Binomial-Verteilung und Hypothesentest

Voraussetzung und verwandte Themen

Für das Verständnis dieser Beschreibung sind Grundlagen der Statistik erforderlich, insbesondere über die Normalverteilung und über Hypothesentests:

www.versuchsmethoden.de/Normalverteilung.pdf

www.versuchsmethoden.de/Poisson-Verteilung.pdf

www.versuchsmethoden.de/Hypothesentests.pdf

Stichworte: Binomial-Verteilung, diskrete Verteilung, Fehleranteil, Hypothesentest

Einführung

Allgemein beschreibt die Binomial-Verteilung bei unabhängigen Versuchen die Wahrscheinlichkeit erfolgreich oder nicht erfolgreich zu sein. Sie ist also eine diskrete Verteilung mit zwei möglichen Ergebnissen. Die Ausgangssituation für den folgenden Versuch ist unverändert, man sagt es handelt sich um „Ziehen mit Zurücklegen“.

Ziel und Nutzen

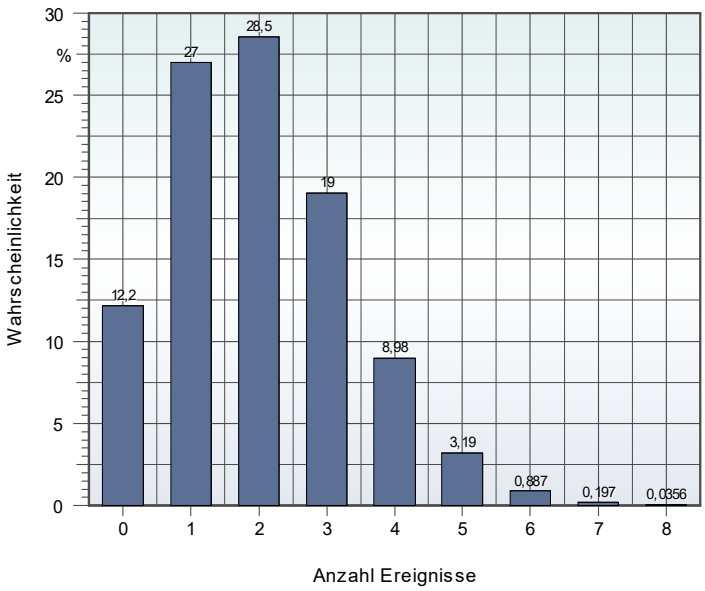
Für technische Fragestellungen wird mit der Binomial-Verteilung z.B. die Anzahl fehlerhafter Einheiten in einer Stichprobe beschrieben. Über den Binomialansatz kann dann eine Aussagewahrscheinlichkeit über die Zuverlässigkeit von Bauteilen gemacht werden.

Grundlagen

Da die Binomial-Verteilung eine diskrete Verteilung ist, muss die Variable x ganzzahlig sein. Die Wahrscheinlichkeitsdichte lautet:

$$h = \binom{n}{x} p^x (1-p)^{n-x} = \frac{n!}{x!(n-x)!}\, p^x (1-p)^{n-x}$$

x : Variable (0, 1, 2 …)
n : Umfang
p : relativer Anteil fehlerhafter Teile



Beispiel für
n = 20
p = 0,1

© 2021 CRGRAPH

Für das gezeigte Beispiel lässt sich aussagen, dass die Wahrscheinlichkeit, dass kein Fehler auftritt = 12,2% ist (erster Balken). Die Wahrscheinlichkeit für einen Fehler ist 27% (zweiter Balken). Die Wahrscheinlichkeit, dass mindestens ein Fehler oder mehr auftreten, ist die Summe aller Wahrscheinlichkeiten für 1 .. ∞ Fehler = 27% + 28,5% + 19% + 8,98% + ….. = 100 – 12,2% = 87,8%.

Vertrauensbereiche

Mit einem Vertrauensbereich drückt man aus, in welchem Bereich sich künftig die Wahrscheinlichkeit p der Ereignisse befinden wird. Hierfür wird normalerweise ein Vertrauensbereich von VB = 90% festgelegt. In den folgenden Formeln wird dieser Vertrauensbereich über die sogenannte Irrtumswahrscheinlichkeit α = 1 - VB definiert.

