Serie 8

Aufgabe 8.1

Der skeptische Konsument gibt nicht auf und versucht weiterhin, den Weinhändler des Betrugs zu überführen. Der Weininhalt der 12 erworbenen Weinflaschen lauten:

$$71, 69, 67, 68, 73, 72, 71, 71, 68, 72, 69, 72 \quad \text{(in Zentiliter).}$$

a) Nun zweifeln wir daran, ob die Daten wirklich gut durch eine Normalverteilung beschrieben werden können (diese Annahme haben wir sowohl beim z- als auch beim t-Test gemacht). Wenn die Normalverteilungsannahme nicht gemacht werden kann, können wir den Vorzeichen-Test durchführen. Führen Sie also den Vorzeichen-Test auf dem 5 %-Signifikanzniveau durch. Wie lautet nun das Ergebnis? **Python**-*Hinweis:*

```
import scipy.stats as st
st.binom.cdf(k=..., n=..., p=...)
st.binom_test(x=..., n=..., p=..., alternative="...")
```

b) Wie lautet das Ergebnis mit dem Wilcoxon-Test? **Python**-*Hinweis:*

```
import scipy.stats as st
st.wilcoxon(...,  zero_method="wilcox", correction=True)
```

Aufgabe 8.2

Untenstehend finden Sie mehrere Beispiele für Vergleiche von 2 Stichproben. Beantworten Sie für jedes Beispiel *kurz* die folgenden Fragen:

- Handelt es sich um gepaarte oder um ungepaarte Stichproben? Begründen Sie!
- Ist der Test einseitig oder zweiseitig durchzuführen? Begründen Sie!
- Wie lautet die Nullhypothese in Worten?
- Wie lautet die Alternativhypothese in Worten?

a) In einem Experiment sollte der Effekt von Zigarettenrauchen auf Blutplättchenanhäufungen untersucht werden. Dazu wurden 11 Probanden vor und nach dem Rauchen einer Zigarette Blutproben entnommen, und es wurde gemessen, wie stark sich die Blutplättchen anhäuften. Es interessiert, ob sich Blutplättchen durch das Rauchen vermehrt anhäufen.

b) Die nächsten Daten sind aus einer Studie von Charles Darwin über die Fremd- und Selbstbefruchtung. 15 Paare von Setzlingen mit demselben Alter, je einer durch Selbst- und einer durch Fremdbefruchtung produziert, wurden gezüchtet. Beide Teile je eines Paares hatten nahezu gleiche Bedingungen. Das Ziel bestand darin zu sehen, ob die fremdbefruchteten Pflanzen mehr Lebenskraft besitzen als die selbstbefruchteten (d.h., ob sie grösser werden). Es wurden die Höhen jeder Pflanze nach einer fixen Zeitspanne gemessen.

c) Beeinflusst der Kalziumgehalt in der Nahrung den systolischen Blutdruck? Zur Überprüfung dieser Frage wurde einer Versuchsgruppe von 10 Männern während 12 Wochen ein Kalziumzusatz verabreicht. Einer Kontrollgruppe von 11 Männern gab man ein Placebopräparat.

d) In einem Experiment wurde untersucht, ob Mäuse zwei Formen von Eisen (Fe^{2+} und Fe^{3+}) unterschiedlich gut aufnehmen. Dazu wurden 36 Mäuse in zwei Gruppen zu je 18 unterteilt und die eine Gruppe mit Fe^{2+} und die andere mit Fe^{3+} „gefüttert". Da das Eisen radioaktiv markiert war, konnte sowohl die Anfangskonzentration wie auch die Konzentration einige Zeit später gemessen werden. Daraus wurde für jede Maus der Anteil des aufgenommenen Eisens berechnet.

Aufgabe 8.3

Zwei Tiefen-Messgeräte messen für die Tiefe einer Gesteins-Schicht an 9 verschiedenen Orten die folgenden Werte: Kennzahlen für die Differenz: $\overline{d}_n$ beträgt -5.78, die

Messgerät A	120	265	157	187	219	288	156	205	163
Messgerät B	127	281	160	185	220	298	167	203	171
Differenz d_i	-7	-16	-3	2	-1	-10	-11	2	-8

Standardabweichung $\sigma_D = 6.2$.

