests auf linearen Trend für den Datensatz "Test_09.dfq"

Kritische Werte werden aus dem (1-α) Quantil der Chi²-Verteilung für die Signifikanzniveaus berechnet $\alpha = 5\,\%, 1\,\%$ und $0.1\,\%$ und $df = m - 2$ Freiheitsgraden, wobei m die Anzahl der Stichproben-Untergruppen ist.

Hinweis:

Das Ergebnis des Tests auf linearen Trend wird ignoriert, wenn gilt: $p_{db} \times G^{-1}(0.95) > |p_b|$

Hierbei ist $G^{-1}()$ die Quantilfunktion der Standardisierten Normalverteilung. In diesem Fall ist die Steigung zu klein und die Testentscheidung wird per Default auf H1 gesetzt (Dabei wird das H0-Ergebnis des Tests auf linearen Trend ignoriert).

1.3.6.4 Abschnittweisen linearen Trend

Multifunktionsleiste:
Registerkarte <Ergebnisse> | <Testverfahren> | <Abschnittsweisen linearen Trend>

Das Verfahren ist komplex und gliedert sich in mehrere Teilschritte, wie nachfolgend beschrieben.

1.3.6.4.1 Verfahren für den Test auf abschnittsweisen Trend nach VW 101 31: 2005-02

Häufig zeigen Daten mit abschnittsweisen linearen Trends ein sägezahnartiges Muster im Wertestrahl der Einzelwerte, bedingt durch Werkzeugwechsel oder Werkzeugkorrekturen. Wir verwenden für dieses Kapitel die *alte und zurückgezogene* Auswertungsstrategie "VW-Konzern 10131 (06/2012)" der Volkswagen AG, da diese Auswertungsstrategie Optionseinstellungen nach dem internen Standard "VW 10130: 2005" der Volkswagen AG, S. 31 - 34, verwendet.

Schritt 1: Erkennung von Sprung-Diskontinuitäten
Anhand des folgenden simulierten Datenbeispiels soll das gesamte Verfahren erläutert werden:

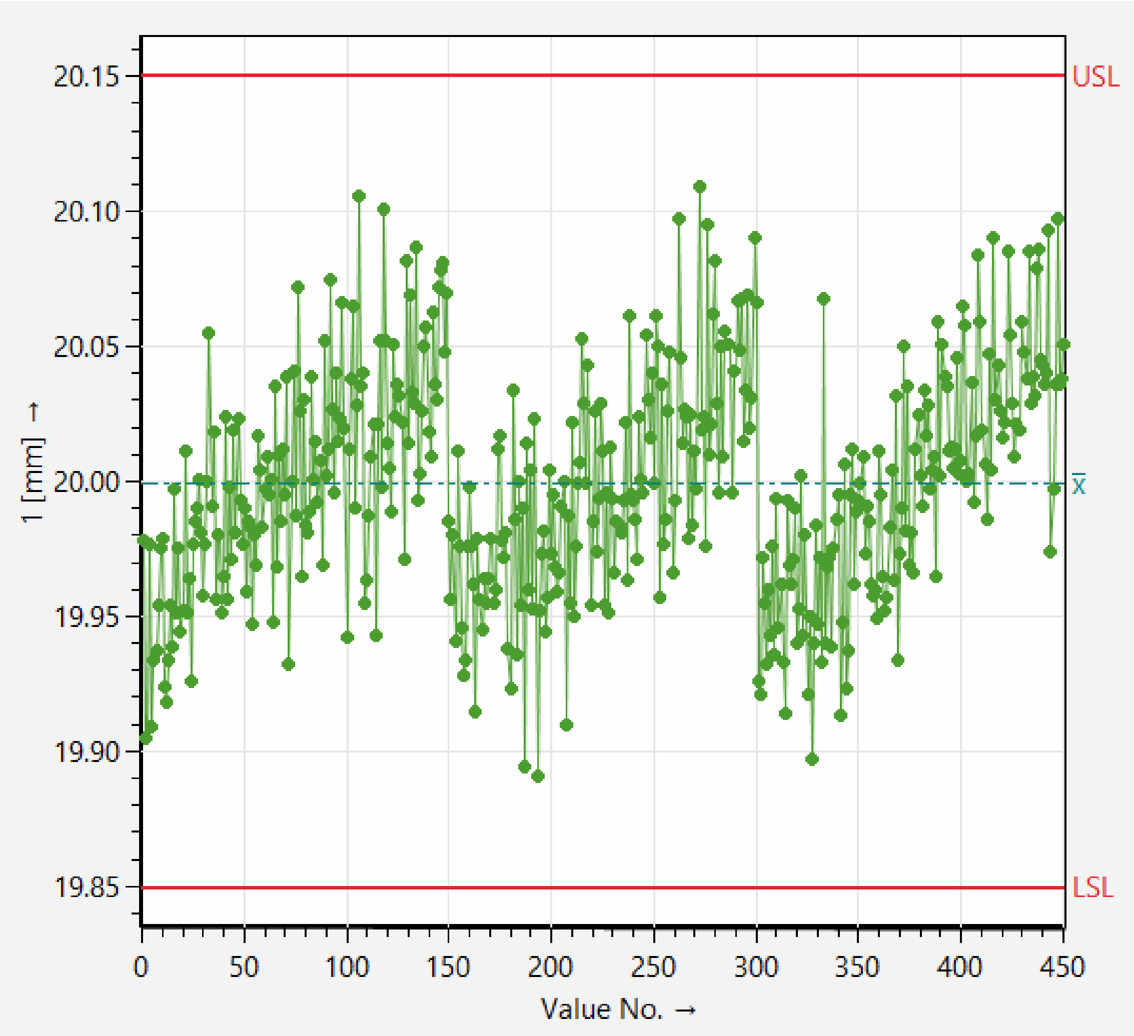


Abbildung 34: Grafik eines simulierten Datensatzes mit drei Datenabschnitten mit linearem Trend.

Die Abbildung 34 zeigt einen Datensatz, der ein sägezahnähnliches Muster aufweist. Für diese Art von Daten ist es zunächst erforderlich, die Stellen der Sprung-Diskontinuitäten zu identifizieren. Dann ist es möglich, eine Regressionsgerade an jeden einzelnen Abschnitt anzupassen.
In der Auswertungsstrategie werden die folgenden Optionseinstellungen für die Erkennung von Sprung-Diskontinuitäten verwendet:

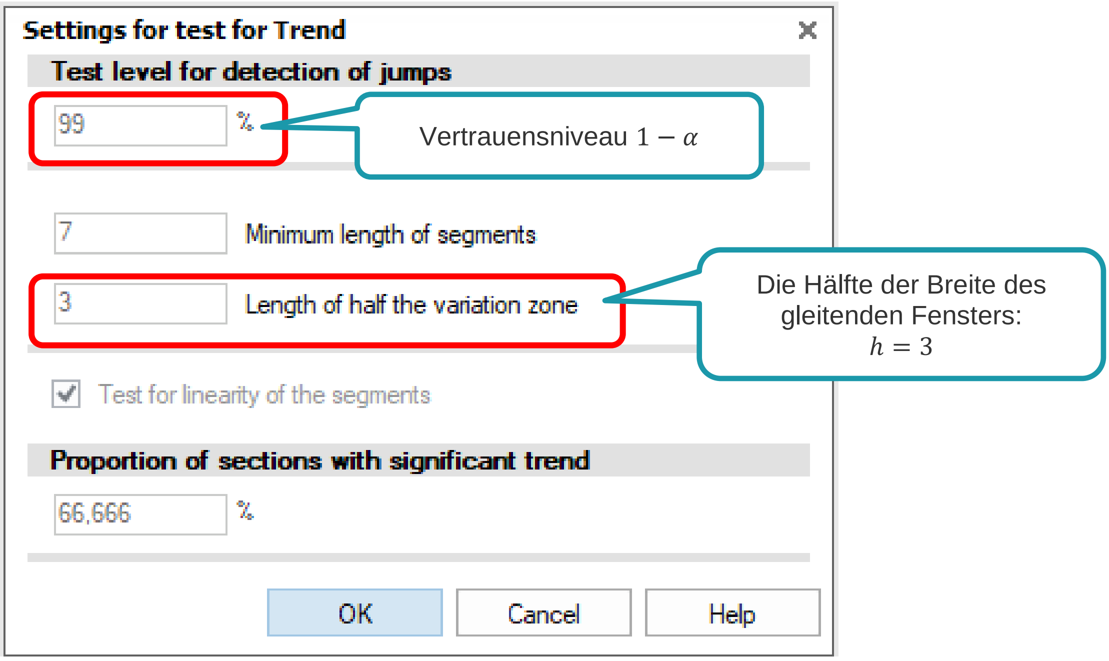


Abbildung 35: Auszug aus der Auswertungsstrategie (Symbol "Test auf Trend") mit den Optionseinstellungen des Tests auf Sprung-Diskontinuitäten.