Der zweiseitige Vertrauensbereich für $p_o = \frac{n_{fehlerhaft}}{n_{geprüft}}$ ist mit Hilfe der Fisher-Verteilung:

$$p_{un} = \frac{(x+1)F_{f1,f2,1-\alpha/2}}{n-x+(x+1)F_{f1,f2,1-\alpha/2}} = \frac{2p_o n + u^2_{1-\alpha/2} - u_{1-\alpha/2}\sqrt{u^2_{1-\alpha/2} + 4p_o n(1-p_o)}}{2(n+u^2_{1-\alpha/2})}$$

mit $f1 = 2(x+1)$ und $f2 = 2(n-x)$

$$p_{ob} = \frac{x}{x+(n-x+1)F_{f1,f2,1-\alpha/2}} = \frac{2p_o n + u^2_{1-\alpha/2} + u_{1-\alpha/2}\sqrt{u^2_{1-\alpha/2} + 4p_o n(1-p_o)}}{2(n+u^2_{1-\alpha/2})}$$

mit $f1 = 2(n-x+1)$ und $f2 = 2x$

Der einseitig nach oben abgegrenzte Vertrauensbereich ist:

$$p_{ob} = \frac{(x+1)F_{f1,f2,1-\alpha}}{n-x+(x+1)F_{f1,f2,1-\alpha}}$$

mit $f1 = 2(x+1)$ und $f2 = 2(n-x)$

Hypothesentest

Für die Fragestellung, ob ein aktuell ermittelter p-Wert im „erwarteten Bereich“ liegt, wird der oben gezeigte Vertrauensbereich verwendet. Liegt p außerhalb dieses VBs, so ist z.B. die Nullhypothese abzulehnen, dass der „aktuelle Fehleranteil p“ der bisherigen Produktionsqualität entspricht (siehe Beispiel in der Visual-XSel Anwendung am Schluss).

Eine häufig genannte Näherung kann über die Hilfsvariable Z angewendet werden Die Nullhypothese „m Ereignisse entsprechen p_o“, wird abgelehnt, wenn $Z > u_{1-\alpha/2}$ ist (Voraussetzung n > 9/(p*(1-p))).

$$Z = \frac{m - n p_o}{\sqrt{n p_o (1-p_o)}}$$

© 2021 CRGRAPH

CRGRAPH

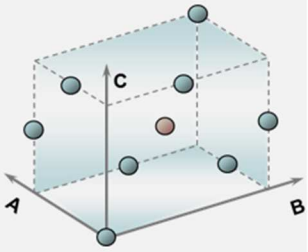
Literatur - Weiterführende Beschreibungen

Ausführliche softwareunabhängige Beschreibungen zum Thema DoE und der dazugehörigen Auswertungen gibt es im

Taschenbuch der statistischen Qualitäts- und Zuverlässigkeitsmethoden

Definitive Screening Designs DSD

Sogenannte Definitive Screening Designs sind sehr neu von Jones und Nachtsheim entwickelte Versuchspläne mit sehr geringem Versuchsumfang.
Sie ermöglichen die Auswertung von quadratischen Modellen und basieren deshalb auf 3 Stufen. Zwischen den Hauptfaktoren untereinander und den quadratischen Termen gibt es keine Vermengung (orthogonal). Die Wechselwirkungen sind nicht zu 100% vermengt.



Nr	A	B	C	D
1	0	1	-1	-1
2	0	-1	1	1
3	-1	0	-1	1
4	1	0	1	-1
5	-1	-1	0	-1
6	1	1	0	1
7	-1	1	1	0
8	1	-1	-1	0
9	0	0	0	0

In der generischen Erzeugung dieser Versuchspläne (iterativ mit Hilfe der Determinante) ergibt sich regulär die Anzahl Versuche mit n = 2*p+2. Manche Pläne, z.B. für p=5 sind dann allerdings teilweise zwischen den Hauptfaktoren vermengt. Hier müssen bis zu 3 Versuchszeilen ergänzt werden. Der Gesamtumfang ergibt sich somit zu:

n = 2*p+2+(1..3)

Alle Faktoren müssen durchgehend auf 3 Stufen sein und es lassen sich keine kategorialen Faktoren darstellen. Nachteilig ist auch, dass keine Auswertung aller möglichen

