

Aufgaben zu Kapitel 8

Aufgabe 1

- Berechnen Sie einen U -Test für das in Kapitel 8.1 besprochene Beispiel mit verbundenen Rängen. Die entsprechende Testvariable „punkte2“ finden Sie im Datensatz „Rangdaten.sav“. Entsprechen die Ergebnisse denen im Buch?
- Berechnen Sie mit Hilfe eines Mann-Whitney U -Tests, ob es einen Unterschied in den mittleren Rängen zwischen den Bedingungen 2 und 3 auf der Variable „av“ gibt (zweiseitige Testung).
- Wie hoch ist in diesem Beispiel die Wahrscheinlichkeit einen tatsächlich vorhandenen Effekt der Größe $d = 0,5$ zu entdecken? Das Signifikanzniveau beträgt $\alpha = 0,05$, die Daten sind annähernd normalverteilt und die Testung erfolgt einseitig.
- Wie viele Schüler müsste der Lehrer untersuchen, um einen mittleren Effekt von $d = 0,5$ mit 80%iger Sicherheit bei einseitiger Testung und unter Annahme der Normalverteilung der Daten zu entdecken? Das Signifikanzniveau wird auf $\alpha = 0,05$ festgelegt.

Aufgabe 2

- Betrachten Sie sich das Beispiel des Entspannungstrainings in Kapitel 8.2. Der Therapeut befragt seine Patienten eine Woche nach der Übung ein weiteres Mal bzgl. ihrer Entspannung. Gibt es noch immer einen positiven Effekt der Übung (verglichen zum Ausgangszeitpunkt), auch nachdem eine Woche vergangen ist? Im Datensatz finden Sie die notwendige Variable „woche“.
- Wie hoch ist in diesem Beispiel die Wahrscheinlichkeit einen tatsächlich vorhandenen Effekt der Größe $d = 0,5$ zu entdecken? Das Signifikanzniveau beträgt $\alpha = 0,05$, die Daten sind annähernd normalverteilt und die Testung erfolgt einseitig.
- Wie viele Patienten müsste der Therapeut untersuchen, um einen Effekt von $d = 0,5$ mit 80%iger Sicherheit bei einseitiger Testung und unter Annahme der Normalverteilung der Daten zu entdecken? Das Signifikanzniveau wird auf $\alpha = 0,05$ festgelegt.

Aufgabe 3

- Bei einem Jonglage-Kurs gibt es drei Gruppen: Eine Gruppe bekommt nur die Bälle in die Hand gedrückt. Die Schüler dieser Gruppe sollen sich das jonglieren selber beibringen. Die zweite Gruppe bekommt schriftliche Unterlagen über die Technik des Jonglierens. Mit der dritten Gruppe übt ein erfahrener Jongleur und gibt Tipps. Nach einer Stunde kommen die Gruppen zusammen alle Schüler jonglieren ein Mal so lange, bis ein Ball auf dem Boden fällt. Lassen sich Unterschiede im Lernerfolg feststellen? Die Gruppenvariable heißt „bedingung“, die Testvariable „sekunden“.
- Wie hoch ist in diesem Beispiel die Wahrscheinlichkeit einen tatsächlich vorhandenen mittleren Effekt der Größe $f = 0,25$ zu entdecken? Das Signifikanzniveau beträgt $\alpha = 0,05$ und pro Bedingung wurden 20 Personen untersucht.

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

- c) Wie viele Personen müssten untersucht werden, um einen mittleren Effekt von $f = 0,25$ mit 95%iger Sicherheit zu entdecken? Das Signifikanzniveau wird auf $\alpha = 0,05$ festgelegt.