Testhypothesen für den Test auf Sprung-Diskontinuitäten:

- **Null-Hypothese H0**: Die Lage von µ ändert sich nicht oder die Änderung ist kontinuierlich.
- **Alternativhypothese H1**: Die Lage von µ ändert sich bei einem Sprung.

1) Berechnung des *Mittelwertes der Stichprobe* für jede Stichprobe.
2) Berechnung einer *gleitenden Statistik für* den Test $(\Delta\bar{x}_j)^2$ auf der Grundlage eines *gleitenden Fensters mit der* Breite $2\times h=6$:

   $(\Delta\bar{x}_j)^2=(\bar{x}_{j+2}+\bar{x}_{j+1}+\bar{x}_j-\bar{x}_{j-1}-\bar{x}_{j-2}-\bar{x}_{j-3})^2;\ j=3,\ 4,\ \ldots,\ m-2$

   wobei m die Gesamtzahl der Stichproben ist und die $\bar{x}_{jth}$ Stichproben-Mittelwerte sind.
3) Berechnung der Prüfgrößen des Tests E:

   $$E=\chi^2_{99\,\%,1}\cdot\frac{6}{n}\cdot\hat{\sigma}_m^2+\frac{27}{4}\cdot\hat{\sigma}_R^2$$

   wobei

   $$\hat{\sigma}_R^2=\hat{\sigma}^2-\hat{\sigma}_m^2$$

   $$\hat{\sigma}=s=\sqrt{\frac{1}{m-1}\cdot\sum_{j=1}^{m}n_j\cdot(\bar{x}_j-\bar{\bar{x}})^2}$$

   $$\hat{\sigma}_m=s_m=\sqrt{\frac{1}{m}\cdot\sum_{j=1}^{m}s_j^2}$$

   $\chi^2_{99\,\%,\,1}$ das $(1-\alpha=)$ 99 %-Quantil der Chi²-Verteilung mit einem Freiheitsgrad.

4) Für jede der gleitend berechneten Prüfgrößen ist der folgende Test durchzuführen:
$wenn\ (\Delta\bar{x}_j)^2 > E, dann\ lehne\ H0\ ab$

5) Wurde H0 verworfen, wird der Index der Stichprobe J mit der Sprungstelle wie folgt ausgewertet:

$$J = \begin{cases} j+2 & if\ (\Delta\bar{x}_{j+2})^2 = max\left\{(\Delta\bar{x}_{j+2})^2, (\Delta\bar{x}_{j+1})^2, (\Delta\bar{x}_j)^2\right\} \\ j+1 & if\ (\Delta\bar{x}_{j+1})^2 = max\left\{(\Delta\bar{x}_{j+2})^2, (\Delta\bar{x}_{j+1})^2, (\Delta\bar{x}_j)^2\right\} \\ j & if\ (\Delta\bar{x}_j)^2 = max\left\{(\Delta\bar{x}_{j+2})^2, (\Delta\bar{x}_{j+1})^2, (\Delta\bar{x}_j)^2\right\} \end{cases}$$

