Übungsaufgaben Kapitel 16
Lösungen

Prof. Dr. Torben Kuhlenkasper

Aufgabe 1

Im Sommersemester 2015 haben die Teilnehmer der Vorlesung *Grundlagen der Deskriptiven Statistik* einen Fragebogen ausgefüllt, bei dem sie u.a. Angaben zum Geschlecht und zur Anzahl der Geschwister machen sollten.

Die 224 weiblichen Vorlesungsteilnehmerinnen hatten durchschnittlich 1.31 Geschwister ($\overline{x}$).

Die 144 männlichen Vorlesungsteilnehmer hatten durchschnittlich 1.4 Geschwister ($\overline{y}$)

Überprüfen Sie mit einem **geeigneten** Test ($\alpha = 0.05$), ob es unterschiedlich viele Geschwister bei männlichen und weiblichen Studenten gibt. (Auf Grund der großen Stichprobe ist von Normalverteilung auszugehen)

(Rechenhilfen: $s_x^2 = 0.951$, $s_y^2 = 0.954$, $F_{143;223;0.95} = 1.28$)

Es handelt sich um zwei unverbundene Stichproben mit Normalverteilung, zweiseitiger Test.

$\Longrightarrow$ zunächst F-Test und dann entweder der t- oder Welch-Test.

1. $H_0 : \sigma_x^2 = \sigma_y^2$ vs. $H_1 : \sigma_x^2 \neq \sigma_y^2$

2. $F = \frac{s_y^2}{s_x^2}$ (weil $s_y^2 > s_x^2$)

3. $F \sim F_{143;223}$

4. Lehne H_0 ab, wenn $F > F_{143;223;0.95} = 1.28$

5. $F = \frac{0.954}{0.951} = 1.003$

6. Da $1.003 \not> 1.28$, lehne H_0 nicht ab.

   Es kommt also der klassische t-Test für unverbundene Stichproben zum Einsatz.

1. $H_0 : \mu_x = \mu_y$ vs. $H_1 : \mu_x \neq \mu_y$

2. $t = \frac{\overline{x} - \overline{y}}{\hat{\sigma} \cdot \sqrt{\frac{1}{m} + \frac{1}{n}}}$

3. $t \sim t_{m+n-2}$

4. Lehne H_0 ab, wenn $|t| > t_{144+224-2;0.975} = 1.96$

5. $\hat{\sigma^2} = \frac{143}{144+224-2} \cdot 0.954 + \frac{223}{144+224-2} \cdot 0.951 = 0.952$

   $\hat{\sigma} = \sqrt{0.952} = 0.976$

   $t = \frac{1.31 - 1.4}{0.976 \cdot \sqrt{\frac{1}{144} + \frac{1}{224}}} = -0.86$

6. Da $|-0.86| \not> 1.96$, lehne H_0 nicht ab.

   Es gibt keinen signifikanten Unterschied bei der Anzahl der Gesschwister zwischen Studentinnen und Studenten.

Aufgabe 2

Bei fünf Personen wurde der Hautwiderstand jeweils zweimal gemessen, einmal bei Tag (X) und einmal bei Nacht (Y). Man erhielt für das metrische Merkmal Hautwiderstand folgende Daten

X_i	24	28	21	27	23
Y_i	20	25	15	22	18

(a) Die Vermutung in Forscherkreisen geht dahin, daß der Hautwiderstand nachts absinkt. Läßt sich diese Vermutung durch die vorliegende Untersuchung erhärten? Testen Sie ohne Annahme der Normalverteilung mit $\alpha = 0.05$.

Es handelt sich um zwei verbundene Stichproben mit $D_i = X_i - Y_i$

1. Festlegung von H_0 und H_1

   $H_0 : M_D \leq 0$ vs. $H_1 : M_D > 0$
2. Festlegung der Teststatistik

   S als die Anzahl aller beobachteten Differenzen mit einem Wert größer als 0.
3. Bestimmung der Verteilung der Teststatistik unter H_0

   Diese ist unter H_0 binomialverteilt mit Parametern $n = 5$ und $p = 0.5$.
4. Aufstellung der Entscheidungsregel

   Lehne H_0 ab, falls $S \geq n - s_\alpha$
5. Berechnungen

   Es ergeben sich $d_1 = 4, d_2 = 3, d_3 = 6, d_4 = 5, d_5 = 5$

   und somit $S = 5$.

