

Aufgaben zu Kapitel 9

Aufgabe 1

Für diese Aufgabe benötigen Sie den Datensatz „Nominaldaten.sav“.

```
library(foreign)
nominal <- read.spss("Nominaldaten.sav",
                    to.data.frame = TRUE)
```

- Sie arbeiten für eine Marktforschungsfirma und sollen überprüfen, ob die in diesem Datensatz untersuchte Stichprobe repräsentativ ist hinsichtlich der Geschlechterverteilung (50:50) sowie dem Landesteil (Westdeutschland und Ostdeutschland, Verhältnis 80:20).
- Berechnen Sie für die Variable „Landesteil“ die zugehörige Effektstärke und bestimmen Sie die Teststärke mit G*Power.
- Unterscheidet sich die Geschlechterverteilung in der Stichprobe je nach Landesteil?
- Berechnen Sie die Effektstärke und die Teststärke a posteriori für den Zusammenhang von Landesteil und Geschlechterverteilung. Welcher Stichprobenumfang wäre nötig, um diesen Effekt mit einer Power von mindestens 80% zu entdecken, wenn man dabei einen α -Fehler von 20% akzeptiert?

Quelle: <https://lehrbuch-psychologie.springer.com/content/zu-den-spss-r-und-gpower-aufgaben-und-ergaenzungen>

Aus: Rasch, Frieze, Hofmann & Naumann (2021). *Quantitative Methoden. Band 2*, 5. Auflage. Heidelberg: Springer.

© Springer-Verlag GmbH Deutschland, ein Teil von Springer Nature

Lösungen

Aufgabe 1

- a) Den entsprechenden eindimensionalen Chi-Quadrat-Test für das Geschlecht erhalten wir über die Funktion `chisq.test()`. Erstellen Sie zunächst eine Häufigkeitstabelle mit der Variable „geschlecht“ und geben Sie diese dann in die Funktion ein. Geben Sie ansonsten nichts zusätzlich an, da beim Geschlecht auf Gleichverteilung (50:50) getestet werden muss, welches der Standardeinstellung entspricht.

```
geschl <- table(nominal$geschl)
chisq.test(geschl)
```

Sie erhalten folgenden Output:

```
Chi-squared test for given probabilities

data:  geschl
X-squared = 0.22774, df = 1, p-value = 0.6332
```

Der Unterschied ist bei weitem nicht signifikant. Der Datensatz ist hinsichtlich der Geschlechterverteilung nahezu repräsentativ.

Für den Landesteil erstellen Sie wieder eine Häufigkeitstabelle. Geben Sie diese in die Funktion `chisq.test()` ein und tragen Sie danach mit dem Argument `p` die erwarteten relativen Häufigkeiten ein. Achten Sie auf die Reihenfolge, damit die erwartete relative Häufigkeit auch dem richtigen Landesteil zugeordnet wird:

$p_w = 0,80$ für die westdeutsche Gruppe (kodiert mit 1) und

$p_o = 0,20$ für die ostdeutsche Gruppe (kodiert mit 2)

```
Landesteil <- table(nominal$Landesteil)
chisq.Land <- chisq.test(Landesteil, p = c(0.80, 0.20))
chisq.Land
```

Sie erhalten folgenden Output:

```
Chi-squared test for given probabilities

data:  Landesteil
X-squared = 351.63, df = 1, p-value < 2.2e-16
```

Die beobachteten Häufigkeiten entsprechen nicht den erwarteten Häufigkeiten. Wie Sie sehen, wurde der Chi-Quadrat-Test dem Objekt `chisq.Land` zugewiesen. Dies hat den Grund, dass Sie sich nun mit den Befehlen `chisq.Land$expected` und `chisq.Land$observed` die beobachteten und erwarteten Häufigkeiten ausgeben lassen können. Mit der Funktion `cbind()` können Sie sich auch eine Tabelle erstellen. Diese verrät, wie die Verhältnisse der Gruppen zueinander sind:

Quelle: <https://lehrbuch-psychologie.springer.com/content/zu-den-spss-r-und-gpower-aufgaben-und-ergaenzungen>

Aus: Rasch, Friese, Hofmann & Naumann (2021). *Quantitative Methoden. Band 2*, 5. Auflage. Heidelberg: Springer.