Es wird vermutet, dass Gerät B systematisch grössere Werte misst. Bestätigen die Messwerte diese Vermutung oder ist eine zufällige Schwankung als Erklärung plausibel?

a) Handelt es sich um verbundene (gepaarte) oder um unabhängige Stichproben?

b) Führen Sie einen t-Test auf dem Niveau $\alpha = 0.05$ durch. Formulieren Sie explizit: Modellannahmen, Nullhypothese, Alternative, Teststatistik, Verwerfungsbereich und Testergebnis.

c) Sei Z die Zufallsvariable, die zählt, bei wie vielen der 9 Messungen Gerät A einen grösseren Wert misst, als Gerät B. Wie ist Z verteilt, wenn die Geräte bis auf Zufallsschwankungen das Gleiche messen?

Aufgabe 8.4

In der folgenden Tabelle sind die Kieferlängen von 10 männlichen und 10 weiblichen Goldschakalen eingetragen: Einige Kennzahlen: $\bar{x}_n = 113.4$, $\bar{y}_n = 108.6$, $\hat{\sigma}_x^2 = 13.82$,

männlich x_i	120	107	110	116	114	111	113	117	114	112
weiblich y_j	110	111	107	108	110	105	107	106	111	111

$\hat{\sigma}_y^2 = 5.16$. Es stellt sich nun die Frage, ob es einen Unterschied zwischen der Kieferlänge von Männchen und Weibchen gibt.

a) Handelt es sich um gepaarte oder ungepaarte Stichproben? Begründen Sie Ihre Antwort.

b) Formulieren Sie Nullhypothese und Alternativhypothese.

c) Führen Sie den t-Test nun noch mit Hilfe von **Python** durch. Geben Sie den resultierenden p-Wert sowie den daraus folgenden Testentscheid an.

```
import pandas as pd
import scipy.stats as st

jackals  = pd.read_table(r"*/jackals.txt", sep = " ")

jackals

st.ttest_ind(jackals["M"], jackals["W"], equal_var=False)
```

Für $*$ muss Ihr Pfad stehen, wo sich die Datei befindet.

d) Führen Sie mit Hilfe von **Python** einen Wilcoxon-Test durch. Geben Sie wiederum p-Wert und Testentscheid an.

```
st.mannwhitneyu(jackals["M"], jackals["W"], alternative = "two-sided")
```

e) Falls die Resultate der beiden Tests unterschiedlich ausgefallen wären, welchem würden Sie eher vertrauen? Weshalb?

Aufgabe 8.5

Im Jahr 2013 wurden im Rahmen einer internationalen Zusammenarbeit unter der Leitung der EAWAG in Dübendorf Konzentrationen von illegalen Substanzen im Abwasser von 42 europäischen Städten während einer Woche untersucht (Ort C. et all, *Spatial differences and temporal changes in illicit drug use in Europe quantified by wastewater analysis*, Addiction 2014 Aug).

Dabei wurden an 7 aufeinanderfolgenden Tagen (6.-12. März) neben anderen Substanzen die medianen Konzentrationen von Ecstasy (MDMA) im Abwasser gemessenen. Aufgrund dieser Studie war eine Aussage einer vielgelesenen Schweizer Gratiszeitung, dass in Zürich viel mehr Drogen konsumiert werden als anderswo.

In der nachfolgenden Tabellle sind für die Städte Zürich und Basel die an den untersuchten Tagen ausgeschiedenen Mengen MDMA aufgeführt. Die Angaben sind in mg pro 1000 Einwohner pro Tag. Nehmen Sie an, dass die täglichen Differen-

Wochentage	Mi	Do	Fr	Sa	So	Mo	Di
Zürich	16.3	12.7	14.0	53.3	117	62.6	27.6
Basel	10.4	8.91	11.7	29.9	46.3	25.0	29.4

zen D_i zwischen den pro tausend Einwohner ausgeschiedenen Mengen von MDMA im Abwasser von Zürich und Basel unabhängig voneinander normalverteilt mit Erwartungswert μ_D und Standardabweichung σ_D sind.