Weitere Informationen und Leseproben:
crgraph.de/Literatur

Consulting & Schulungen

Bei unseren Inhouse- oder Online-Schulungen wird die praxisnahe Anwendung von statistischen Methoden vermittelt. Wir haben über 25 Jahre Erfahrung, insbesondere in der Automobilindustrie und unterstützen Sie bei Ihren Problemstellungen, führen Auswertungen für Sie durch, oder erstellen firmenspezifische Auswertevorlagen.
Weitere Informationen finden Sie unter:
crgraph.de/schulungen
Sie haben ein konkretes Qualitätsproblem, oder wollen ein Produkt effizient und zuverlässig entwickeln? Sie wollen keine Statistik-Software anschaffen, weil diese voraussichtlich zu selten gebraucht wird, oder weil zu wenig Zeit zur Einarbeitung vorhanden ist? Dann sind unsere Q-Support Pakete genau das Richtige:
crgraph.de/consulting

Hotline

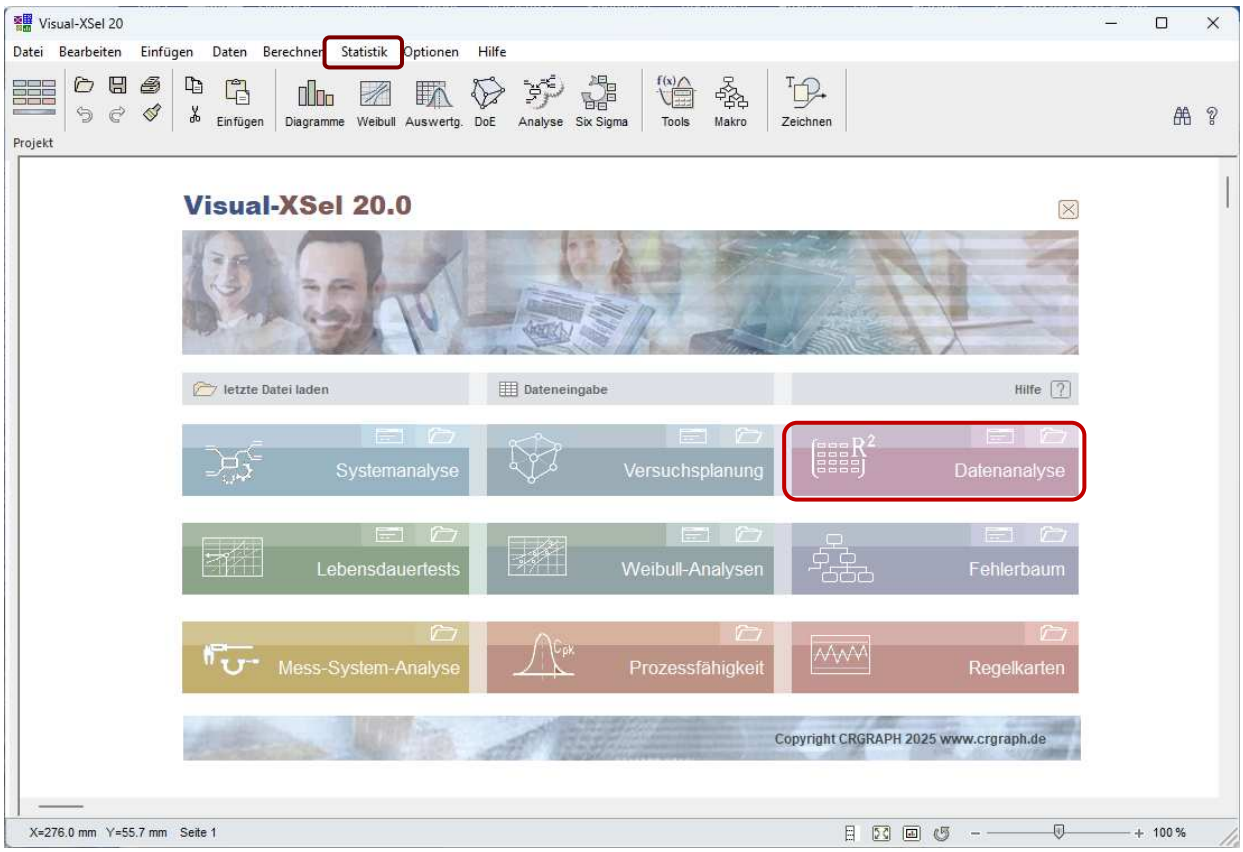
Haben Sie noch Fragen, oder Anregungen? Wir stehen Ihnen gerne zur Verfügung:
Tel. +49 (0)8151-9193638
E-Mail: info@crgraph.de
Besuchen Sie uns auf unserer Home-Page: www.crgraph.de

