

Übungsaufgaben zum selber rechnen - Statistik II

Auf den folgenden Seiten finden Sie Übungsaufgaben zu den in der Vorlesung Statistik II behandelten Themenbereichen (die Aufgaben sind entsprechend der Themenbereiche nummeriert).

- Wir empfehlen die Aufgaben zur **Nachbereitung** der Themenbereiche zu nutzen, **um die statistischen Konzepte und Verfahren einzuüben**.
- Die Aufgaben **sollen zu einem tieferen Verständnis beitragen und auf die Prüfung vorbereiten!** (Allerdings liegt - anders als in der Modulprüfung - den Aufgaben keine Formelsammlung bei!)
- Die Aufgaben zu den einzelnen Themenbereichen gliedern sich in
 1. Aufgaben, die händisch zu berechnen sind
 2. Aufgaben, die mit Hilfe des Statistikprogramms „R“ zu berechnen sind

Hinweise zur den angegebenen Lösungen der Aufgaben, die mit „R“ zu berechnen sind:

a) Die Lösungen setzen voraus, dass bekannt ist, wie Datensätze geladen (und gespeichert) werden. Einen Überblick wie verschiedene Datensätze geladen und gespeichert werden findet sich im Moodle.

[Statistik I & II -> Statistikprogramm R -> Lernmodule zur Einführung in R]

b) Die angegebenen Lösungen setzen voraus, dass der jeweilige Datensatz zuvor mit dem Befehl „attach“ aktiviert wurde. Dieser Befehl aktiviert ein Data Frame dauerhaft, so dass alle nachfolgend eingegebenen Befehle automatisch auf diesen angewendet werden. *Beispiel: attach(datensatzXY).*

c) Die angegebenen Lösungen stellen lediglich eine Lösungsmöglichkeit dar. Es ist durchaus möglich mit Hilfe einer anderen Funktion / eines anderen Pakets auf dieselbe Lösung zu kommen!

Die Themenbereiche gliedern sich anhand der Prüfungsliteratur wie folgt:

Themenbereich 1:

Arbeiten mit Verteilungen

Themenbereich 2:

Grundlagen der Statistik II

Themenbereich 3:

Vergleich von Mittelwerten

Themenbereich 4:

Vergleich von Varianzen

1. Arbeit mit Verteilungen

→ Benutzen Sie für die Aufgaben bitte die Tabellen 1, 2, 3, 5 und 8 im Anhang der Prüfungsliteratur

(S. 985ff im Buch)

1.1 Binomialverteilung

a) Bestimmen Sie für die binomialverteilte Zufallsgröße X die Wahrscheinlichkeit $P(X=1)$ für

- I. $n=10$ und $p=0.1$
- II. $n=30$ und $p=0.9$

b) Bestimmen Sie für die binomialverteilte Zufallsgröße X die Wahrscheinlichkeiten für

- I. $P(X \leq 4)$ und $P(X > 4)$ bei $n=20$ und $p=0.4$
- II. $P(X \leq 6)$ und $P(X > 6)$ bei $n=25$ und $p=0.7$

c) Bestimmen Sie

- I. $F_{10;0,40}(2)$
- II. $F_{30;0,8}(15)$

d) Aus einem 32-Blatt-Kartenspiel wird zufällig eine Karte gezogen (Hinweis: In einem Kartenspiel sind je 8 Karten einer Farbe im Spiel). Mit welcher Wahrscheinlichkeit zieht jemand bei 20 Durchgängen (Wiederholungen) eine bestimmte Anzahl Karten der Farbe Pik? Anmerkung: Es wird also jedes Mal aus einem kompletten 32er Blatt gezogen!

- I. genau 5 mal
- II. zwischen 5 und 10mal

e) In der Mensa werden Freitags ein Fischgericht und 2 weitere Menüs angeboten. Erfahrungsgemäß wählen 9 von 30 Gästen das Fischgericht. Die Küche bereitet 12 Fischgerichte vor. Wie groß ist die Wahrscheinlichkeit, dass diese 12 Gerichte für die zu erwartenden 20 Gäste nicht ausreichen?