a) Schätzen Sie aus den Daten den Mittelwert und die Standardabweichung der Differenzen, d.h., $\widehat{\mu}_D$ und $\widehat{\sigma}_D$.

b) Handelt es sich um gepaarte oder ungepaarte Stichproben? Begründen Sie Ihre Antwort.

c) Formulieren Sie die Nullhypothese und die Alternativhypothese, wenn Sie die Aussage der besagten Gratiszeitung überprüfen möchten.

d) Führen Sie einen statistischen Test mit Hilfe von **`Python`** auf dem Signifikanzniveau 95 % durch, unter der Annahme, dass die Daten normalverteilt sind.

   Wie lautet die Teststatistik und wie ist diese unter der Nullhypothese verteilt?

e) Führen Sie nun einen statistischen Test mit Hilfe von **`Python`** auf dem Signifikanzniveau 95 % durch, unter der Annahme, dass die Daten nicht normalverteilt sind.

Kurzlösungen vereinzelter Aufgaben

A 8.1: a) $K = [0, 2]$ b) p-Wert: 0.66

A 8.3: a) gepaart b) $K = (-\infty, -1.86]$

A 8.4:

a) ungepaart

b) $H_0 : \mu_1 = \mu_2$
$H_A : \mu_1 \neq \mu_2$

c) p-Wert: 0.0034

d) p-Wert: 0.0048

Musterlösungen zu Serie 8

Lösung 8.1

a) In dieser Teilaufgabe bezeichnet μ den Median der stetigen Zufallsvariablen X, also $P(X \leq \mu) = 0.5$. Um Hypothesen über den Median der Verteilung von X_i zu testen, verwendet man den Vorzeichentest.

1. *Modell*: $X_1, \cdots, X_{12}$ i.i.d. , wobei X_i eine beliebige Verteilung hat.
2. *Nullhypothese*:
$$H_0 : \ \mu = \mu_0 = 70$$
Alternative:
$$H_A : \ \mu < \mu_0$$
3. *Teststatistik*: V: Anzahl X_i's mit $(X_i > \mu_0)$ *Verteilung der Teststatistik unter* H_0:
$$V \sim \text{Bin}(12, 0.5)$$
4. *Signifikanzniveau*:
$$\alpha = 0.05$$
5. *Verwerfungsbereich für die Teststatistik*:
$$K = [0, c]$$
mit
$$c = \max\{v : \ P(V \leq v) \leq \alpha\}$$
Es gilt
$$P(V \leq 2) = 0.019$$
und
$$P(V \leq 3) = 0.073$$
Also $c = 2$. Diese Werte ermitteln sich mit Hilfe von **Python**:
```
import scipy.stats as st

st.binom.cdf(k=[2,3], n=12, p=0.5)

## [0.01928711 0.07299805]
```
6. *Testentscheid*: $v = 7; 7 \notin K; \rightarrow H_0$ beibehalten.

Oder

```
import scipy.stats as st

st.binom_test(x=7, n=12, p=0.5, alternative="less")

## 0.80615234375
```

Der P-Wert ist also 0.806. Der P-Wert ist deutlich über dem Signifikanzniveau, deshalb wird die Nullhypothese nicht verworfen. Alternativ kann der P-Wert auch wie folgt berechnet werden :

```
import scipy.stats as st

st.binom.cdf(k=7, n=12, p=0.5)

## 0.80615234375
```

b) Wir führen einen einseitigen Wilcoxon-Test mit Nullhypothese $\mu_0 = 70\,\text{cl}$ mit **Python** durch

```
import scipy.stats as st
from pandas import Series

x = Series([71, 69, 67, 68, 73, 72, 71, 71, 68, 72, 69, 72])

x = Series([71, 69, 67, 68, 73, 72, 71, 71, 68, 72, 69, 72])
st.wilcoxon(x-70, zero_method="wilcox", correction=True)
print(st.wilcoxon(x-70, zero_method="wilcox", correction=True))

## WilcoxonResult(statistic=33.5, pvalue=0.6902117434795202)
```

Da wir hier einen einseitigen Test haben, **Python** den P-Wert aber für eine zweiseitige Alternative berechnet und der beobachtete Mittelwert grösser als 70 ist, müssen wir den ausgegebenen Wert halbieren und von 1 abziehen. Der P-Wert ist also $1 - 0.69/2 = 0.66$. Der P-Wert beträgt also 0.66, woraus wir schliessen, dass die Nullhypothese belassen werden kann.