© 2021 CRGRAPH

Anwendung in Visual-XSel

www.crgraph.de

Unsere Software **Visual-XSel** ist ein leistungsfähiges Tool für alle wichtigen statistischen Qualitäts- und Zuverlässigkeitsmethoden. Die Binomial-Verteilung lässt sich als Grafik darstellen und ein Hypothesentest hierzu ist unter Datenanalyse möglich:



Hier finden Sie eine Übersicht und Einstiegsvideos:
crgraph.de/visual-xsel-software/

Youtube ein Video:

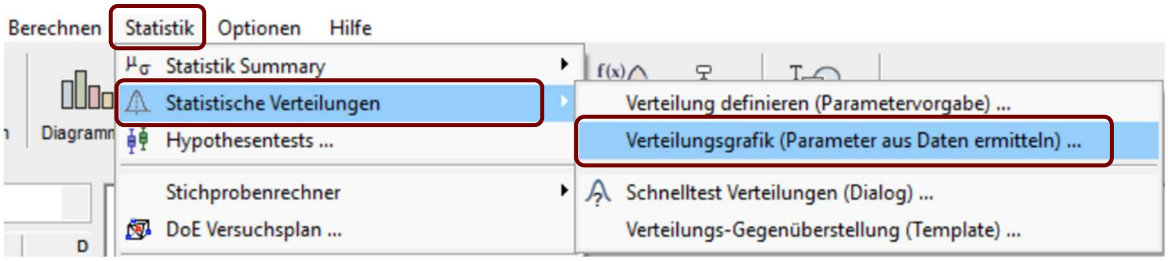
Nicht umsonst ist diese Software in vielen namhaften Firmen im Einsatz:
crgraph.de/Referenzen.

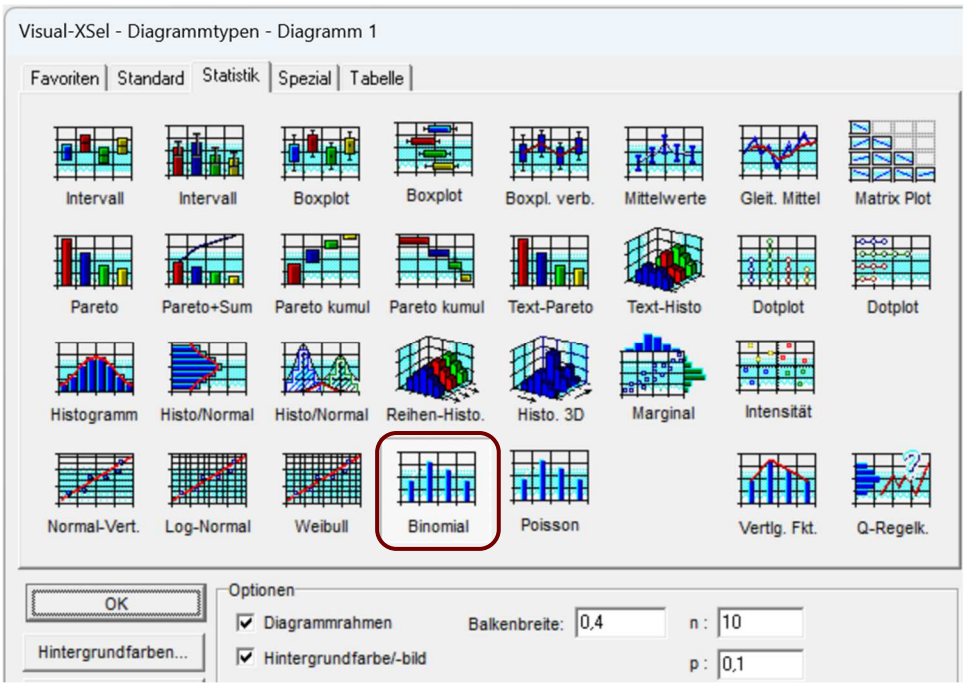
Die folgende Beschreibung ist eine Anleitung und Einführung in die Behandlung der Binomial-Verteilung in Visual-XSel.