Lösungen

Aufgabe 1

- a) Verwenden Sie die Funktion `wilcox.test()` „sex“ ist die Gruppierungsvariable und „punkte2“ die abhängige Variable. Um eine gerichtete Hypothese zu testen, ergänzen Sie `alternative = „greater“`, da ein Mittelwertunterschied größer als Null erwartet wird.

```
wilcox.test(rang$punkte2 ~ rang$sex,  
            exact = FALSE,  
            correct = FALSE,  
            alternative = "greater")
```

Sie erhalten folgenden Output:

```
wilcoxon rank sum test  
  
data:  rang$punkte2 by rang$sex  
w = 61.5, p-value = 0.02767  
alternative hypothesis: true location shift is greater than 0
```

Mit der Funktion `qnorm()` können Sie nun den empirischen z -Wert ermitteln. Sie setzen dafür den empirischen p -Wert in die Klammern ein und erhalten den dazugehörigen empirischen z -Wert.

```
qnorm(0.02767)  
[1] -1.916197
```

Dieser Wert entspricht dem im Buch per Hand ermittelten z -Wert. Da die Hypothese einseitig formuliert ist und die Rangplatzunterschiede in der vorhergesagten Richtung vorliegen, ist dieses Ergebnis als hypothesenkonform signifikant zu interpretieren.

- b) Um nur Werte aus den Bedingungen 2 und 3 zu betrachten, müssen die entsprechenden Fälle ausgewählt werden. Verwenden Sie hierfür die Funktion `subset()` und führen Sie danach den U -Test durch.

```
aufgabe1b <- subset(rang,  
                   rang$bedingung == "Gruppe 2"  
                   | rang$bedingung == "Gruppe 3")  
wilcox.test(aufgabe1b$av ~ aufgabe1b$bedingung,  
            correct = FALSE,  
            exact = FALSE)
```

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.

Aufgaben mit R & G*Power

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

```
wilcoxon rank sum test
```

```
data: rang$punkte2 by rang$sex
```

```
w = 61.5, p-value = 0.02767
```

```
alternative hypothesis: true location shift is greater than 0
```

- c) Hier soll die Teststärke a posteriori berechnet werden. Öffnen Sie hierzu bitte G*Power und wählen Sie „Test family“: t-tests, „Statistical test“: Means: Wilcoxon-Mann-Whitney test (two groups), „Type of power analysis“ „Post hoc: Compute achieved power“.

The screenshot shows the G*Power interface for a posteriori power analysis. The 'Test family' is set to 't tests' and the 'Statistical test' is 'Means: Wilcoxon-Mann-Whitney test (two groups)'. The 'Type of power analysis' is 'Post hoc: Compute achieved power - given α , sample size, and effect size'. Under 'Input Parameters', 'Tail(s)' is 'One', 'Parent distribution' is 'Normal', 'Effect size d' is 0.5, ' α err prob' is 0.05, 'Sample size group 1' is 8, and 'Sample size group 2' is 10. Under 'Output Parameters', 'Noncentrality parameter δ ' is 1.0300645, 'Critical t' is 1.7516210, 'Df' is 15.1887339, and 'Power (1- β err prob)' is 0.2544778.

In unserem Beispiel liegt die Wahrscheinlichkeit einen tatsächlich vorhandenen Effekt zu entdecken bei gerade einmal 25,4%.

- d) Hier ist die Teststärke a priori, d.h. der benötigte Stichprobenumfang gefragt. Öffnen Sie G*Power und wählen „Test family“: t-tests, „Statistical test“: Means: Wilcoxon-Mann-Whitney test (two groups), „Type of power analysis“: A priori: Compute required sample size.

The screenshot shows the G*Power interface for an a priori power analysis. The 'Test family' is 't tests' and the 'Statistical test' is 'Means: Wilcoxon-Mann-Whitney test (two groups)'. The 'Type of power analysis' is 'A priori: Compute required sample size - given α , power, and effect size'. Under 'Input Parameters', 'Tail(s)' is 'One', 'Parent distribution' is 'Normal', 'Effect size d' is 0.5, ' α err prob' is 0.05, 'Power (1- β err prob)' is 0.80, and 'Allocation ratio N2/N1' is 1. Under 'Output Parameters', 'Noncentrality parameter δ ' is 2.5152354, 'Critical t' is 1.6603560, 'Df' is 99.2225438, 'Sample size group 1' is 53, 'Sample size group 2' is 53, 'Total sample size' is 106, and 'Actual power' is 0.8032180.