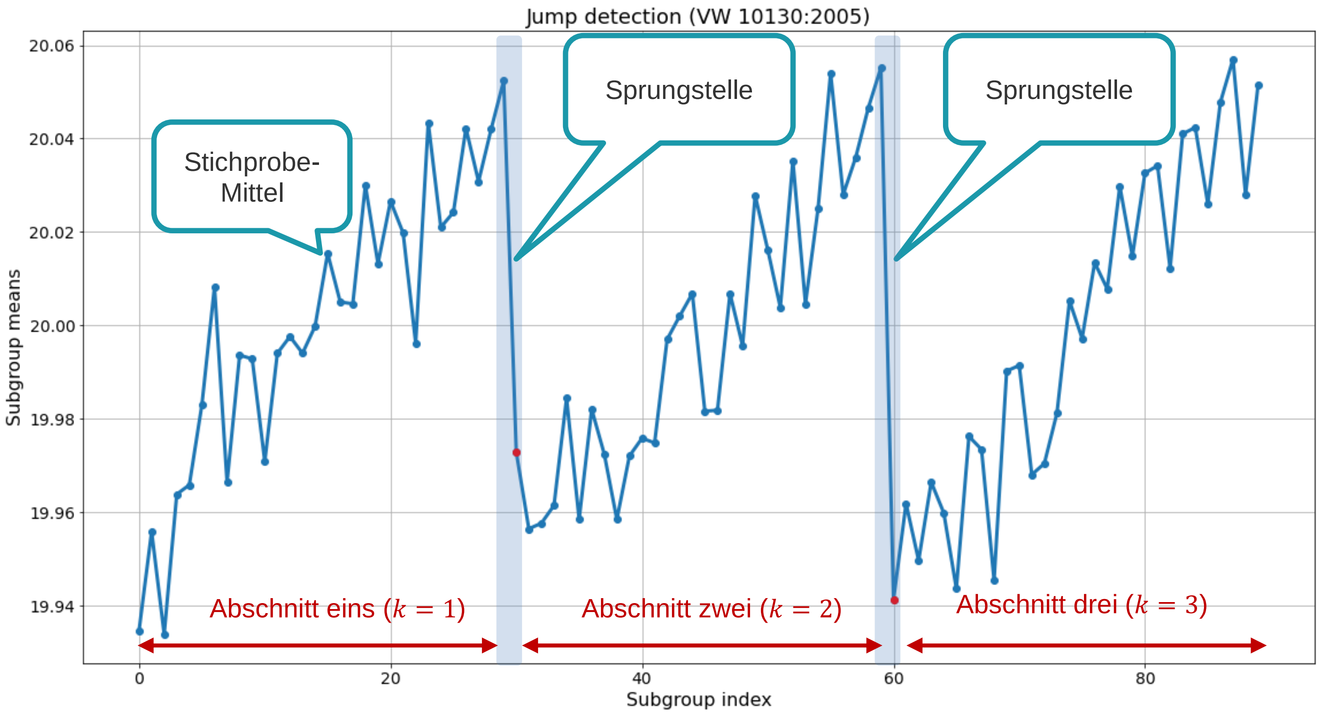


Abbildung 36: Grafik eines simulierten Test-Datensatzes mit drei Datenabschnitten mit linearem Trend.

Schritt 2: Anpassung eines Regressionsmodells an jeden Abschnitt (Geradenanpassung)

1) Wenn ein Abschnitt eine Mindestlänge[3] von $m_{min} = 7$ (dies ist der Wert der Option "Mindestlänge der Abschnitte" in Abbildung 35), wird eine Regressionsgerade angepasst (Methode der kleinsten Quadrate)
$\hat{b}_0 = \bar{\bar{x}} - \hat{b}_1 \bar{J}_i$
$\hat{b}_1 = \frac{s_{\bar{x}j}}{s_j^2}$
wobei
$s_{\bar{x}j} = \frac{1}{m_k}\left(\sum_{j=j_{lo}}^{j_{up}} j \times \bar{x}_j - m_k \times \bar{J}_k \times \bar{\bar{x}}_k\right)$ Kovarianz
$s_j^2 = \frac{1}{m_k - 1}\left(\sum_{j=j_{lo}}^{j_{up}} j^2 - m_k \times \bar{J}_k^2\right)$ Varianz der Stichproben-Indizes
$\bar{J}_k$: Der Mittelwert der Stichproben-Indizes (des $k$ten Datenabschnitts)
$\bar{\bar{x}}_k$: Der Mittelwert der Stichproben-Mittelwerte (des $k$ten Datenabschnitts)

[3] Liegt ein Abschnitt zwischen zwei Sprungpunkten, werden die erste Stichprobe nach dem linken Sprungpunkt und die letzte Stichprobe vor dem rechten Sprungpunkt bei der Ermittlung der Länge nicht mitgezählt.

m_k: Die Anzahl der Stichproben-Mittelwerte (des $k$ten Datenabschnitts)

Schritt 3: Einen Test auf signifikante Steigung für jede Regressionsgerade durchführen.
Der t-Test für die Steigung prüft, ob die Steigung einer Regressionsgerade signifikant von Null verschieden ist:

- **Null-Hypothese H0**: Die Steigung ist gleich Null ($\beta_1 = 0$)
- **Alternativ-Hypothese H1**: Die Steigung ist ungleich Null ($\beta_1 \neq 0$)