© 2021 CRGRAPH

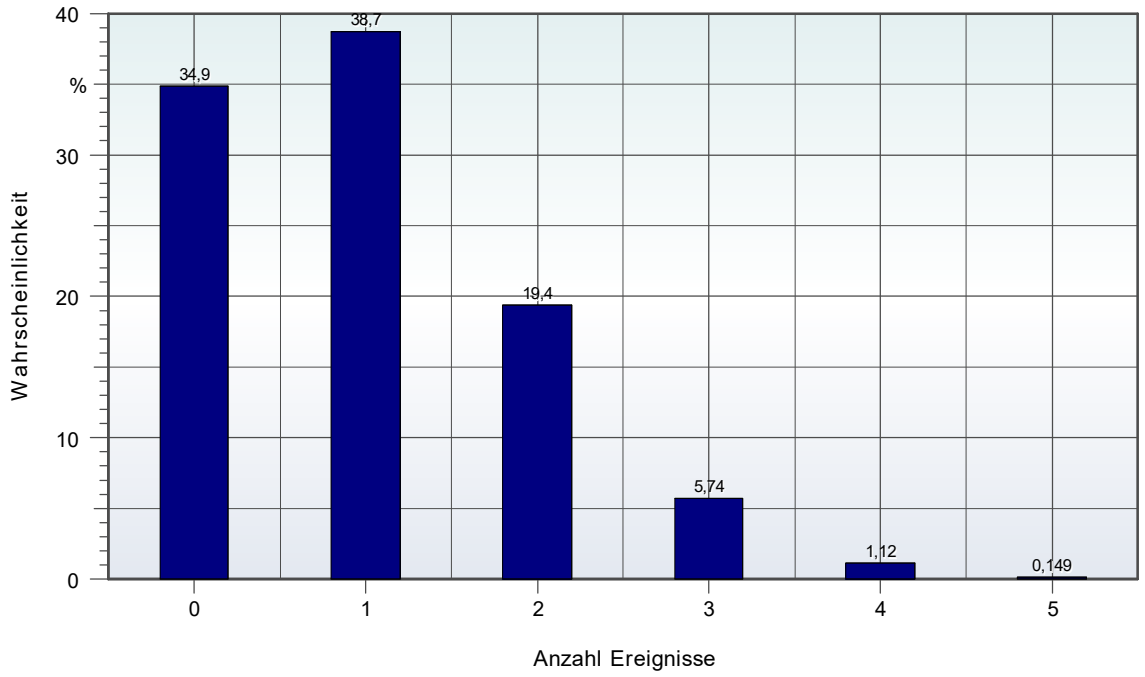
Binomial Grafik

Eine Binomial-Verteilung kann wie folgt als Grafik erzeugt werden:



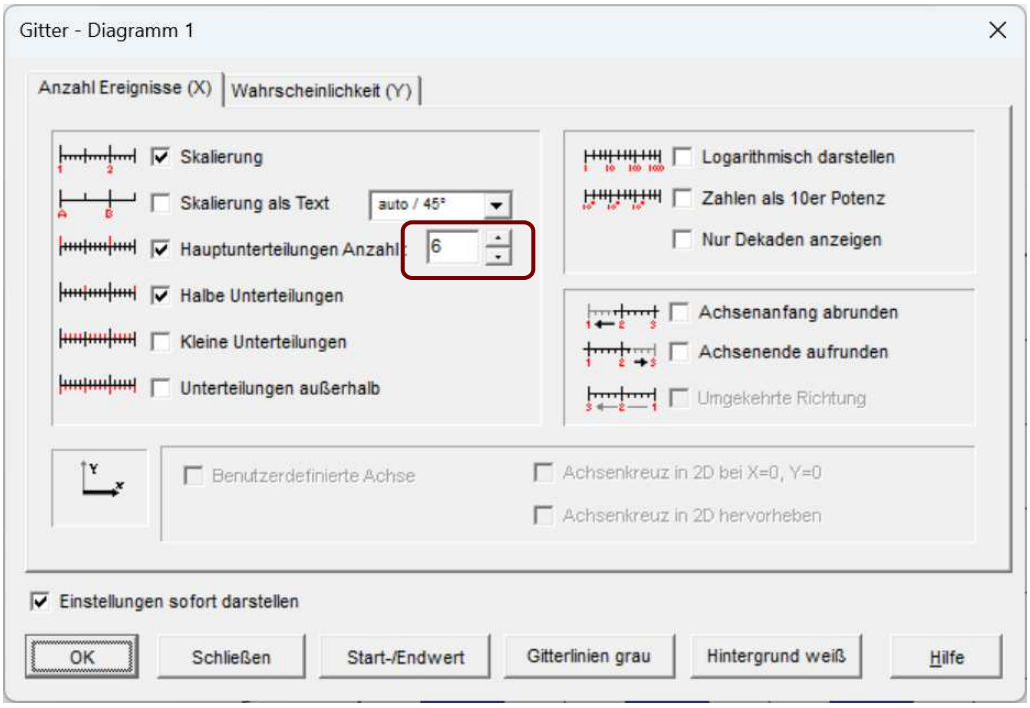


Wählen Sie unten rechts eine Stichprobengröße n und ein p. Für das gezeigte Beispiel ergibt sich:



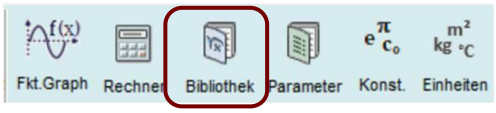
© 2021 CRGRAPH

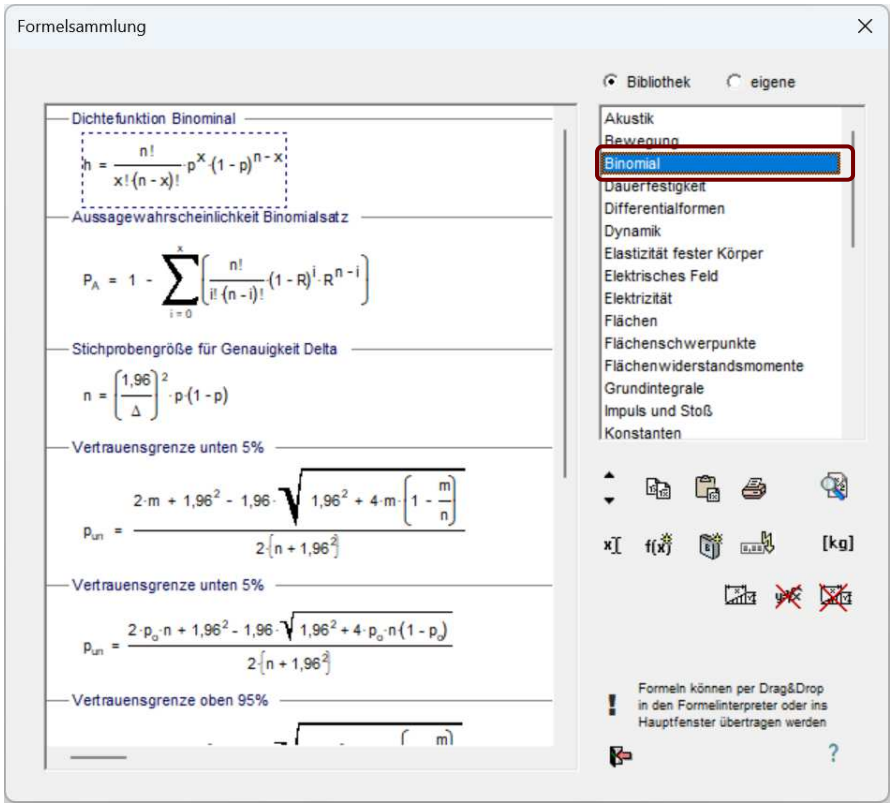
Hinweis: Es kann noch der Achsenbereich definiert werden, der hier auf 5 begrenzt wurde. Die Achsenskalierung kann u.U. nicht ganzzahlig sein. Dies kann in dem gezeigten Beispiel durch die Ikone ***Diagramm / Gitter*** und Reduzierung der Hauptunterteilung geändert werden:



Binomial Formeln

Über die Ikone ***Tools / Bibliothek*** lassen sich die Formeln der Binomial-Verteilung und die Vertrauensbereiche anzeigen und weiterverwenden.