Lösung 8.2

a) *Gepaarte Stichprobe*: Zu jeder Blutplättchenmenge vor dem Rauchen gehört die Blutplättchenmenge derselben Person nach dem Rauchen.

Einseitiger Test: Wir wollen nicht wissen, ob sich die Bluplättchenmenge *verändert* hat, sondern ob sie sich *erhöht* hat.

H_0: Rauchen hat keinen Einfluss auf die Anhäufung der Blutplättchen: $\mu_R = \mu_{NR}$

H_A: Durch Rauchen erhöht sich die Anhäufung der Blutplättchen: $\mu_R > \mu_{NR}$

b) *Gepaarte Stichprobe*: Zu jeder Höhe eines selbstbefruchteten Setzlinge gehört die Höhe des fremdbefruchteten „Partners".

Einseitiger Test: Wir wollen nicht wissen, ob sich die Höhen *unterscheiden*, sondern ob die fremdbefruchteten Setzlinge *grösser* werden als die selbstbefruchteten.

H_0: Die Höhen unterscheiden sich nicht: $\mu_f = \mu_s$

H_A: Fremdbefruchtete Setzlinge werden grösser als selbstbefruchtete: $\mu_f > \mu_s$

c) *Ungepaarte Stichprobe*: Ungleiche Anzahl in den Gruppen. Zu einem Blutdruck aus der Versuchsgruppe gehört nicht ein spezifischer aus der Kontrollgruppe.

Zweiseitiger Test: Wir wollen nur wissen, ob das Kalzium einen Einfluss hat auf den Blutdruck, *egal* ob nach oben oder unten.

H_0: Kalzium hat keinen Einfluss auf den Blutdruck: $\mu_{\text{Kalz}} = \mu_{\text{Kontr}}$

H_A: Kalzium hat einen Einfluss auf den Blutdruck: $\mu_{\text{Kalz}} \neq \mu_{\text{Kontr}}$

d) *Ungepaarte Stichprobe*: Die Anzahlen in den beiden Gruppen brauchen nicht gleich zu sein. Zur Eisenmessung einer „Fe^{2+}-Maus" gehört nicht eine bestimmte Messung einer „Fe^{3+}-Maus".

Zweiseitiger Test: Wir wollen nur wissen, ob die Mäuse die verschiedenen Eisenformen *unterschiedlich* gut aufnehmen.

H_0: Die Eisenaufnahme ist von der Form unabhängig: $\mu_2 = \mu_3$

H_A: Die Eisenaufnahme ist von der Form abhängig :$\mu_2 \neq \mu_3$

Lösung 8.3

a) Es handelt sich um *gepaarte* Stichproben. Am gleichen Ort wird mit beiden Geräten gemessen.

b)
- *Modell:*

$$D_1, \ldots, D_n \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\mu, \sigma^2)$$

  σ wird durch $\hat{\sigma}$ geschätzt.

- *Nullhypothese:*

$$H_0: \ \mu_D = \mu_0 = 0$$

  Alternative:

$$H_A: \ \mu_D < \mu_0$$

- *Teststatistik:* t Verteilung mit Freiheitsgrad 8

- *Signifikanzniveau:*

$$\alpha = 5\,\%$$

- *Verwerfungsbereich:*

$$K = (-\infty, -3.843]$$

```
from scipy.stats import t
import numpy as np

t.ppf(q=0.05, loc=0, scale=6.2/np.sqrt(9), df=8)

## -3.8430659442138757
```

- *Testentscheid:*

  Der Wert

$$\overline{d}_n = -5.78$$

  liegt im Verwerfungsbereich, d.h. eine neue Eichung der Geräte ist angezeigt.

Mittel Standardisierung:

- *Modell:*

$$D_1, \ldots, D_n \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\mu, \sigma^2)$$

  σ wird durch $\hat{\sigma}$ geschätzt.