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

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

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

Aufgaben mit R & G*Power

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

Insgesamt müsste der Lehrer 106 Schüler – also je 53 Mädchen und 53 Jungs – untersuchen, um mit 80%iger Sicherheit einen mittleren Effekt zu entdecken.

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 & G*Power

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

Aufgabe 2

- a) Verwenden Sie die Funktion `wilcox.test()`, ergänzen Sie `paired = TRUE`, weil es sich um abhängige Daten handelt. Da es verbundene Ränge gibt, kann keine exakte Signifikanz berechnet werden, weshalb Sie `exact = FALSE` ergänzen müssen. Wir erwarten einen Unterschied kleiner als Null, weshalb zusätzlich `alternative = "less"` angegeben wird, damit Sie direkt überprüfen, ob der Unterschied in der erwarteten Richtung vorliegt. Hierbei ist jedoch auf die Reihenfolge zu achten, in der Sie die Variablen eingeben. Wenn Sie zuerst die Variable „woche“ und danach die Variable „vorher“ eingeben würden, müssten Sie `alternative = "greater"` eingeben.

```
wilcox.test(rang$vorher, rang$woche,  
           paired = TRUE,  
           correct = FALSE,  
           exact = FALSE,  
           alternative = "less")
```

wilcoxon signed rank test

data: rang\$vorher and rang\$woche

V = 1, p-value = 0.01262

alternative hypothesis: true location shift is less than 0

Der Unterschied ist signifikant, weshalb er in der vorhergesagten Richtung vorliegt. Die Richtung entspricht also der Hypothese: Auch nach einer Woche geben die Patienten an, sich entspannter zu fühlen als vor dem Training.

- b) Um die Teststärke a posteriori zu bestimmen, wählen Sie in G*Power „Test family“: t-tests und „Statistical test“: Means: Wilcoxon signed-rank test (matched pairs) und „Type of power analysis“ „Post hoc: Compute achieved power – given α , sample size, and effect size“ „Post hoc: Compute achieved power“.

The screenshot shows the G*Power software interface with the following settings:

- Test family:** t tests
- Statistical test:** Means: Wilcoxon signed-rank test (matched pairs)
- Type of power analysis:** Post hoc: Compute achieved power – given α , sample size, and effect size
- Input Parameters:**
 - Tail(s): One
 - Parent distribution: Normal
 - Effect size dz: 0.5
 - α err prob: 0.05
 - Total sample size: 9
- Output Parameters:**
 - Noncentrality parameter δ : 1.4658075
 - Critical t: 1.8725063
 - Df: 7.5943669
 - Power (1- β err prob): 0.3783428

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 & G*Power

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

Unter den Bedingungen in unserem Beispiel lag die Wahrscheinlichkeit den Effekt von $d = 0,5$ zu entdecken nur bei ca. 37,8%.

- c) Um die benötigte Stichprobengröße zu bestimmen wählen Sie in G*Power „Test family“: t-tests, „Statistical test“: Means: Wilcoxon-Mann-Whitney test (two groups Wilcoxon signed-rank test (matched pairs)), „Type of power analysis“: A priori: Compute required sample size.

Input Parameters		Output Parameters	
Tail(s)	One	Noncentrality parameter δ	2.5854415
Parent distribution	Normal	Critical t	1.7062592
Determine => Effect size dz	0.5	Df	25.7380304
α err prob	0.05	Total sample size	28
Power (1- β err prob)	.80	Actual power	0.8083058

Der Therapeut müsste 28 Personen untersuchen um einen Effekt der Stärke $d = 0,5$ mit 80%iger Sicherheit zu entdecken.