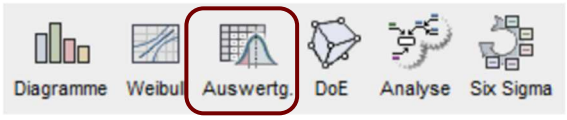


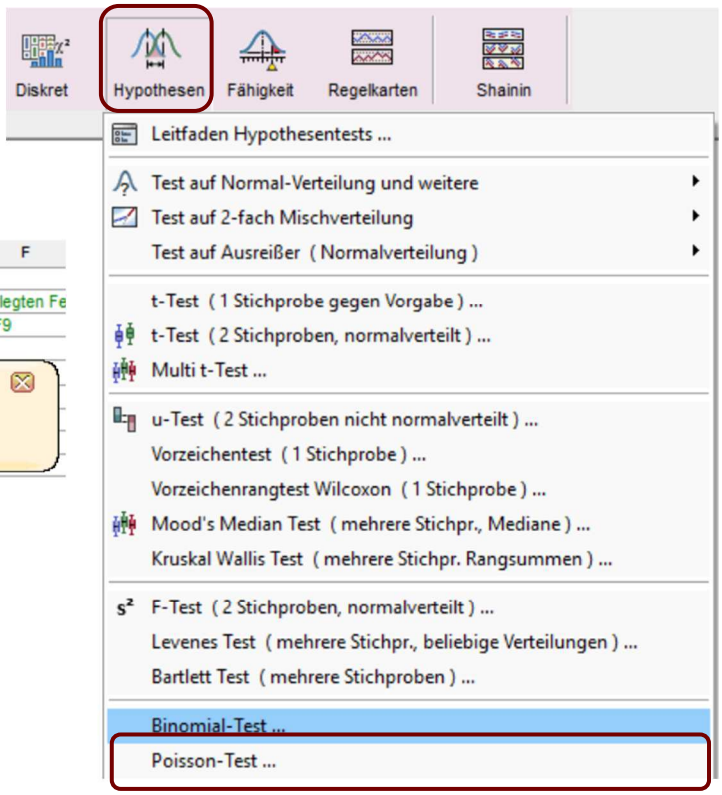


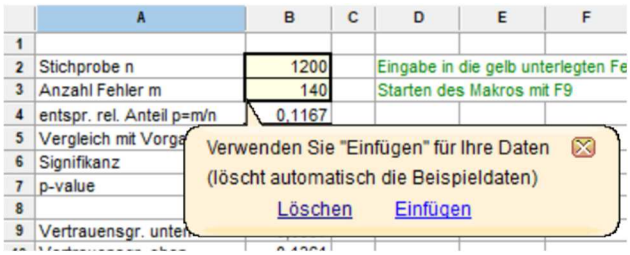
© 2021 CRGRAPH

Binomial Hypothesentest

Der Hypothesentest auf einen Fehleranteil gegen eine Vorgabe erfolgt über ein Template, dass über die Ikone ***Auswertung / Hypothesen*** geöffnet werden kann.
Das Template enthält zunächst Beispieldaten. Für eigene Daten sind in der Tabelle links in den gelben Feldern die Angaben

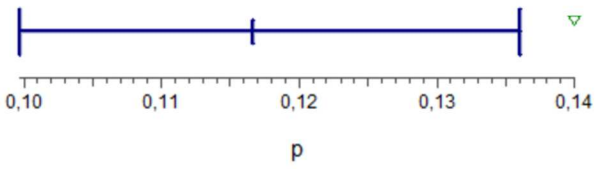






zu machen und das Makro mit F9 zu starten.

Stichprobe n	1200
Anzahl Fehler m	140
entspr. rel. Anteil p=m/n	0,1167
Vergleich mit Vorgabe p	0,1400
Signifikanz	0,05
p-value	0,020
Vertrauensgr. unten	0,0997
Vertrauensgr. oben	0,1361



Ergebnis:
Die Nullhypothese, dass der Fehleranteil m/n der Vorgabe p entspricht, wird verworfen!

