



Voraussetzung und verwandte Themen

Für diese Beschreibungen sind Grundlagen der Statistik vorteilhaft. Weiterführende und verwandte Themen sind:

www.versuchsmethoden.de/Hypothesentests.pdf

www.versuchsmethoden.de/ChiTests.pdf

Einführung

Der exakte Test nach Fisher ist ein Signifikanztest auf Unabhängigkeit von Häufigkeiten in einer Kontingenztafel (Vierfeld-Tabelle). Das Anwendungsgebiet ist das gleiche, wie beim χ^2 -Test.

Ziel und Nutzen

Im Gegensatz zum χ^2 -Test liefert der exakte Test nach Fisher jedoch auch bei sehr kleinen Stichproben (Einzelhäufigkeiten < 5) zuverlässige Ergebnisse.

Grundlagen

Beispiel: Es werden zwei Gruppen befragt (Damen und Herren), wie Ihnen ein Produkt gefällt:

	Damen	Herren
gefällt	4	5
gefällt nicht	3	0

Die Frage ist, ob die Damen und die Herren das Produkt gleich bewerten, oder nicht. Allgemein formuliert wird die Kontingenztafel folgendermaßen dargestellt (inkl. Zeilen- und Spaltensummen = Randsummen):

	Gruppe 1	Gruppe 2	Σ
Merkmal 1	a	b	a+b
Merkmal 2	c	d	c+d
Σ	a+c	b+d	n = a+b+c+d

a,b,c und d sind ganzzahlige Häufigkeiten.

Der exakte Test bildet Kombinationen von Zellhäufigkeiten, die entstehen könnten. Der jeweilige Extremfall ist, dass ein Feld zu 0 wird. Für diese Kombinationen werden die Wahrscheinlichkeiten berechnet und addiert. Diese Wahrscheinlichkeiten folgen einer hypergeometrischen Verteilung. Die Wahrscheinlichkeit ist dann:

$$P = \frac{\binom{a+b}{a} \binom{c+d}{c}}{\binom{n}{a+c}} = \frac{(a+b)!(c+d)!(a+c)!(b+d)!}{n!a!b!c!d!}$$

Exakter Fisher Test

Die eingangs beschriebene Fragestellung wird nun zu einer Nullhypothese formuliert H_0 : Die Häufigkeiten in den Gruppen sind gleich, gegen die Alternativhypothese: H_1 : Die Häufigkeiten der Gruppen sind unterschiedlich.

Hierfür sind die Wahrscheinlichkeiten zu berechnen, mit der man die verschiedenen möglichen Kombinationen der Kontingenztafel beobachten kann (unter der Gültigkeit von H_0). Es reicht aus, nur das Feld a zu variieren. Reduziert man a bis in die Extrembereiche, so müssen die anderen Felder unter Berücksichtigung konstanter Randsummen angepasst werden. Die Kombinationen gehen soweit, bis Felder < 0 entstehen. Das gleiche macht man auf der anderen Seite durch Erhöhung von a. Danach werden alle Wahrscheinlichkeiten der Variationen addiert, die kleiner der Ausgangswahrscheinlichkeit P sind. Der p_{value} errechnet sich nun durch:

$$p_{value} = \sum_{P_i < P} P_i + P$$

i = Index aller möglichen Kombinationen bis in die Extrembereiche

H_0 wird abgelehnt, wenn der p_{value} kleiner einem vorgegebenen Signifikanzniveau ist (in der Regel $\alpha = 0,05$).

Man spricht hier von einem exakten Test, weil bei der Festlegung von H_0 und des Signifikanzniveaus die exakte Verteilung der Teststatistik verwendet wird und nicht eine Approximation, wie bei der χ^2 -Statistik.

Zurück zu dem Beispiel am Anfang:

	Damen	Herren	Σ
gefällt	4	5	9
gefällt nicht	3	0	3
Σ	7	5	n = 12

Die erste Variation ist die Reduzierung des oberen linken Feldes a reduziert um 1. Diese Kombination fällt jedoch aus, da sonst das Feld d < 0 werden würde. Die erst mögliche Kombination nach oben ist:

	Damen	Herren	Σ
gefällt	5	4	9
gefällt nicht	2	1	3
Σ	7	5	n = 12

Das Gesamtergebnis ist: $p_{value} = 0,205$

Obwohl den Herren im Gegensatz zu den Damen das Produkt ausschließlich gefällt, kann die Nullhypothese auf Gleichheit der befragten Gruppen nicht abgelehnt werden. Erst ab dem folgendem Verhältnis wären mit $p_{value} = 0,048$ die Gruppen als unterschiedlich anzusehen:

	Damen	Herren
gefällt	1	5
gefällt nicht	4	0

Exakter Fisher Test

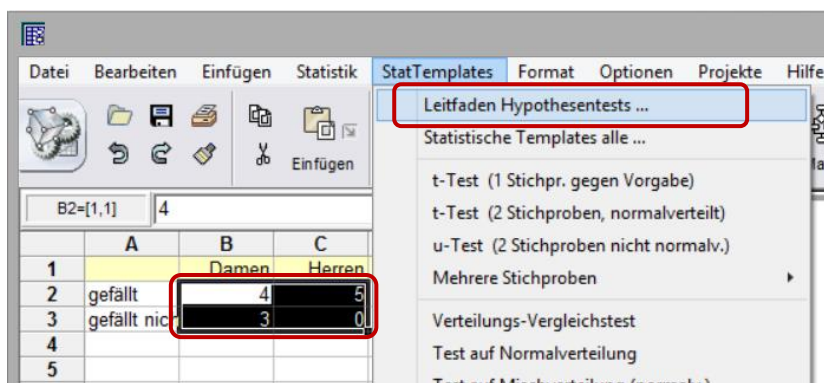
Auf der einen Seite lässt sich zwar eine Kontingenztabelle mit geringen Häufigkeiten exakt berechnen, anhand von diesem Beispiel erkennt man aber, dass zur Erkennung von Unterschieden letztlich doch eine deutlich größere Stichprobe sinnvoll ist.

Für diese Methode steht das Visual-XSel Template *StatTest_FisherExactTest.vxd* unter StatistikTests zur Verfügung. Hiermit können auch relativ große Zahlen verarbeitet werden, was in manchen anderen Anwendungen wegen der Fakultät oft ein Problem ist.

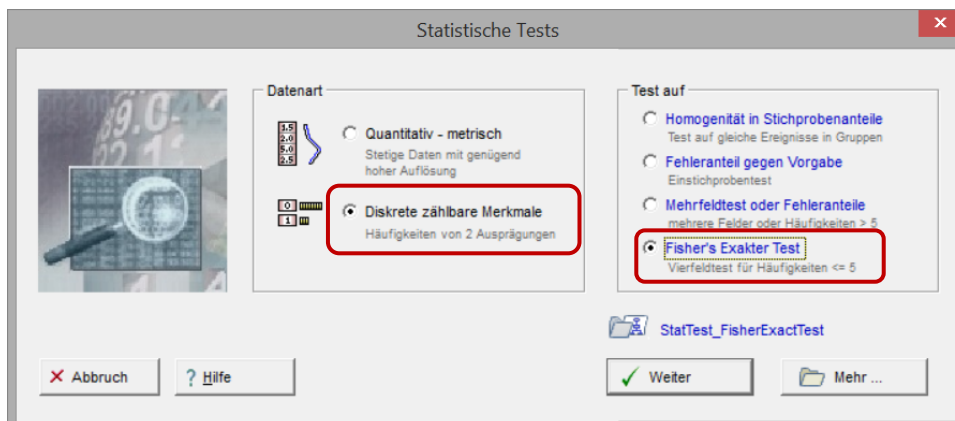
Anwendung in Visual-XSel 14.0

www.crgraph.de

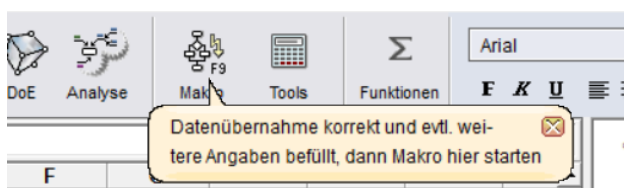
Das Template für den exakten Fisher-Test kann direkt über *Datei/Templates/ Statistische Tests*, oder über den Leitfaden aufgerufen werden. Vorher müssen in der Tabelle die folgenden 4 Felder markiert worden sein:



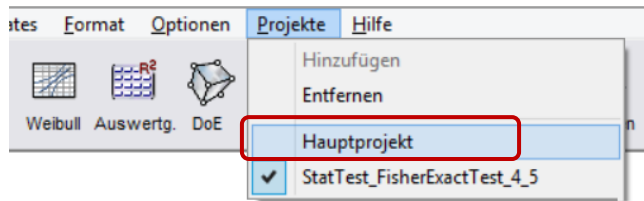
Es folgt eine Abfrage, in der folgende Auswahl zu treffen ist:



Nach der Auswahl wird das Template in die bestehende Datei hinzugeladen und es ist das Makro zu starten:



Um wieder auf die ursprüngliche Tabellenseite zu kommen ist unter dem Menüpunkt Projekte wieder das Hauptprojekt auszuwählen:



Literatur

Taschenbuch der statistischen Qualitäts- und Zuverlässigkeitsmethoden

Die wichtigsten Methoden und Verfahren für die Praxis.

Beinhaltet statistische Methoden für Versuchsplanung & Datenanalyse, sowie Zuverlässigkeit & Weibull.

- Statistische Verteilungen und Tests & Mischverteilungen
- Six Sigma Einführung und Zyklen
- Systemanalysen Wirkdiagramm, FMEA, FTA, Matrizen-Methoden
- Shainin- und Taguchi-Methoden
- Versuchsplanung DoE, D-Optimal
- Korrelations- und Regressionsverfahren
- Multivariate Datenauswertungen
- Prozessfähigkeit – Messmittelfähigkeit MSA 4 und VDA 5
- Regelkarten
- Toleranzrechnung und Monte-Carlo-Simulation
- Statistische Hypothesentests
- Weibull und Lebensdaueranalysen
- Stichprobengröße

190 Seiten, Ringbuch

ISBN: 978-3-00-043678-9