- *Nullhypothese:*

$$H_0: \; \mu_D = \mu_0 = 0$$

  Alternative:

$$H_A: \; \mu_D < \mu_0$$

- *Teststatistik:*

$$T = \frac{\sqrt{n}(\overline{D}_n - \mu_0)}{\hat{\sigma}}$$

  Verteilung der Teststatistik unter H_0:

$$T \sim t_{n-1}$$

- *Signifikanzniveau:*

$$\alpha = 5\,\%$$

- *Verwerfungsbereich:* (zu **R**)

$$K = (-\infty, t_{9-1,0.05}] = (-\infty, -1.86]$$

```
from scipy.stats import t
import numpy as np

t.ppf(q=0.05, df=8)

## -1.8595480375228428
```

- *Testentscheid:*

$$t = \frac{\overline{d}_n - 0}{\hat{\sigma}/\sqrt{9}} = -2.8$$

  Der Wert t der Teststatistik liegt im Verwerfungsbereich, d.h. eine neue Eichung der Geräte ist angezeigt.

c) Z wäre binomialverteilt mit Parametern $n = 9$ und $\pi = 0.5$ unter der Nullhypothese $\mu = 0$. Darauf aufbauend kann man auch einen Vorzeichentest durchführen. Der Vorteil ist, dass man keine Normalverteilung mehr annehmen muss.

Lösung 8.4

a) Es handelt sich um ungepaarte Stichproben, da zu den einzelnen Männchen nicht jeweils ein bestimmtes Weibchen gehört. Die Anzahlen in den beiden Stichproben brauchen auch gar nicht gleich gross zu sein.

b) Wir führen folgende Bezeichnungen ein:

   X_i: i-ter Wert der Kieferlänge der männlichen Tiere, $i = 1, \ldots, n = 10$

   Y_j: j-ter Wert der Kieferlänge der weiblichen Tiere, $j = 1, \ldots, m = 10$

   Nullhypothese H_0: X_i i.i.d. $\mathcal{N}(\mu_1, \sigma^2)$, Y_j i.i.d. $\mathcal{N}(\mu_2, \sigma^2)$ unabhängig, $\mu_1 = \mu_2$

   Alternative H_A: $X_i \sim (\mu_1, \sigma^2)$, $Y_j \sim \mathcal{N}(\mu_2, \sigma^2)$ mit $\mu_1 \neq \mu_2$

c) Der **Python**-Output für den t-Test sieht folgendermassen aus: (zu **R**)

```
import pandas as pd
import scipy.stats as st
jackals = pd.read_table("../../../Themen/Statistik_Messdaten/Uebungen_de/Da

st.ttest_ind(jackals["M"],jackals["W"], equal_var=False)

## Ttest_indResult(statistic=3.4843242131699643, pvalue=0.00335995243519250
```

Der p-Wert ist $0.0034 < 0.05$, also wird die Nullhypothese verworfen.

d) Der **Python**-Output für den Wilcoxon-Test sieht folgendermassen aus: (zu **R**)

```
import pandas as pd
import scipy.stats as st
jackals = pd.read_table("../../../Themen/Statistik_Messdaten/Uebungen_de/Da

st.mannwhitneyu(jackals["M"],jackals["W"], alternative="two-sided")

## MannwhitneyuResult(statistic=87.5, pvalue=0.004845462884722891)
```

Der P-Wert ist somit $0.0049 < 0.05$, also wird auch bei diesem Test die Nullhypothese verworfen.

e) Das Resultat des Wilcoxon-Tests ist vertrauenswürdiger, da er im Gegensatz zum t-Test nicht annimmt, dass die Daten normal-verteilt sind und wir diese Voraussetzung in keiner Weise überprüft haben. Allerdings ist die stark unterschiedliche Standardabweichung in den zwei Gruppen problematisch für beide Tests.