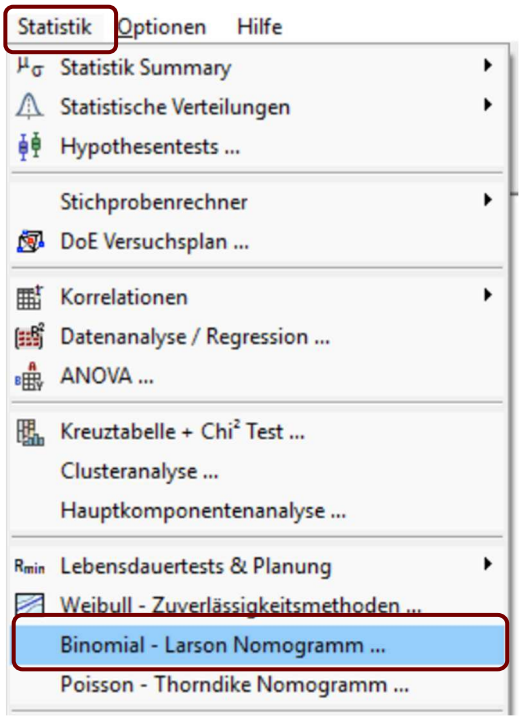
© 2021 CRGRAPH

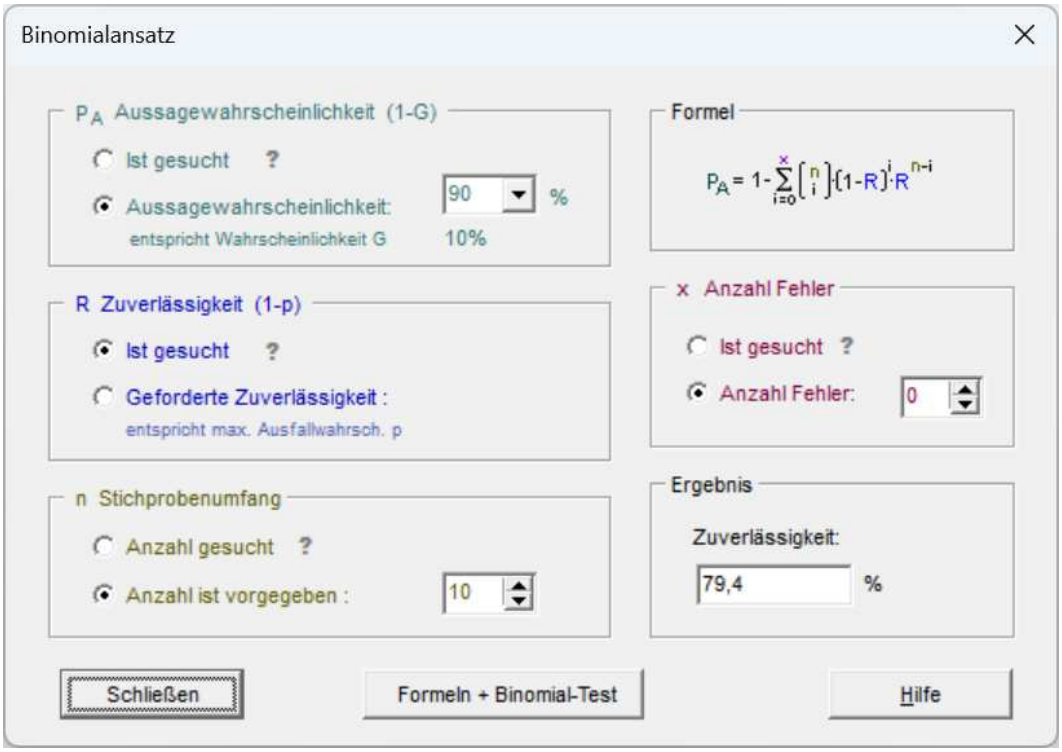
Binomialansatz – Larson-Nomogramm

Wie Eingangs beschrieben, kann über den Binomialansatz eine Aussagewahrscheinlichkeit über die Zuverlässigkeit von Bauteilen gemacht werden. In diesem Zusammenhang gibt es historisch das sogenannte Larson-Nomogramm.

In Visual-XSel lassen sich die Wahrscheinlichkeiten über eine Dialogbox berechnen. Der Aufruf erfolgt über den Menüpunkt ***Statistik / Binomial – Larson-Nomogramm***

Die Berechnung bezieht sich hier wieder aus technischer Sicht auf Fehler eines Bauteils oder eines Systems. Anstelle der Verteilungsfunktion G gibt es hier P_A = 1-G. Es kann hiermit indirekt das Ergebnis des Larson-Nomogramms berechnet werden.





© 2021 CRGRAPH