1.2 Normalverteilung

a) Wie lautet der Wert der Normalverteilungsfunktion für

- I. $z=1.68$
- II. $z=-0.85$

b) Die Zufallsgröße X ist normalverteilt mit $\mu=120$ und $\sigma=10$. Berechnen Sie die Wahrscheinlichkeiten

- I. $P(X \leq 120)$
- II. $P(X > 125)$
- III. $P(110 \leq X \leq 130)$

c) Eine Zufallsgröße X ist normalverteilt mit $\mu=5$ und $\sigma=1.2$.

- I. Unterhalb welcher Zahl liegen die Werte von X mit 80%iger Wahrscheinlichkeit?
- II. Oberhalb welcher Zahl liegen die Werte von X mit 60%iger Wahrscheinlichkeit?
- III. Zwischen welchen Zahlen liegen die Werte von X mit 95%iger Wahrscheinlichkeit? 7

1.3 T-Verteilung

a) Wie lautet der Wert der t-Verteilungsfunktion für

- I. $df=8, p=0.60$

b) Eine Zufallsvariable sei t-verteilt mit 20 Freiheitsgraden. Bestimmen Sie

- I. $P(t > 2.5)$
- II. $P(t \leq -0.86)$
- III. $P(-1.325 \leq t \leq 1.725)$

c) Die tatsächlichen Ankunftszeiten eines Linienbusses seien um die fahrplanmäßige Ankunftszeit t -verteilt (d.h. der Bus kann zu früh, pünktlich oder zu spät am Zielbusbahnhof ankommen).

- I. Wie groß ist die Wahrscheinlichkeit, dass ein Bus bei 10 Freiheitsgraden mit einer Abweichung von maximal ± 1.37 Min. an einer Haltestelle eintrifft?
- II. In welchem Bereich bewegen sich bei 13 Freiheitsgraden 90% der (symmetrischen) Abweichungen von der fahrplanmäßigen Ankunftszeit?
- III. In welchem Bereich bewegen sich bei 20 Freiheitsgraden 95% der (symmetrischen) Abweichungen von der fahrplanmäßigen Ankunftszeit?

1.4 Chi-Quadrat-Verteilung

a) Wie lautet der Wert der χ^2 -Verteilungsfunktion für

- I. $df=8, p=0.10$
- II. $df=15, p=0.99$

b) Eine Zufallsvariable sei χ^2 -verteilt mit 20 Freiheitsgraden. Bestimmen Sie

- I. $P(\chi^2 > 9.5)$
- II. $P(\chi^2 \leq 12.5)$
- III. $P(31 \leq \chi^2 \leq 34)$

c) Die Verteilung einer Zufallsvariablen sei χ^2 -verteilt.

- I. Wie groß ist die Wahrscheinlichkeit, dass bei 30 Freiheitgraden der Chi-Quadrat-Wert ≤ 15 ist?
- II. Wie groß ist die Wahrscheinlichkeit, dass bei 15 Freiheitgraden der Chi-Quadrat-Wert > 8.5 ist?

1.5 F-Verteilung

a) Wie lautet der Wert der F-Verteilungsfunktion für

- I. $df_{\text{Zähler}} = 5, df_{\text{Nenner}} = 2, p=0.99$
- II. $df_{\text{Zähler}} = 100, df_{\text{Nenner}} = 5, p=0.05$

b) Eine Zufallsvariable sei F-verteilt mit 6 Zähler- und 60 Nennerfreiheitsgraden. Bestimmen Sie

- I. $P(F > 2.67)$
- II. $P(F \leq 3.12)$
- III. $P(2.67 \leq F \leq 3.12)$

c) Eine Zufallsvariable sei F-verteilt.

- I. Wie groß ist die Wahrscheinlichkeit, dass bei 10 Zähler- und 20 Nennerfreiheitsgraden der F-Wert ≤ 2 ist?
- II. Wie groß ist die Wahrscheinlichkeit, dass bei 40 Zähler- und 16 Nennerfreiheitsgraden der F-Wert ≥ 3 ist?

2. Grundlagen der Statistik II

1. Was besagt der Zentrale Grenzwertsatz?
2. Nennen und definieren Sie die 4 Qualitätskriterien von Kennwerten.
3. Was ist ein Konfidenzintervall ?
4. Was gibt der p-Wert an?

3. Vergleich von Mittelwerten

Aufgabe 3.1

Sie kaufen ein neues Auto, das laut Betriebsanleitung mit einer Tankfüllung (40l) im Schnitt 700 Kilometer fährt.