Lösung 8.5

a) Wir schätzen mit **Python** den Mittelwert und die Standardabweichung folgendermassen: (zu **R**)

```
import numpy as np
from pandas import Series

mdma_zuerich = Series([16.3, 12.7, 14.0, 53.3, 117, 62.6, 27.6])
mdma_basel = Series([10.4, 8.91, 11.7, 29.9, 46.3, 25.0, 29.4])

d = mdma_zuerich - mdma_basel

d.mean()
d.std()

## 20.27
## 26.272304175056032
```

Auf den ersten Blick scheint die Sachlage klar : in Zürich wird mehr MDMA konsumiert. Wir stellen dann allerdings fest, dass die Differenzen eine ziemlich grosse Streuung aufweisen.

b) Die Städte Zürich und Basel können als unterschiedliche Versuchseinheiten aufgefasst werden. Wir fassen die Stichproben also als ungepaart auf.

c) Die Nullhypothese lautet, dass es keinen Unterschied zwischen den beiden Städten in Bezug auf die ausgeschiedene Menge an MDMA gibt, also $\mu_D =$

$\mu_0 = 0$. Die Alternativhypothese entspricht der Behauptung der Gratiszeitung, nämlich dass in Zürich mehr Drogen konsumiert werden und damit mehr MDMA ausgeschieden wird, also $\mu_D > \mu_0 = 0$.

d) Da es sich um ungepaarte Stichproben handelt, gilt: (zu **R**)

```
import numpy as np
from pandas import Series
import scipy.stats as st

mdma_zuerich = Series([16.3, 12.7, 14.0, 53.3, 117, 62.6, 27.6])
mdma_basel = Series([10.4, 8.91, 11.7, 29.9, 46.3, 25.0, 29.4])

st.ttest_ind(mdma_zuerich,mdma_basel, equal_var=False)

## Ttest_indResult(statistic=1.3273296255450922, pvalue=0.2232566584677775)
```

In diesem Fall ist der zweiseitige P-Wert $pvalue = 0.22$, d.h., der einseitige P-Wert ist grösser als 0.05, womit die Nullhypothese beibehalten wird.

Würden wir die Daten als gepaarte Stichproben auffassen, dann haben wir für $\overline{D}_n \sim \mathcal{N}\left(\mu_D, \frac{\sigma_D^2}{n}\right)$ (Zentraler Grenzwertsatz). Dann ist die Teststatistik:

$$T = \frac{\overline{D}_n - \mu_0}{\widehat{\sigma}_D / \sqrt{n}} = \frac{\sqrt{n}\overline{D}_n}{\widehat{\sigma}_D}$$
$$t = 2.0413$$

Verteilung von T unter H_0: $T \sim t_{n-1} = t_6$. Mit **Python** ergibt sich : (zu **R**)

```
import numpy as np
from pandas import Series
import scipy.stats as st

mdma_zuerich = Series([16.3, 12.7, 14.0, 53.3, 117, 62.6, 27.6])
mdma_basel = Series([10.4, 8.91, 11.7, 29.9, 46.3, 25.0, 29.4])

st.ttest_rel(mdma_zuerich,mdma_basel)

## Ttest_relResult(statistic=2.041289516059924, pvalue=0.08728840175305683)
```

Da wir hier einen einseitigen Test haben und **Python** den P-Wert für eine zweiseitige Alternative berechnet, müssen wir den Wert für **pvalue** halbieren und somit ist der P-Wert 0.04364. Er ist (knapp) kleiner als $\alpha = 0.05$ und somit wird auf dem 5 % Signifikanzniveau die Nullhypothese verworfen. Falls die Daten also als geppaarte Stichproben aufgefasst würden, lautet die Schlussfolgerung, dass signifikant mehr Ecstasy in Zürich als in Basel konsumiert wird.

e) (zu **R**)

```
import numpy as np
from pandas import Series
import scipy.stats as st

mdma_zuerich = Series([16.3, 12.7, 14.0, 53.3, 117, 62.6, 27.6])
mdma_basel = Series([10.4, 8.91, 11.7, 29.9, 46.3, 25.0, 29.4])

st.mannwhitneyu(mdma_zuerich, mdma_basel, alternative='greater')

## MannwhitneyuResult(statistic=34.0, pvalue=0.12507650397969117)
```

Für den Mann-Whitney-U-Test kann in **Python** die Richtung der Alternativhypothese definiert werden, und zwar mit **alternative='...'**. Wir können folglich den Wert von **pvalue** direkt für den Testentscheid herbeiziehen: Der (einseitige) P-Wert beträgt in diesem Fall 0.13 und ist somit grösser als das Signifikanzniveau 5 %. Wir behalten die Nullhypothese also bei. Offenbar hängt der Testentscheid stark von der Verteilungsannahme der Differenzen ab.