   Aus der Tabelle ergibt sich $s_{0.05} = 0$ (Spalte bei $n = 5$) und somit 5-0=5.

   Da $5 \geq 5$, lehne H_0 ab.
6. Antwort

   Das Absinken des Hautwiderstands ist signifikant zu $\alpha = 0.05$.

(b) Man überprüfe die Nullhypothese aus (a), wenn bekannt ist, daß der Hautwiderstand normalverteilt ist.

1. Festlegung von H_0 und H_1

   $H_0 : \mu_D \leq 0$ vs. $H_1 : \mu_D > 0$
2. Festlegung der Teststatistik

   $t = \frac{\sqrt{n}\bar{d}}{s}$
3. Bestimmung der Verteilung der Teststatistik unter H_0

   Diese ist unter H_0 t-verteilt mit n-1 Freiheitsgraden.

4. Aufstellung der Entscheidungsregel

   Lehne H_0 ab, falls $t > t_{n-1,1-\alpha}$

5. Berechnungen

   Es ergibt sich aus den Daten $\overline{d} = 4.6$ und $s = 1.14$

   und somit $t = \frac{\sqrt{5} \cdot 4.6}{1.14} = 9.023$

   Da $9.0233 > t_{4,0.95} = 2.1318$, lehne H_0 ab.

6. Antwort

   Auch bei Annahme der Normalverteilung ist das Absinken des Hautwiderstands signifikant zu $\alpha = 0.05$.

Aufgabe 3

Im Rahmen der PISA-Studie wurden die Leistungen in den Bereichen `Lesekompetenz` und `Mathematische Grundbildung` bestimmt. Tabelle 1 zeigt die Punkte von 10 Ländern.

Table 1: Punkte in den Bereichen `Lesekompetenz` und `Mathematische Grundbildung`

Land	Lesekompetenz	Mathematische Grundbildung
BR	396	334
D	484	490
F	505	517
I	487	457
FL	483	514
L	441	446
A	507	515
PL	479	470
P	470	454
E	493	476

Quelle: Deutsches PISA-Konsortium: PISA 2000

Testen Sie, ob die erwartete Punktezahl in beiden Bereichen sich unterscheidet. ($\alpha = 0.05$).

(a) Gehen Sie von Normalverteilung aus.

(b) Gehen Sie nicht von Normalverteilung aus.

(a) **verbundene zwei Stichproben, t-Test auf Grund der angenommenen Normalverteilung**

1. $H_0 : \mu_D = 0$ gegen $H_1 : \mu_D \neq 0$ mit $D = X - Y$

2. t-Test mit der Teststatistik $t = \frac{\sqrt{n} \cdot \bar{d}}{s_D}$

3. t ist unter H_0 t-verteilt mit $n - 1$ Freiheitsgraden.

   Tabellenwert für den zweiseitigen Test: $t_{9,0.975} = 2.3060$

4. Wir lehnen H_0 ab, wenn $|t| > t_{9,0.975}$

5. Es gilt $n = 10$ und $d_1 = 62$, $d_2 = -6$, $d_3 = -12$,$d_4 = 30$,$d_5 = -31$,$d_6 = -5$,$d_7 = -8$, $d_8 = 9$ $d_9 = 16$ und $d_{10} = 17$.

   Wir erhalten somit $\bar{d} = 7.2$ und damit $s_D^2 = \frac{1}{n-1} \sum_{i=1}^{n} (d_i - \bar{d}) = \ldots = 675.73$ sowie $s = 25.99$

   Wir erhalten für den Wert der Teststatistik $t = \frac{\sqrt{10} \cdot 7.2}{25.99} = 0.88$

6. Wegen $|0.88| \not> 2.3060$ lehnen wir H_0 nicht ab.

7. Wir können also nicht davon ausgehen, dass es eine Unterschied zwischen den unterschiedlichen Fächern bei den Ländern gibt.