- a) Die 1. Tankfüllung ist nach 696 Kilometern verbraucht. Besteht Anlass zu glauben, dass die Angabe aus der Betriebsanleitung ein Fehler ist? Gehen Sie von einer zu erwartenden Streuung von $\sigma = 3$ (die auch in der Betriebsanleitung angegeben ist!) aus! Formulieren Sie vor Ihren Berechnungen eine Nullhypothese und eine Alternativhypothese. Die Sicherheitswahrscheinlichkeit soll $\alpha = .05$ betragen.
- b) Sie bleiben skeptisch und nehmen sich vor in Zukunft genau darauf zu achten, wieviele Kilometer ihre Tankfüllung jeweils hält. Bestimmen Sie den optimalen Stichprobenumfang zur Testung der Fragestellung, ob die Angabe aus der Betriebsanleitung ein Fehler ist, mit Hilfe der Angaben aus a) und Sicherheitswahrscheinlichkeit $\alpha = .0228$, β -Fehler = .1587 (bzw. Power = .8413). Betrachten Sie es als praktisch bedeutsam, wenn Sie mit einer Tankfüllung weniger als 696km weit kommen. Runden Sie Ihr Ergebnis auf eine ganze Zahl!
- c) Daraufhin führen Sie eine kontrollierte Studie mit 5 Durchgängen durch. Jede Tankfüllung fahren Sie bis auf den letzten Tropfen. Berechnen Sie anhand dieser Daten, ob Anlass besteht zu glauben, dass die Angabe aus der Anleitung ein Fehler ist. Gehen Sie dabei davon aus, dass die Streuung nicht bekannt ist und der Standardfehler aus den Daten geschätzt werden muss (runden Sie bei der Schätzung der Streuung, des Standardfehlers und bei der Bestimmung des t-Werts auf eine ganze Zahl)! Formulieren Sie vor Ihren Berechnungen eine Nullhypothese und eine Alternativhypothese. Die Sicherheitswahrscheinlichkeit soll $\alpha = .05$ betragen. Betrachten Sie es als praktisch bedeutsam, wenn Sie mit einer Tankfüllung weniger als 696km weit kommen. Hinweis: $\sqrt{5} = 2.24$, $\sqrt{6} = 2.45$, $\sqrt{7} = 2.65$

Hier die Ergebnisse der erreichten Kilometer

1	2	3	4	5
685	699	681	689	691

- d) Sie führen die Studie mit 100 Durchgängen erneut durch. Gehen Sie dabei von einem mittleren Literzahl von 685 und einer wahren Streuung von $\sigma = 4$ aus! Berechnen Sie anhand dieser Daten, ob Anlass besteht zu glauben, dass die Angabe aus der Anleitung ein Fehler ist. Formulieren Sie vor Ihren Berechnungen eine Nullhypothese und eine Alternativhypothese. Die Sicherheitswahrscheinlichkeit soll $\alpha = .05$ betragen. Betrachten Sie es als praktisch bedeutsam, wenn Sie mit einer Tankfüllung weniger als 696km weit kommen.

NUN ZU :

- e) Sie haben in einer Zeitung gelesen, dass Psychologiestudierende in ganz Deutschland im Durchschnitt schlecht über ihren Studiengang informiert sind (mit $\mu = 3,5$ und $\sigma =$ unbekannt).
[Gemessen auf einer Skala von 1-6, wobei 1=sehr gut informiert und 6=gar nicht informiert].
Diese Information überrascht Sie, da Sie das Gefühl haben, dass die Psychologiestudierenden Ihres Jahrgangs an der Universität Bielefeld sehr gut über den Studiengang informiert sind. Sie haben die Hypothese, dass Ihre Kohorte besser über den Studiengang informiert ist, als Psychologiestudierende deutschlandweit.

1. Formulieren Sie ihre Hypothesen (Null- und Alternativhypothese)
2. Lesen Sie die Datei „Erstis1516.sav“ in R ein.
3. Erstellen Sie nun einen neuen Datensatz, der nur Personen enthält, die die Variable „info_stu“ (=Informiertheit Studiengang Psychologie) ausgefüllt haben, d.h. Personen mit fehlenden Werten (hier definiert als 99) der Variable „info_stu“ (=Informiertheit Studiengang Psychologie) werden ausgeschlossen.
4. Welche Test-Statistik muss berechnet werden und welches Signifikanzniveau würden Sie wählen?
5. Berechnen Sie in R die passende Test-Statistik und formulieren Sie einen Ergebnissatz (gehen Sie von einem Signifikanzniveau von 5% aus).
Tipp: Die Hilfsfunktion `help(t.test)` liefert alle nötigen Informationen.