Die Prüfgröße t_{slope} ist die Schätzung der Steigung geteilt durch den Standardfehler der Steigung:

$$t_{slope} = \frac{\hat{\beta}_1}{se_{\beta_1}}$$

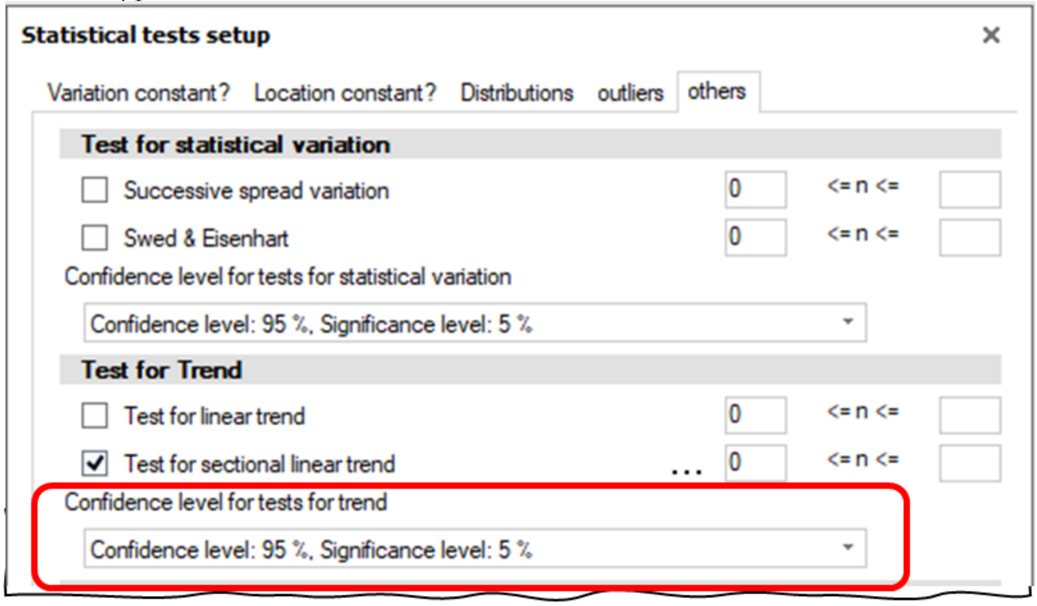


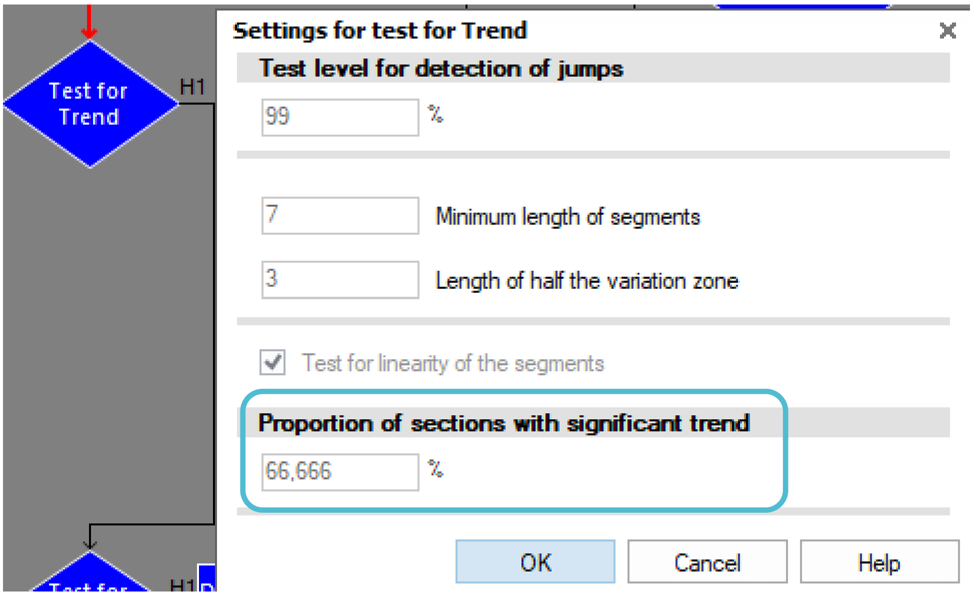
Abbildung 37: Auszug aus der Auswertungsstrategie (Symbol "Test auf Trend"), der die Option für das Signifikanzniveau des Tests auf Trend zeigt.