© Springer-Verlag GmbH Deutschland, ein Teil von Springer Nature

Aufgaben mit R und G*Power

Rasch, Frieze, Hofmann & Naumann (2021). *Quantitative Methoden. Band 2* (5. Auflage). Heidelberg: Springer.

```
cbind(erwartet = chisq.Land$expected,  
      beobachtet = chisq.Land$observed)
```

	erwartet	beobachtet
westdeutschland	1015.2	748
ostdeutschland	253.8	521

In der Stichprobe sind Westdeutsche signifikant unterrepräsentiert und Ostdeutsche überrepräsentiert.

b) Effektstärke für den Faktor Landesteil:

Test family: χ^2 Tests. Statistical test: Goodness-of-fit-tests: Contingency tables. Klicken Sie auf „Determine“, um die Effektstärke zu bestimmen. Erwartet waren 80% West- und 20% Ostdeutsche. Die empirischen relativen Wahrscheinlichkeiten berechnen Sie als Anteil der empirischen Häufigkeiten an der Stichprobengröße und tragen diese unter p(H1) ein. Sie erhalten das folgende Fenster.

Cell	p(H0)	p(H1)
1	0.8	0.5894
2	0.2	0.4106

Number of cells: 2

Effect size w: 0.5265

Die Analyse ergibt ein w von 0,53. Das entspricht einem großen Effekt. Ostdeutsche sind also in der Stichprobe bedeutsam überrepräsentiert. Analog können Sie w an Hand der in Kapitel 9

vorgestellten Formel berechnen: $w = \sqrt{\frac{\chi^2}{N}} = \sqrt{\frac{350,901}{1269}} = 0,53$

Quelle: <https://lehrbuch-psychologie.springer.com/content/zu-den-spss-r-und-gpower-aufgaben-und-ergaenzungen>

Aus: Rasch, Frieze, Hofmann & Naumann (2021). *Quantitative Methoden. Band 2*, 5. Auflage. Heidelberg: Springer.

© Springer-Verlag GmbH Deutschland, ein Teil von Springer Nature

Aufgaben mit R und G*Power

Rasch, Friese, Hofmann & Naumann (2021). *Quantitative Methoden. Band 2* (5. Auflage). Heidelberg: Springer.

Teststärkeanalyse mit G*Power:

Da es sich um eine Post hoc Analyse handelt, muss unter dem Menü „Type of power analysis“ die Option „Post hoc“ ausgewählt werden. G*Power berechnet die folgende Teststärke:

The screenshot shows the G*Power software interface. At the top, 'Test family' is set to 'χ² tests' and 'Statistical test' is 'Goodness-of-fit tests: Contingency tables'. The 'Type of power analysis' is 'Post hoc: Compute achieved power - given α, sample size, and effect size'. Under 'Input Parameters', 'Effect size w' is 0.5265000, 'α err prob' is 0.05, 'Total sample size' is 1269, and 'Df' is 1. Under 'Output Parameters', 'Noncentrality parameter λ' is 351.7697, 'Critical χ²' is 3.8414588, and 'Power (1-β err prob)' is 1.0000000.

Input Parameters		Output Parameters	
Determine =>	Effect size w	Noncentrality parameter λ	351.7697
	α err prob	Critical χ²	3.8414588
	Total sample size	Power (1-β err prob)	1.0000000
	Df		

Die Teststärke approximiert den theoretisch möglichen Maximalwert 1. Ein Effekt dieser Größe hatte damit maximale Chancen, entdeckt zu werden.