Aufgabe 3.2

Sie befragen im Rahmen Ihrer Bachelorarbeit Patienten in 2 psychosomatischen Kliniken in Bielefeld und legen den 100 Patienten (45 aus Klinik A, 55 aus Klinik B) u.a. einen Fragebogen zur Erfassung von Depressionen vor. Patienten beantworten 15 Items zur Depressivität und 10 Items zur sozialen Phobie und erhalten dann einen Gesamtscore für Depression, der zwischen 0 (keine Anzeichen für Depressivität) und 10 (ausgeprägte Depressivität) liegt und einen Gesamtscore für das Ausmaß der sozialen Phobie, der zwischen 0 (keine Anzeichen für soziale Phobie) und 10 (ausgeprägte soziale Phobie) liegt.

- a) Bei der Datenauswertung fällt Ihnen auf, dass Patienten der einen Klinik (A) im Mittel einen Depressionswert von 6 (gehen Sie von einer bekannten Populationsstreuung von $\sigma = 1$ aus) erreichen, während Patienten der anderen Klinik (B) einen mittleren Depressionswert von 8 (gehen Sie von einer bekannten Populationsstreuung von $\sigma = 1$ aus) erreichen. Kann davon ausgegangen werden, dass Patienten aus Klinik B depressiver sind als Patienten aus Klinik A? Formulieren Sie eine Nullhypothese und eine Alternativhypothese und führen Sie den T-Test für unabhängige Stichproben durch. Die Sicherheitswahrscheinlichkeit soll $\alpha = .05$ betragen.

Hinweis: $\sqrt{\frac{1}{16}} = \frac{1}{4} = 0.25$, $\sqrt{\frac{1}{24}} = 0.20$, $\sqrt{\frac{1}{32}} = 0.18$

- b) Bei der Datenauswertung fällt Ihnen außerdem auf, dass Patienten der einen Klinik (A) im Mittel einen mittleren Wert für soziale Phobie von 5 (gehen Sie diesmal von einer bekannten Populationsstreuung von $\sigma = 3$ aus) erreichen, während Patienten der anderen Klinik (B) einen mittleren Wert für soziale Phobie von 4.4 (gehen Sie diesmal von einer bekannten Populationsstreuung von $\sigma = 3$ aus) erreichen. Kann davon ausgegangen werden, dass Patienten aus Klinik B weniger sozial phobisch sind als Patienten aus Klinik A? Formulieren Sie eine Nullhypothese und eine Alternativhypothese und führen Sie den T-Test für unabhängige Stichproben durch. Die Sicherheitswahrscheinlichkeit soll $\alpha = .05$ betragen.

Hinweis: $\sqrt{\frac{3}{8}} = 0.6$, $\sqrt{\frac{5}{8}} = 0.8$, $\sqrt{\frac{7}{8}} = 0.9$

- c) Sie haben einen Fragebogen vorliegen, den Herr Schmidt ausgefüllt hat, es ist aber vergessen worden, die Klinik einzutragen, aus der Herr Schmidt kommt. Herr Schmidt erzielt einen Depressionswert von 6.5. Wenn man davon ausgeht, dass Patienten mit ähnlichen Depressionswerten aus den entsprechenden Kliniken (A vs. B) kommen, aus welcher Klinik kommt Herr Schmidt dann eher?

Aufgabe 3.3

Sie befragen im Rahmen Ihres Empiriepraktikums 80 Abiturienten einen Tag vor ihrer Klausur im Leistungskurs und einen Tag nach ihrer Klausur im Leistungskurs und erfassen u.a. aktuelle Lebenszufriedenheit (4 Items) und psychische Beanspruchung auf jeweils 8-stufigen Likertskalen von 1=gering bis 8=hoch. Die vorliegende Tabelle zeigt die mittleren Skalenwerte der Abiturienten zu den Messzeitpunkten 1 und 2.