Ist die folgende Bedingung erfüllt, so ist die Null-Hypothese zu verwerfen:
$t_{slope} > t_{95\,\%;\,m_k-2}$
wobei
$t_{95\,\%;\,m_k-2}$ 95 %-Quantil der t-Verteilung: 95 %-Quantil der t-Verteilung
$m_k - 2$ Anzahl von Freiheitsgraden
m_k Anzahl der Stichproben-Mittelwerte in der kAbschnitt

Hinweis 1: Es gibt keine Ausgabepunkte für die Ergebnisse des Sprungstellen-Ermittlungsverfahrens oder der berechneten Regressionskoeffizienten und auch nicht für die Ergebnisse des Tests auf eine signifikante Steigung.

Schritt 4: Den Anteil der Datenabschnitte mit einem linearen Trend zählen

Wenn eine Auswertungsstrategie verwendet wird, die für jedes der zeitabhängigen Verteilungsmodelle C1, C2, C3 und C4 (gemäß ISO 22514-2) einen anderen Auswertungspfad vorsieht, wird dieser "Test" verwendet, um zwischen einem C1-Modell und einem C3-Modell unterscheiden zu können. Wenn wir die alte (und zurückgezogene) Auswertungsstrategie "*VW-Konzern 10131 (06/2012)*" öffnen, finden wir im Fenster "*Einstellungen für* Test auf Trend" die Option: "*Anteil der Abschnitte mit signifikantem Trend*". Diese Option ist auf 2/3 (genauer: 66,666 %) eingestellt.



Wenn ein Datensatz Sprungstellen aufweist (➔ Siehe "Test auf Trend"), werden die Daten auf beiden Seiten der Sprungstelle als separate Daten-"abschnitte" betrachtet.

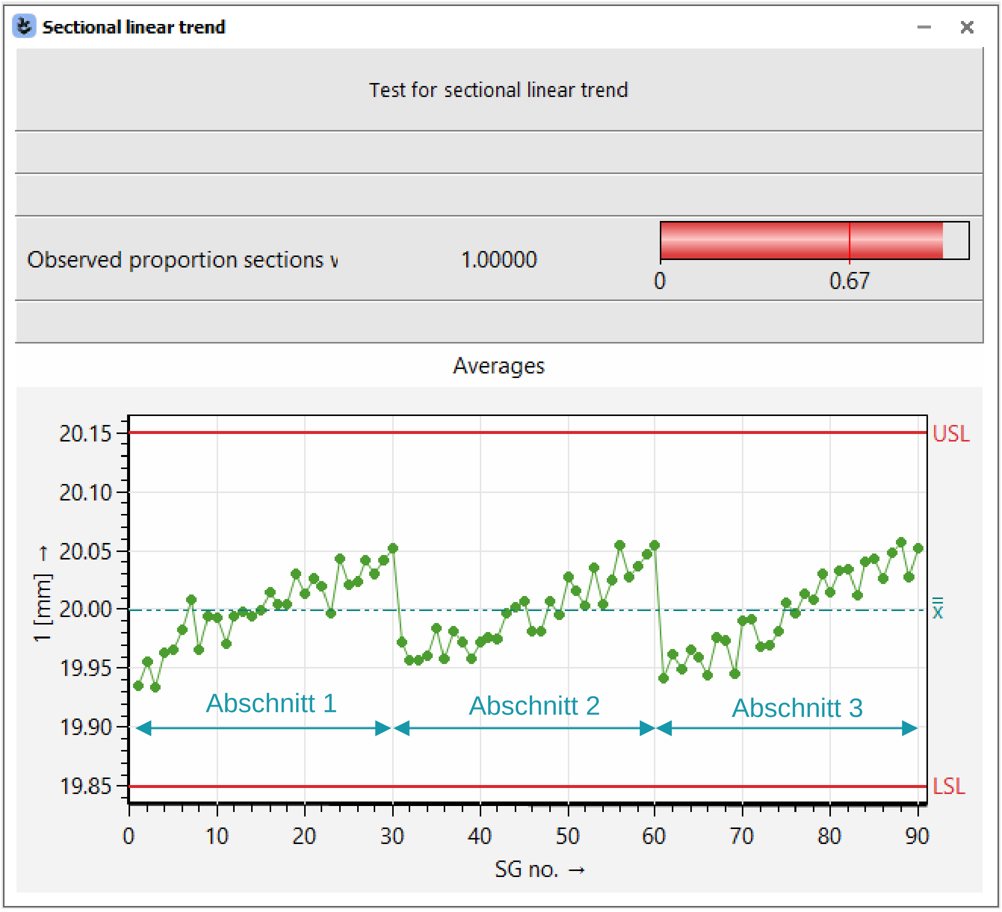


Abbildung 38: Fenster "Abschnittweisen linearen Trend" mit Ergebnissen für einen simulierten Datensatz mit drei Abschnitten mit linearem Trend (Die Grafik zeigt Stichprobenmittelwerte, keine Einzelwerte!).

