Kapitel 9: Verfahren für Nominaldaten

Eindimensionaler Chi²-Test

Berechnen der Effektgröße w²

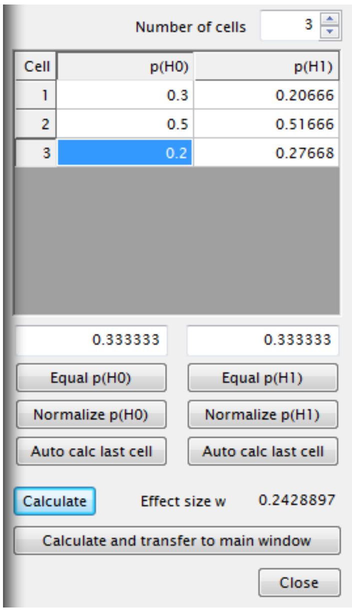
Die empirische Effektgröße w^2 gibt die Größe der Abweichung zwischen den erwarteten und beobachteten Häufigkeiten in standardisierter Form an. G*Power verwendet die Wurzel aus w^2 als Effektgröße. Sowohl w als auch w^2 sind inhaltlich schwer zu interpretieren. Nach Cohen (1988) gelten folgende Konventionen (vgl. Kap. 9.1.4).

$w^2 = 0{,}01$	$w = 0{,}1$	kleiner Effekt
$w^2 = 0{,}09$	$w = 0{,}3$	mittlerer
$w^2 = 0{,}25$	$w = 0{,}5$	großer Effekt

Diese Werte können als Anhaltspunkt verwendet werden, um eine Stichprobenumfangsplanung bzw. eine Teststärkeanalyse a priori durchzuführen. Alternativ kann die erwartete Effektstärke aus vorangegangenen Untersuchungen geschätzt werden. Es gibt mehrere Möglichkeiten, w² zu schätzen. Zum einen ist die möglich über die folgende Formel, die die absolute Größe der Abweichungen (Chi²-Wert) an der Stichprobengröße relativiert: $\hat{w}^2 = \frac{\chi^2}{N}$.

Durch diese Standardisierung lassen sich die Effektgrößen auch aus verschiedenen Untersuchungen miteinander vergleichen. Mit Hilfe dieser Formel lässt sich ein empirischer Effekt berechnen, der benötigt wird, um die Teststärke für eine Untersuchung a posteriori zu berechnen.

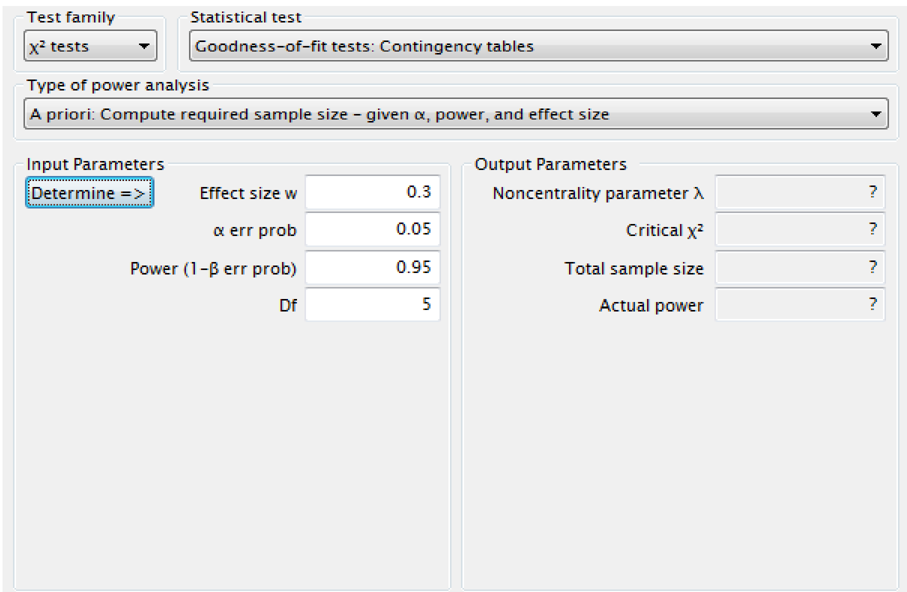
Alternativ kann in G*Power die Effektstärke mit Hilfe von Zellwahrscheinlichkeiten berechnet werden, die unter der Null- und Alternativhypothese erwartet werden („Test family“: χ^2 tests, „Statistical test: Goodness-of-fit tests: Contingeny tables“, Fenster „Determine“). Dafür müssen Sie die theoretisch erwarteten Häufigkeiten in Wahrscheinlichkeiten umrechnen. In die Felder p(H0) notieren Sie die erwarteten Wahrscheinlichkeiten, in die Felder p(H1) notieren Sie die empirischen Wahrscheinlichkeiten. Die Tabellen 9.3 und 9.4 in Kapitel 9.1 liefern die empirischen und die erwarteten Häufigkeiten für das Beispiel „Klinik“. Wenn Sie diese Häufigkeiten an der Gesamthäufigkeit relativieren und in die entsprechenden Felder übertragen, erhalten Sie folgendes Bild:



Wie erwartet entspricht der ermittelte empirische Effekt von w = .2429 dem in Kapitel 9.1.4 per Hand ermittelten Effekt von w^2 = 0,059.