	Lebenszufriedenheit T1	Lebenszufriedenheit T2	Psychische Beanspruchung T1	Psychische Beanspruchung T2
1	3	6	5	5
2	4	6	5	5
3	2	3.0	4	3
4	3	5	6	7

Verändern sich Lebenszufriedenheit und psychische Beanspruchung von T1 nach T2? Formulieren Sie jeweils eine Nullhypothese und eine Alternativhypothese und führen Sie den T-Test für abhängige Stichproben durch. Die Sicherheitswahrscheinlichkeit soll $\alpha = .05$ betragen. Hinweis: $\sqrt{\frac{1}{3}} = 0.6$, $\sqrt{\frac{2}{3}} = 0.8$, $\sqrt{\frac{1}{4}} = 0.5$ (jeweils auf eine Nachkommastelle gerundet!)

Aufgabe 3.4

NUN ZU :

- a) Im Datensatz „Erstis_vergleich“ finden Sie die Daten zur Studienwahlsicherheit der Psychologiestudierenden Ihrer Kohorte ("stuwasi1_1516") und der letzten Kohorte die zum WiSe 14/15 ihr Studium begonnen hat ("stuwasi1_1415").
 - b.1 Sie haben die Hypothese, dass sich die Studienwahlsicherheit zwischen den Kohorten unterscheidet. Formulieren Sie Null- und Alternativhypothese.
 - b.2 Welche Test-Statistik muss berechnet werden und welches Signifikanzniveau würden Sie wählen?
 - b.3 Berechnen Sie den T-Test in R. Formulieren Sie das Ergebnis.

4. Vergleich von Varianzen I

Aufgabe 4.1

Es sei bekannt, dass in der Population der Studierenden in Deutschland der mittlere IQ bei 110 liegt, wobei $\sigma = 10$.

Sie sprechen nun für eine Studie 26 Personen in der Bielefelder Innenstadt an und führen mit diesen einen IQ-Test durch. Sie erhalten in dieser Stichprobe zwar einen mittleren IQ von 110, allerdings auch $\hat{\sigma}_x = 7$. Besteht Anlass zu glauben, dass die Varianz in der Stichprobe sich von der Varianz in der Population unterscheidet? Formulieren Sie eine Nullhypothese und eine Alternativhypothese, geben Sie die Anzahl der Freiheitsgrade an und vergleichen Sie die Varianzen mit Hilfe des Einstichprobentests. Die Sicherheitswahrscheinlichkeit soll $\alpha = .05$ betragen.

Aufgabe 4.2

Bei der weiteren Auswertung der Testung aus Aufgabe 4, stellen Sie fest, dass in der Gruppe der von ihnen befragten Frauen $\sigma_1 = 6$ ($n=13$), während in der Gruppe der von ihnen befragten Männer $\sigma_2 = 10$ ($n=13$). Besteht Anlass zu glauben, dass die Varianz in der Frauen-Stichprobe sich von der Varianz in der Männer-Stichprobe unterscheidet? Formulieren Sie eine Nullhypothese und eine Alternativhypothese und vergleichen Sie die Varianzen mit Hilfe des Zweistichprobentests. Die Sicherheitswahrscheinlichkeit soll $\alpha = .05$ betragen.

4. Vergleich von Varianzen II

Aufgabe 4.3

I) Sie zeigen 6 Personen im Rahmen Ihres empirischen Projektseminars entweder einen lustigen Film („Mr. Bean macht Ferien“), einen traurigen Film („The Green Mile“) oder einen neutralen Film (eine Dokumentation über Tiere) und erfassen anschließend die Leistung der Teilnehmer/innen in einer Wortsuchaufgabe (innerhalb einer vorgegebenen Zeit sollen möglichst viele Wörter in einem "Buchstabensalat" gefunden werden). Die Anzahl gefundener Wörter der 6 Personen ist in der vorliegenden Tabelle abgebildet.