Das Programm zählt die Gesamtzahl der Abschnitte (in Abbildung 38: drei Abschnitte) und es zählt die Anzahl der Abschnitte mit einem linearen Trend (in Abbildung 38: ebenfalls drei Abschnitte).

Als Nächstes berechnet das Programm das Verhältnis, das in folgender Abbildung dargestellt ist Abbildung 38 (der Wert 1,00000):

$$p = \frac{number\ of\ sections\ with\ a\ linear\ trend}{total\ number\ of\ sections}$$

Es war ein Vorschlag der Volkswagen AG, die heuristische Regel anzuwenden, dass ein C3-Modell einen Mindestanteil $p_{min.C3} = 2/3$ (➔ 66,666%) von Abschnitten mit linearem Trend erfordert. Mit anderen Worten:

- wenn der Anteil p gleich oder größer ist als $2/3$ ist, dann führt der Strategiepfad zum zeitabhängigen Verteilungsmodell C3.
- Wenn der Anteil p kleiner ist als $2/3$ ist, dann führt der Strategiepfad zum zeitabhängigen Verteilungsmodell C1.

1.3.7 Zwei-Stichproben-Tests auf gleiche Varianz und gleiche Lage

1.3.7.1 F-Test

Multifunktionsleiste:
Registerkarte <Ergebnisse> | <Testverfahren> | <F-Test>

Um diesen Test durchzuführen, müssen unsere Daten *mindestens zwei Merkmale* aufweisen: Der F-Test für zwei Stichproben prüft für *zwei ausgewählte Merkmale*, ob die Varianz der beiden Merkmale signifikant unterschiedlich ist (oder nicht). Der Test geht von zwei Stichproben aus, die aus normal verteilten Grundgesamtheiten stammen.

Die Prüfgröße ist das Verhältnis der Varianzen der beiden Stichproben (Gesamtstichproben):

$$F = \frac{s_1^2}{s_2^2}$$

F-T

F test

Part no.	MW			Part no.	MW		
Part descr.	minearl water			Part descr.	minearl water		
Char.No.	1			Char.No.	2		
Char.Descr.	Bottle filling volume of filler 1			Char.Descr.	Bottle filling volume of filler 2		
Distribution	Normal Distribution			Distribution	Normal Distribution		
n_{eff}	35	$\bar{x}$	248.2169	n_{eff}	35	$\bar{x}$	250.7629
s^2	1.012	s	1.006	s^2	0.924	s	0.961

H_0	Variances of the total populations are equal
H_1	Variances of the total populations are NOT equal

Test level	critical values lower	critical values upper	Test statistics
α = 5 %	0.50	1.98	1.09527
α = 1 %	0.40	2.47	
α = 0.1 %	0.31	3.21	
Test results	Null hypothesis not rejected		

Abbildung 39: Das Fenster "F-Test" zeigt die Ergebnisse des F-Tests (Test auf gleiche Varianzen) zum Vergleich der Varianz des Füllvolumens, das von zwei verschiedenen Füllern auf derselben Maschine erzeugt wurde.

Wenn ein Datensatz mehr als zwei Merkmale aufweist, kann durch einen Klick in der Multifunktionsleiste auf die Registerkarte <Teil / Merkmal> ein bestimmtes Merkmal ausgewählt werden (das Fenster "F-Test" muss den Fokus haben):