Aufgabe 3

- a) Geben Sie die beiden Variablen entsprechend in die Klammern der Funktion `kruskal.test()` ein.

```
kruskal.test(rang$sekunden ~ rang$bedingung)
```

```
Kruskal-wallis rank sum test
```

```
data: rang$sekunden by rang$bedingung
```

```
kruskal-wallis chi-squared = 9.7171, df = 2, p-value = 0.007762
```

Der Einfluss der Bedingung auf die Anzahl der jonglierten Sekunden ist signifikant.

Tipp: Wenn Sie die Mittelwerte für die abhängige Variable „sekunden“ getrennt für jede Gruppe einzeln berechnen wollen, können Sie die Funktion `aggregate()` verwenden.

```
aggregate(rang$sekunden, list(rang$bedingung), mean)
```

Sie erhalten folgenden Output, der die Unterschiede zwischen den Gruppen deutlich macht.

```
Group.1    x
1 Gruppe 1  7.2
2 Gruppe 2 10.4
3 Gruppe 3 20.4
```

Eine Überprüfung auf statistische Signifikanz ersetzt dieses Vorgehen natürlich nicht. In Fällen, in denen Sie die Rangvarianzanalyse mit ordinalskalierten Daten verwenden, ist dieses
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.

Aufgaben mit R & G*Power

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

Vorgehen gar nicht erst zulässig, denn die Berechnung von Mittelwerten erfordert mindestens Intervallskalenqualität (siehe Kapitel 1). In Fällen, in denen Sie die Rangvarianzanalyse einsetzen, weil die Stichprobengröße für ein parametrisches Verfahren zu klein ist, kann diese Analyse aber informativ sein.

- b) Um die Teststärke unter diesen Bedingungen zu bestimmen öffnen Sie G*Power und wählen Sie „Test family“: F- tests, „Statistical test“: ANOVA: Fixed Effects, omnibus, one-way. Sowie “Type of power analysis”: Post hoc: Compute achieved power.

The screenshot shows the G*Power interface with the following settings:

- Test family: F tests
- Statistical test: ANOVA: Fixed effects, omnibus, one-way
- Type of power analysis: Post hoc: Compute achieved power - given α , sample size, and effect size
- Input Parameters:
 - Determine =>
 - Effect size f: 0.25
 - α err prob: 0.05
 - Total sample size: 60
 - Number of groups: 3
- Output Parameters:
 - Noncentrality parameter λ : 3.7500000
 - Critical F: 3.1588427
 - Numerator df: 2
 - Denominator df: 57
 - Power ($1-\beta$ err prob): 0.3744311

Die Teststärke in diesem Beispiel liegt bei ca. 37,4%.

- c) Um zu Bestimmen wie groß die Stichprobe insgesamt sein müsste, um einen tatsächlich existierenden Effekt zu entdecken wählen Sie in G*Power „Test family“: F-tests, „Statistical test“: ANOVA: Fixed Effects, omnibus, one-way. Sowie “Type of power analysis”: A priori: Compute required sample size.

The screenshot shows the G*Power interface with the following settings:

- Test family: F tests
- Statistical test: ANOVA: Fixed effects, omnibus, one-way
- Type of power analysis: A priori: Compute required sample size - given α , power, and effect size
- Input Parameters:
 - Determine =>
 - Effect size f: 0.25
 - α err prob: 0.05
 - Power ($1-\beta$ err prob): 0.95
 - Number of groups: 3
- Output Parameters:
 - Noncentrality parameter λ : 15.7500000
 - Critical F: 3.0320649
 - Numerator df: 2
 - Denominator df: 249
 - Total sample size: 252
 - Actual power: 0.9514888

In diesem Beispiel müssen insgesamt 252 Personen – 84 pro Bedingung – untersucht werden, um mit 95%iger Sicherheit einen mittleren Effekt zu entdecken.

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

Literatur

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

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