(b) **verbundene zwei Stichproben, Vorzeichentest**

1. $H_0 : M_D = 0$ gegen $H_1 : M_D \neq 0$ mit $D = X - Y$

2. Vorzeichentest mit der Teststatistik $S =$ Anzahl positiver Differenzen

3. S ist unter H_0 Binomial-veretilt mit $n = 10$ und $p = 0.5$

4. Wir lehnen H_0 ab, falls $S \leq s_{0.025}$ or $S > n - s_{0.025}$.

   Aus der Tabelle mit der Binomialverteilung erhalten wir für $n = 10$ den Wert $s_{0.025} = 1$

5. Wir beobachten $S = 5$ positive Differenzen im Datensatz.

6. Da $5 \nleq 1$ oder $5 \ngeq 9$, lehnen wir H_0 nicht ab.

7. Wir können also auch hier nicht davon ausgehen, dass es eine Unterschied zwischen den unterschiedlichen Fächern bei den Ländern gibt.

Aufgabe 4

Im Rahmen der PISA-Studie wurde in den einzelnen Ländern die durchschnittliche Klassengröße ermittelt. Wir bilden zwei Klassen, wobei die Länder der ersten Klasse eine kleine und die Länder der zweiten Klasse eine hohe Klassengröße besitzen. Aus jeder Klasse wurden jeweils 4 Länder ausgewählt. Die Tabelle 2 zeigt die Ergebnisse im Bereich `Naturwissenschaftliche Grundbildung` dieser Länder.

Table 2: Ergebnisse im Bereich Naturwissenschaftliche Grundbildung

klein	hoch
496	487
481	513
476	532
500	499

Quelle: Deutsches PISA-Konsortium: PISA 2000

Überprüfen Sie, ob die erwartete Punktezahl in den beiden Gruppen unterschiedlich ist. ($\alpha = 0.05$). Gehen Sie von Normalverteilung aus.

zwei unverbundene Stichproben (zwei Gruppen von Ländern) mit Normalverteilung

Zunächst: F-Test zur Überprüfung der Varianzen.

1. $H_0 : \sigma_x^2 = \sigma_y^2$ gegen $H_1 : \sigma_x^2 \neq \sigma_y^2$

2. F-Test mit der Teststatistik $F = \frac{\sigma_x^2}{\sigma_y^2}$ oder $F = \frac{\sigma_y^2}{\sigma_x^2}$

3. F ist unter H_0 F-verteilt mit $n-1$ und $m-1$ Freiheitsgraden. Aus der Tabelle erhalten wir als kritischen Wert $F_{3,3,0.95} = 9.28$

4. Wir lehnen H_0 ab, wenn $F > F_{n-1,m-1,0.95}$

5. Es gilt hier $s_x^2 = \frac{1}{m-1} \sum_{i=1}^{n} (x_i - \bar{x}) = \ldots = 133.58$

und $s_y^2 = \frac{1}{n-1} \sum_{i=1}^{n} (y_i - \bar{y}) = \ldots = 374.25$

Somit erhalten wir $F = \frac{374.25}{133.58} = 2.8$

6. Da $2.8 \not> 9.28$, lehnen wir H_0 nicht ab.

7. Wir gehen also von gleichen Varianzen in den beiden Gruppen aus und verwenden nun den klassischen t-Test.

t-Test für zwei unverbundene Stichproben (mit gleichen Varianzen)

1. $H_0 : \mu_x = \mu_y$ gegen $H_1 : \mu_x \neq \mu_y$

2. t-Test mit der Teststatistik $t = \frac{\bar{x}-\bar{y}}{s \cdot \sqrt{\frac{1}{m}+\frac{1}{n}}}$

3. t ist unter H_0 t-verteilt mit $n + m - 2$ Freiheitsgraden.

   Wir erhalten aus der Tabelle als kritischen Wert: $t_{6,0.975} = 2.4469$

4. Wir lehnen H_0 ab, wenn $|t| > t_{6,0.975}$

5. Aus den Daten erhalten wir $\bar{x} = 488.25$ und $\bar{y} = 507.75$.

Wir nutzen außerdem (von den Berechnungen des F-Tests): $s^2 = \frac{3}{6} \cdot 133.58 + \frac{3}{6} \cdot 374.25 = 253.92$ und somit $s = 15.93$

Wir erhalten dann $t = \frac{488.25 - 507.75}{15.93 \cdot \sqrt{\frac{1}{4} + \frac{1}{4}}} = -1.73$

6. Da $|-1.73| \not> 2.4469$, lehnen wir H_0 nicht ab.

7. Wir können also nicht davon ausgehen, dass die beiden Gruppen von Ländern sich signfikant unterscheiden.