- c) Um zu untersuchen, ob sich die Geschlechterverteilung je nach Landesteil unterscheidet, muss ein zweidimensionaler Chi-Quadrat-Test durchgeführt werden. Verwenden Sie hierfür die Funktion `CrossTable()` des Pakets `gmodels`. Dabei spielt es keine Rolle, welche der beiden Variablen den Zeilen, entspricht der ersten, und welche den Spalten, entspricht der zweiten, zugewiesen wird. Mit `chisq = TRUE` geben Sie an, dass ein Chi-Quadrat-Test durchgeführt werden soll, mit `expected = TRUE` erhalten Sie die erwarteten Häufigkeiten und `format = "SPSS"` bekommen Sie den Output im SPSS-Format ausgegeben. Für die Effektstärken verwenden Sie die Funktionen `Cramerv()` und `Phi()` des Pakets `DescTools`. Dies erleichtert die Interpretation.

Quelle: <https://lehrbuch-psychologie.springer.com/content/zu-den-spss-r-und-gpower-aufgaben-und-ergaenzungen>

Aus: Rasch, Friese, Hofmann & Naumann (2021). *Quantitative Methoden. Band 2*, 5. Auflage. Heidelberg: Springer.

© Springer-Verlag GmbH Deutschland, ein Teil von Springer Nature

```
install.packages("gmodels")
library(gmodels)
crossTable(nominal$Landesteil, nominal$geschl,
           chisq = TRUE, expected = TRUE, format = "SPSS")
install.packages("DescTools")
library(DescTools)
Cramerv(nominal$Landesteil, nominal$geschl, conf.level = 0.95)
Phi(nominal$Landesteil, nominal$geschl)
```

Der Output sieht dann folgendermaßen aus:

Quelle: <https://lehrbuch-psychologie.springer.com/content/zu-den-spss-r-und-gpower-aufgaben-und-ergaenzungen>

Aus: Rasch, Frieze, Hofmann & Naumann (2021). *Quantitative Methoden. Band 2*, 5. Auflage. Heidelberg: Springer.

© Springer-Verlag GmbH Deutschland, ein Teil von Springer Nature

Aufgaben mit R und G*Power

Rasch, Frieze, Hofmann & Naumann (2021). *Quantitative Methoden. Band 2* (5. Auflage). Heidelberg: Springer.

```

Cell Contents
|-----|
|              Count |
|      Expected Values |
| Chi-square contribution |
|              Row Percent |
|              Column Percent |
|              Total Percent |
|-----|

Total Observations in Table: 1269

              | nominal$geschl
nominal$Landesteil | weiblich | männlich | Row Total |
|-----|-----|-----|-----|
westdeutschland |      398 |      350 |      748 |
|      368.990 |      379.010 |      |
|      2.281 |      2.221 |      |
|      53.209% |      46.791% |      58.944% |
|      63.578% |      54.432% |      |
|      31.363% |      27.581% |      |
|-----|-----|-----|-----|
ostdeutschland |      228 |      293 |      521 |
|      257.010 |      263.990 |      |
|      3.275 |      3.188 |      |
|      43.762% |      56.238% |      41.056% |
|      36.422% |      45.568% |      |
|      17.967% |      23.089% |      |
|-----|-----|-----|-----|
Column Total |      626 |      643 |      1269 |
|      49.330% |      50.670% |      |
|-----|-----|-----|-----|

Statistics for All Table Factors

Pearson's Chi-squared test
-----
chi^2 = 10.96385    d.f. = 1    p = 0.0009290668

Pearson's Chi-squared test with Yates' continuity correction

```

Quelle: <https://lehrbuch-psychologie.springer.com/content/zu-den-spss-r-und-gpower-aufgaben-und-ergaenzungen>

Aus: Rasch, Frieze, Hofmann & Naumann (2021). *Quantitative Methoden. Band 2*, 5. Auflage. Heidelberg: Springer.

© Springer-Verlag GmbH Deutschland, ein Teil von Springer Nature

Aufgaben mit R und G*Power

Rasch, Frieze, Hofmann & Naumann (2021). *Quantitative Methoden. Band 2* (5. Auflage). Heidelberg: Springer.

```
-----  
Chi^2 = 10.58917    d.f. = 1    p = 0.001137517
```

```
Minimum expected frequency: 257.0102
```

```
Cramer V      lwr.ci      upr.ci  
0.09295028  0.03792824  0.14796854
```

[1] 0.09295028

In der westdeutschen Stichprobe sind Frauen etwas überrepräsentiert (53% vs. 47%). In der ostdeutschen Stichprobe dagegen haben mehr Männer als Frauen teilgenommen (56% vs. 44%). Der Zusammenhang zwischen Landesteil und Geschlecht ist statistisch signifikant. Insgesamt ist das Geschlechterverhältnis in der Stichprobe ausgeglichen (siehe Aufgabe 1a), innerhalb der Landesteile gibt es aber signifikante „Verzerrungen“.