R-Code

Aufgabe 8.1

a)
```
pbinom(c(2, 3), 12, p = 0.5)

## [1] 0.01928711 0.07299805
```

```
binom.test(7, 12, p = 0.5, alternative = "two.sided")

##
##  Exact binomial test
##
## data:  7 and 12
## number of successes = 7, number of trials =
## 12, p-value = 0.7744
## alternative hypothesis: true probability of success is not equal to 0.5
## 95 percent confidence interval:
##  0.2766697 0.8483478
## sample estimates:
## probability of success
##              0.5833333
```

b)
```
wilcox.test(c(71, 69, 67, 68, 73, 72, 71, 71, 68, 72,
    69, 72), mu = 70, alternative = "less")

##
##  Wilcoxon signed rank test with continuity
##  correction
##
## data:  c(71, 69, 67, 68, 73, 72, 71, 71, 68, 72, 69, 72)
## V = 44.5, p-value = 0.6838
## alternative hypothesis: true location is less than 70
```

Aufgabe 8.3

b) (zu **Python**)
```
qt(0.05, df = 8)

## [1] -1.859548
```

Aufgabe 8.4

a)

b) (zu **Python**)

```
jackals <- read.table(file = "./Daten/jackals.txt",
    header = TRUE)
t.test(jackals[, "M"], jackals[, "W"])

##
##  Welch Two Sample t-test
##
## data:  jackals[, "M"] and jackals[, "W"]
## t = 3.4843, df = 14.894, p-value = 0.00336
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  1.861895 7.738105
## sample estimates:
## mean of x mean of y
##     113.4     108.6
```

c) (zu **Python**)

```
wilcox.test(jackals[, "M"], jackals[, "W"])

##
##  Wilcoxon rank sum test with continuity
##  correction
##
## data:  jackals[, "M"] and jackals[, "W"]
## W = 87.5, p-value = 0.004845
## alternative hypothesis: true location shift is not equal to 0
```

Der Wert ist hier leicht grösser als bei Python. Der Grund liegt an *Stetigkeitskorrekturen*, die jeweils leicht anders gehandhabt werden. Die Grössenordnung bleibt allerdings gleich.

Aufgabe 8.5

a) (zu **Python**)

```
mdma_zuerich <- c(16.3, 12.7, 14, 53.3, 117, 62.6,
    27.6)
mdma_basel <- c(10.4, 8.91, 11.7, 29.9, 46.3, 25, 29.4)
d <- mdma_zuerich - mdma_basel
mean(d)

## [1] 20.27

sd(d)
```

```
## [1] 26.2723
```

d) (zu **Python**)

```
t.test(mdma_zuerich, mdma_basel, alternative = "greater",
    paired = TRUE)

##
##  Paired t-test
##
## data:  mdma_zuerich and mdma_basel
## t = 2.0413, df = 6, p-value = 0.04364
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  0.9742245       Inf
## sample estimates:
## mean of the differences
##                   20.27
```

(zu **Python**)

```
t.test(mdma_zuerich, mdma_basel, alternative = "greater",
    paired = FALSE)

##
##  Welch Two Sample t-test
##
## data:  mdma_zuerich and mdma_basel
## t = 1.3273, df = 7.5245, p-value = 0.1116
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  -8.362064       Inf
## sample estimates:
## mean of x mean of y
##  43.35714  23.08714
```

f) (zu **Python**)

```
wilcox.test(mdma_zuerich, mdma_basel, alternative = "greater",
    paired = FALSE)

##
##  Wilcoxon rank sum test
##
## data:  mdma_zuerich and mdma_basel
## W = 34, p-value = 0.1297
## alternative hypothesis: true location shift is greater than 0
```