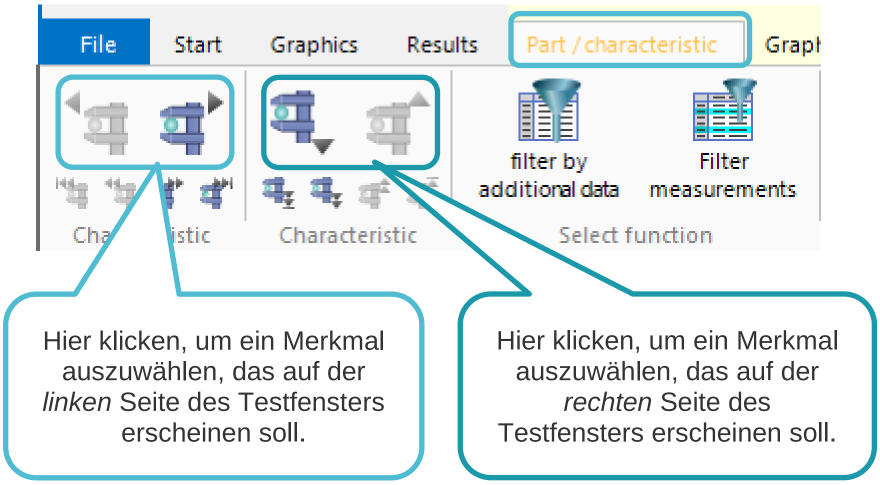


Abbildung 40: Selektion der Merkmale für den F-Test mit zwei Stichproben (und den t-Test mit zwei Stichproben)

Die kritischen Werte werden anhand der umgekehrten Verteilungsfunktion (Quantilfunktion) der F-Verteilung für die Signifikanzniveaus berechnet $\alpha = 5\ \%, 1\ \%$ und $0.1\ \%$:

$$F_{crit.lower} = F_{\frac{\alpha}{2};df1,df2}$$

$$F_{crit.upper} = F_{1-\frac{\alpha}{2};df1,df2}$$

$$df1 = n_1 - 1$$

$$df2 = n_2 - 1$$

wobei n_1 der Gesamtstichprobenumfang des Merkmals auf der linken Seite des Testfensters ist und n_2 der Gesamtstichprobenumfang des Merkmals auf der rechten Seite des Testfensters ist.

1.3.7.2 t-Test

Multifunktionsleiste:
Registerkarte <Ergebnisse> | <Testverfahren> | <t-Test>

Um den t-Test mit zwei Stichproben durchzuführen, muss unser Datensatz mindestens *zwei Merkmale* aufweisen. Bei diesem Verfahren wird für die beiden ausgewählten Merkmale geprüft, ob ein signifikanter Unterschied zwischen den beiden Parametern µ besteht oder nicht. Der Test geht von zwei Stichproben aus, die aus normal verteilten Grundgesamtheiten stammen.

Die Prüfgröße t ist wie folgt definiert:

$$t = \frac{|\bar{x}_1 - \bar{x}_2|}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$s_p = \sqrt{\frac{(n_2 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

wobei

$\bar{x}_1$: Mittelwert der ersten Stichprobe

$\bar{x}_2$: Mittelwert der zweiten Stichprobe

n_1: Gesamtstichprobenumfang der ersten Stichprobe

n_2: Gesamtstichprobenumfang der ersten Stichprobe

s_p: Standardabweichung: gepoolte Standardabweichung

t-test

t-test							
Part no.	MWB250			Part no.	MWB250		
Part descr.	Mineral water bottle 250 ml			Part descr.	Mineral water bottle 250 ml		
Char.No.	Filler 1			Char.No.	Filler 2		
Char.Descr.	Filling volume (Filler 1)			Char.Descr.	Filling volume (Filler 2)		
Distribution	Normal Distribution			Distribution	Normal Distribution		
n_{eff}	35	$\bar{x}$	250.00620	n_{eff}	35	$\bar{x}$	251.97106
s^2	0.0123	s	0.111	s^2	0.0109	s	0.105
H_0	Average values of the total populations are equal						
H_1	Average values of the total populations are NOT equal						
Test level	critical values lower		critical values upper		Test statistics		
α = 5 %	---		2.00		76.2834***		
α = 1 %	---		2.65				
α = 0.1 %	---		3.44				
Test results	Null hypothesis rejected at level α ≤ 0,1%						

Abbildung 41: Das Fenster "t-test" zeigt ein Ergebnis eines t-Tests mit zwei Stichproben.

Kritische Werte werden für die Signifikanzniveaus $\alpha = 5\,\%, 1\,\%$ und $0{,}1\,\%$ unter Verwendung des $\left(1 - \frac{\alpha}{2}\right)$ Quantil der t-Verteilung für $df = (n_1 + n_2 - 2)$ Freiheitsgraden berechnet:

$$t_{crit.\alpha} = t_{1-\frac{\alpha}{2};n_1+n_2-2}$$

1.3.7.3 Assistent (Testverfahren)

Multifunktionsleiste:
Registerkarte <Ergebnisse> | <Testverfahren> | <Assistent (Planen/Durchführen)>

Aufgrund der Komplexität des Assistenten ist die Beschreibung der zugehörigen Testverfahren separat dokumentiert.