- d) Effektstärke: Die Effektstärke berechnen Sie in G*Power, in dem Sie auf „Determine“ klicken und dann die erwarteten relativen Wahrscheinlichkeiten (erwartete Häufigkeit/Gesamtzahl Fälle) und die empirischen relativen Wahrscheinlichkeiten (empirische Häufigkeiten/Gesamtzahl Fälle) in das Fenster eintragen.

Cell	p(H0)	p(H1)
1	0.2908	0.3136
2	0.2987	0.2758
3	0.2025	0.1797
4	0.208	0.2309

Number of cells: 4

0.25 0.25

Equal p(H0) Equal p(H1)

Normalize p(H0) Normalize p(H1)

Auto calc last cell Auto calc last cell

Calculate Effect size w 0.09290627

Calculate and transfer to main window

Close

Quelle: <https://lehrbuch-psychologie.springer.com/content/zu-den-spss-r-und-gpower-aufgaben-und-ergaenzungen>

Aus: Rasch, Frieze, Hofmann & Naumann (2021). *Quantitative Methoden. Band 2*, 5. Auflage. Heidelberg: Springer.

© Springer-Verlag GmbH Deutschland, ein Teil von Springer Nature

Aufgaben mit R und G*Power

Rasch, Friese, Hofmann & Naumann (2021). *Quantitative Methoden. Band 2* (5. Auflage). Heidelberg: Springer.

Die von G*Power ermittelte Effektstärke von $w = 0,09$ erhalten Sie auch, wenn Sie w von Hand mit Hilfe des χ^2 -Wertes bestimmen: $\hat{w}^2 = \frac{\chi^2}{N} = \frac{10,964}{1269} = 0,009$; $w = 0,095$

Teststärke mit G*Power: Kopieren sie den ermittelten Wert für w in das Hauptfenster und vervollständigen Sie die anderen Werte gemäß der Angaben. G*Power ermittelt eine Teststärke Post hoc von 91%.

Test family		Statistical test		
χ^2 tests		Goodness-of-fit tests: Contingency tables		
Type of power analysis				
Post hoc: Compute achieved power - given α , sample size, and effect size				
Input Parameters		Output Parameters		
Determine =>	Effect size w	0.0929063	Noncentrality parameter λ	10.9534758
	α err prob	0.05	Critical χ^2	3.8414588
	Total sample size	1269	Power (1- β err prob)	0.9114343
	Df	1		

Stichprobenumfangsplanung: Minimaler Stichprobenumfang, wenn Teststärke wenigstens 80% betragen soll und $\alpha = 20\%$:

Test family		Statistical test		
χ^2 tests		Goodness-of-fit tests: Contingency tables		
Type of power analysis				
A priori: Compute required sample size - given α , power, and effect size				
Input Parameters		Output Parameters		
Determine =>	Effect size w	0.0929063	Noncentrality parameter λ	4.5056851
	α err prob	0.20	Critical χ^2	1.6423744
	Power (1- β err prob)	0.80	Total sample size	522
	Df	1	Actual power	0.8001882

Es sind mindestens 522 Teilnehmer vonnöten, wenn der Effekt bei einem α von 20% mit mindestens 80%iger Wahrscheinlichkeit entdeckt werden soll.

Quelle: <https://lehrbuch-psychologie.springer.com/content/zu-den-spss-r-und-gpower-aufgaben-und-ergaenzungen>

Aus: Rasch, Friese, Hofmann & Naumann (2021). *Quantitative Methoden. Band 2*, 5. Auflage. Heidelberg: Springer.

© Springer-Verlag GmbH Deutschland, ein Teil von Springer Nature