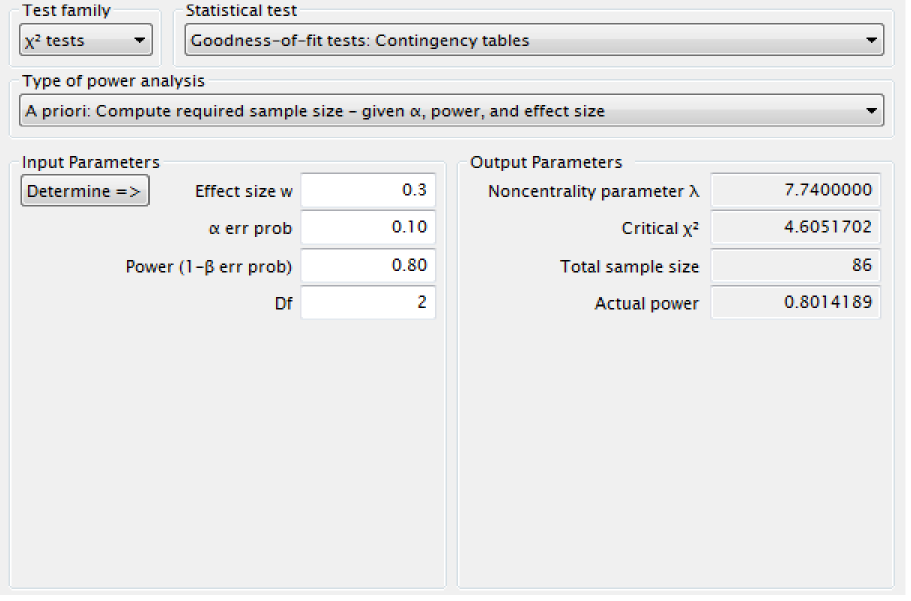
Stichprobenumfangsplanung bzw. Teststärkeanalyse a priori

Um zu ermitteln, wie viele Versuchspersonen benötigt werden, um einen erwarteten Effekt mit einer bestimmten Power zu entdecken, muss eine Teststärkeanalyse durchgeführt werden (siehe Kap. 9.1.6). Öffnen Sie hierzu G*Power und wählen Sie „Test family“: χ^2 tests, „Statistical test: Goodness-of-fit tests: Contingeny tables“, “Type of power analysis: A priori”. Sie sehen nun folgendes Fenster vor sich:



Tragen Sie die erwartete Effektgröße w ein und spezifizieren Sie das α-Fehlerniveau, sowie die gewünschte Teststärke. Die Freiheitsgrade für den eindimensionalen χ^2-Test ergeben sich nach der Berechnungsvorschrift: df = k – 1, wobei k der Anzahl der untersuchten Gruppen oder Kategorien entspricht (Kap. 9.1.2).

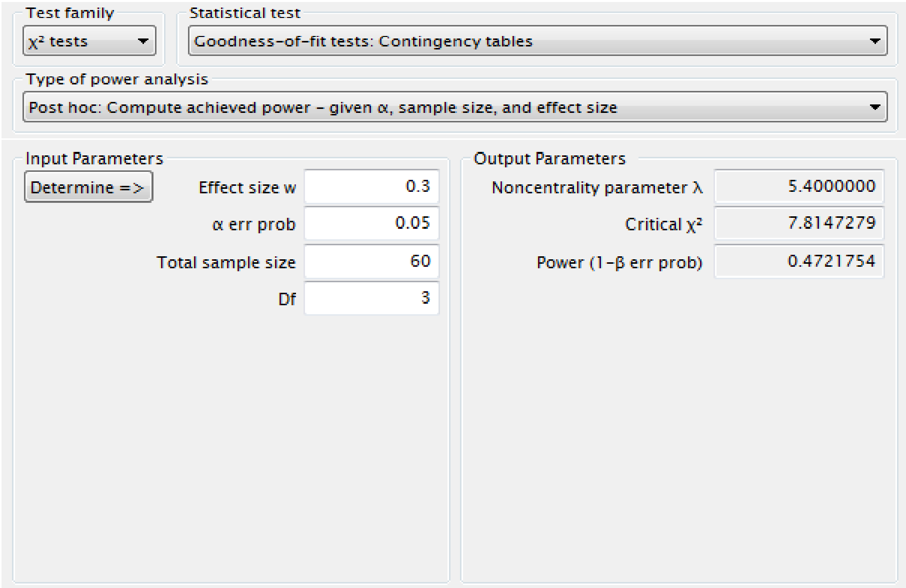
Wie viele Probanden müssen erhoben werden, um in einem eindimensionalen Chi²-Test mit drei Gruppen (df = 2) einen mittleren Effekt der Größe w = 0,3 mit einer Wahrscheinlichkeit von 80%, bei einem α -Fehlerniveau von 10% zu entdecken? Wenn wir die entsprechenden Eingaben machen, liefert das Programm eine geforderte Stichprobengröße von N = 86:



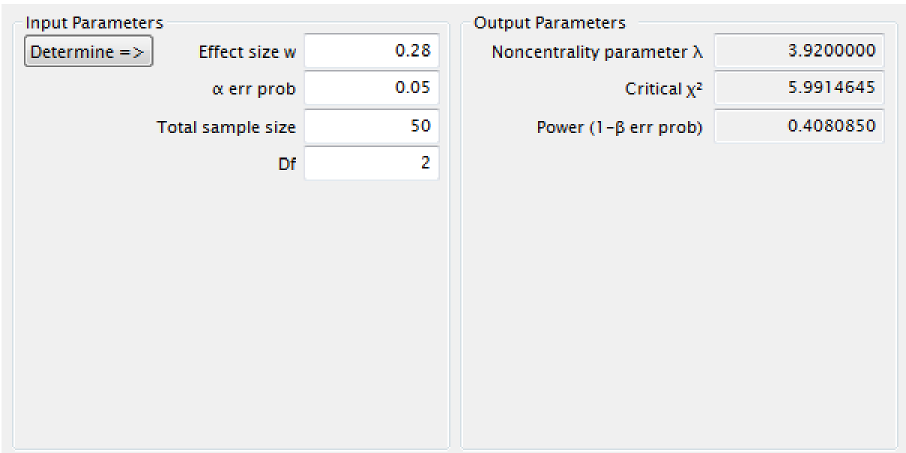
Berechnen der Teststärke a posteriori

Um die Teststärke einer Studie a posteriori zu schätzen, wechseln Sie bei „Type of power analysis“ auf post hoc. Geben Sie den empirisch ermittelten Effekt w (zur Ermittlung siehe oben), das α-Niveau, die Stichprobengröße sowie die Freiheitsgrade des Tests ein. G*Power ermittelt dann die Power des Tests, also die Wahrscheinlichkeit, mit der Sie unter den bestehenden Bedingungen einen Effekt der bestehenden Größe auch tatsächlich entdecken konnten.

Ein Forscher hat über seine vier Bedingungen (df = 3) hinweg ein nicht signifikantes Ergebnis erzielt (α = 5%). In jeder Bedingung erfasste er 15 Versuchspersonen. Der Forscher erachtet einen mittleren Effekt von w = 0,3 für inhaltlich relevant. Wie groß war seine Chance, einen Effekt dieser Größe zu finden? Die Berechnung mit G*Power zeigt, dass der Forscher eben so gut eine Münze hätte werfen können (1-ß = 0,47).



Ein anderer Forscher möchte die Effektstärke für seinen empirisch ermittelten Effekt berechnen. Nehmen wir beispielsweise an, er habe bei 50 Versuchspersonen und zwei Freiheitsgarden einen empirischen Chi²-Wert von 3,90 ermittelt, der bei einem Signifikanzniveau von $\alpha = 5\%$ nicht signifikant geworden ist ($p = .14$). Unter Berücksichtigung der entsprechenden Effektgröße $w = .28$ und der restlichen Daten, erhalten Sie folgenden Output:



Neben dem kritischen Chi²-Wert (5,99), der die Signifikanzgrenze markiert, ermittelt G*Power eine Post hoc Teststärke von .4081. Das heißt, die Wahrscheinlichkeit, den bestehenden Effekt mit der gegebenen Stichprobengröße als signifikant nachzuweisen, war sehr gering. Insofern ist die Schlussfolgerung, es bestehe kein Effekt, in diesem Fall nicht gerechtfertigt. Vielmehr ist die richtige Schlussfolgerung, dass der Effekt zwar von mittlerer Größe ist, aber die Power nicht ausreichend hoch war, um ihn als statistisch signifikant nachzuweisen.

Zweidimensionaler und Vierfelder Chi²-Test

Beim zweidimensionalen χ^2-Test aus Kapitel 9.2 sowie beim Vierfelder Chi²-Test (Kap. 9.3) ist die Vorgehensweise bei der Teststärkenanalyse ähnlich zum eindimensionalen Fall. Auch hier wird die Effektstärke entweder über die Formel $\hat{w}^2 = \frac{\chi^2}{N}$ bestimmt oder in G*Power über die erwarteten und empirischen Wahrscheinlichkeiten (siehe oben). Es gelten dieselben Konventionen für die Effektgröße wie oben vorgestellt. Lediglich die Berechnung der Freiheitsgrade erfolgt nun nach der Beziehung $df = (k - 1) \cdot (l - 1)$, wobei k und l die Anzahl der Gruppen bzw. Kategorien für die beiden untersuchten Dimensionen angibt (beim Vierfelder Chi²-Test betragen die Freiheitsgrade immer 1). Da die Vorgehensweise ansonsten identisch zum eindimensionalen Fall ist, wird hier auf eine Beispielrechnung verzichtet. Allerdings ist noch darauf hinzuweisen, dass G*Power nicht mit den alternativen Effektstärken Cramers Index (Kap. 9.2.3) und dem Phi-Koeffizienten (Kap. 9.3.1) arbeitet, alle Berechnungen erfolgen über w.

Literatur

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NY: Erlbaum.