Lustig	Traurig	Neutral
1	4	10
3	4	8

a) Überprüfen Sie ob die Art des gezeigten Films einen Einfluss auf die Leistung in der Wortsuchaufgabe hat. Formulieren Sie eine Nullhypothese und eine Alternativhypothese. Die Sicherheitswahrscheinlichkeit soll $\alpha = .05$ betragen.

b) Sie wollen wissen, welche der drei Gruppen sich konkret unterscheiden. Berechnen Sie zu diesem Zweck Post-Hoc Vergleiche nach der Methode der „Least significant difference“ (LSD) und interpretieren Sie das Ergebnis. Hinweis bei händischer Berechnung: Runden Sie den kritischen t-Wert vor Ihren Berechnungen auf 2 Nachkommastellen! Runden Sie den errechneten LSD-Wert ebenfalls auf 2 Nachkommastellen!

II) Für Ihre Bachelorarbeit führen Sie eine Untersuchung in einem Kindergarten durch. Sie bilden 3 Gruppen, mit je 2 Kindern und führen mit ihnen das „Bobo-Doll“-Experiment durch. Hierfür zeigen sie den Kindern einen Film, in dem gezeigt wird, wie ein Mann gegenüber einer Puppe aggressive Verhaltensweisen zeigt. Sie variieren in den drei Gruppen das Ende des Films:

1. Gruppe (Belohnung): Der Mann wird am Ende des Films für sein Verhalten belohnt
2. Gruppe (Bestrafung): Der Mann wird am Ende des Films für sein Verhalten bestraft
3. Gruppe (Keine Konsequenz): Das Verhalten des Mannes wird nicht kommentiert

Anschließend beobachten Sie das Verhalten der Kinder und die Anzahl aggressiver Verhaltensweisen, die diese gegenüber einer Puppe zeigen.

Die Rohwerte, sowie die Bedingungsmitelwerte sind in der folgenden Matrix abgebildet:

Stufe j des Faktors		
Belohnung (j=1)	Bestrafung (j=2)	Keine Konsequenz (j=3)
6	3	3
6	1	5

Überprüfen Sie, ob die Art des gezeigten Films Einfluss auf das Verhalten der Kinder hat. Formulieren Sie eine Nullhypothese und eine Alternativhypothese. Die Sicherheitswahrscheinlichkeit soll $\alpha = .05$ betragen.

NUN ZU :

Ihr neu erworbenes Wissen zur Varianzanalyse wollen Sie nun auch an einem praktischen Beispiel in R anwenden. Sie schauen sich erneut den Datensatz „Erstis_T1_T2“ an und haben die Hypothese, dass Personen, die unterschiedlich viel neben dem Studium arbeiten (=BERUFL_T) sich hinsichtlich ihrer Studienwahrscheinlichkeit (STUWA_SI1) unterscheiden.

- a) Erstellen Sie zunächst einen Datensatz, bei dem die Missings in den Variablen „BERUFL_T“ und „STUWA_SI1“ ausgeschlossen werden.
- b) Überprüfen Sie daraufhin die Verteilungsvoraussetzungen für eine einfaktorielle Varianzanalyse.
- c) Berechnen Sie die Varianzanalyse in R und formulieren Sie ihr Ergebnis.
- d) Berechnen Sie abschließend noch eine Post-Hoc Analyse, um herauszufinden, welche Gruppenmittelwerte sich unterscheiden.

4. Vergleich von Varianzen III

Aufgabe 4.4

Sie erheben für Ihre Bachelorarbeit Daten in einer Klinik für posttraumatische Belastungsstörungen. Und befragen 3 Patienten/innen dieser Klinik vor Beginn der Therapie, am Tag ihrer Entlassung sowie zwei Monate später hinsichtlich der Anzahl der Flashbacks, die Patienten/innen am Tag haben. Die vorliegende Tabelle zeigt die Angaben der Personen.

Vor	Nach	Follow-Up
10	4	7
11	2	5
9	3	3

Sie möchten überprüfen, ob die Anzahl der Flashbacks im Verlauf der Studie (d.h. nach der Therapie und/oder im Follow-Up) signifikant geringer ist als vor der Therapie. Formulieren Sie eine Nullhypothese und eine Alternativhypothese. Die Sicherheitswahrscheinlichkeit soll $\alpha = .05$ betragen.

NUN ZU :

Sie haben für Ihre Bachelorarbeit bereits die Varianzanalyse mit Messwiederholung per Hand ausgerechnet, sind sich aber nicht sicher, ob Sie richtig gerechnet haben. Sie wollen Ihr Ergebnis mit ihrem Statistikprogramm R überprüfen.

- a) Laden Sie den Datensatz „Flashback.sav“ in R.
- b) Berechnen Sie nun die Varianzanalyse mit Messwiederholung und vergleichen Sie ihr Ergebnis mit ihrem händisch berechneten Ergebnis.
- c) Berechnen Sie abschließend noch eine Post-Hoc Analyse, um herauszufinden, welche Gruppenmittelwerte sich unterscheiden.

5. G-Power

Aufgabe 5.1

I) Sie stehen vor der Entscheidung, ein Thema für Ihr Empiriepraktikum auszuwählen. Sie wollen Ihr Empira gut abschließen, haben allerdings wenig Zeit, Versuchspersonen zu rekrutieren. Sie können zwischen zwei angebotenen Themen wählen:

Thema eins: Es soll untersucht werden, ob sich eine Intervention auf die Achtsamkeit von Patienten einer psychosomatischen Klinik auswirkt, wozu in der Klinik Personen für eine Experimentalgruppe (mit Intervention) und eine gleich große Kontrollgruppe (ohne Intervention) rekrutiert werden sollen. Personen erhalten im Falle der Teilnahme also entweder die Intervention oder keine Intervention und füllen einen Fragebogen zur Achtsamkeit aus. Frühere Studien mit gesunden Probanden zeigten eine Effektstärke der Intervention von $d = 0.5$.

Thema zwei: Es soll untersucht werden, ob sich drei Körpertherapien unterschiedlich auf die Selbstwahrnehmung von Patienten einer Klinik auswirken, wozu in der Klinik Personen für drei Experimentalgruppen rekrutiert werden sollen - Personen erhalten im Falle der Teilnahme also eine von drei Körpertherapien und füllen anschließend einen Fragebogen zur Selbstwahrnehmung aus. Frühere Studien mit gesunden Probanden zeigten eine Effektstärke von $f = 0.4$.

Sie wollen sich dem Thema anschließen, bei welchem Sie am wenigsten Versuchspersonen benötigen, um den vermuteten Effekt aufzudecken. Bestimmen Sie dazu mit Hilfe des Programms G-Power die jeweiligen optimalen Stichprobengrößen. Die Sicherheitswahrscheinlichkeit soll $\alpha = .05$ betragen, die gewünschte Power ($1 - \beta$) sollte 0.8 betragen.

Bei welchem Thema müssen Sie weniger Versuchspersonen rekrutieren, um den vermuteten Effekt aufzudecken?

II) Eine Freundin von Ihnen hat eine Untersuchung in einer Psychiatrie durchgeführt. Sie hatte die Hypothese, dass eine neu entwickelte Methode zu einem größeren Therapieerfolg führt, als die herkömmliche Therapiemethode und hat in ihrer Studie 20 Patienten mit dieser neuen Therapiemethode behandelt. Ihre Freundin fand nun eine Effektstärke von $d=0.5$, der berechnete t-test gegen einen festen Wert wurde bei einer Sicherheitswahrscheinlichkeit $\alpha = .05$ aber nicht signifikant.

- a) Sie haben die Vermutung, dass Ihre Freundin zu wenige Versuchspersonen untersucht hat, um einen Effekt der Größe $d = 0.5$ mit einer Sicherheitswahrscheinlichkeit von $\alpha = .05$ aufzudecken. Begründen Sie ihre Vermutung, indem Sie die optimale Stichprobengröße mit Hilfe von G-Power bestimmen. Die Power $1 - \beta$ soll erneut bei 0.8 liegen
- b) Ihre Freundin versteht nicht, wieso die Stichprobengröße einen Einfluss auf die Power hat. Erklären Sie ihr diesen Umstand.

6. Multiple Regression

1. Erläutern Sie die einzelnen Schritte der Dummy-Kodierung einer unabhängigen Variablen.
2. Die Regressionskoeffizienten haben bei einer Dummy-Kodierung eine besondere Bedeutung. Erläutern Sie die Bedeutung des Achsenabschnitts b_0 und der Regressionsgewichte b_j
3. Wie lautet die Gleichung der multiplen Regression mit Interaktionsterm (Moderation) für eine abhängige Variable Y und zwei unabhängige Variablen X_1 und X_2 ? Worin liegt der Unterschied zu der Gleichung einer multiplen Regression ohne Interaktionsterm?