Stochastik

Serie 11

Aufgabe 11.1

Ein Weinhändler behauptet, dass die von ihm gefüllten Weinflaschen mindestens 70 Zentiliter enthalten. Ein skeptischer Konsument vermutet aber, dass der Weinhändler zu wenig Wein abfüllt und will diese Behauptung überprüfen. Deshalb kauft er 12 Weinflaschen und misst ihren Inhalt. Er findet:

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

Die Standardabweichung der Abfüllungen ist nicht bekannt. Man muss sie also aus den gemachten Stichproben schätzen:

$$\hat{\sigma}_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 \approx 1.96^2$$

Da die Standardabweichung geschätzt wurde und nicht mehr exakt bekannt ist, kann der z-Test nicht durchgeführt werden. Verwenden Sie nun den t-Test auf dem 5 %-Signifikanzniveau. Geben Sie die Modellannahmen, H_0, H_A, den Verwerfungsbereich, den Wert der Teststatistik und das Testergebnis explizit an. Was ändert sich an obigem Test?

Aufgabe 11.2

Unterhalb einer Kläranlage wurden 16 unabhängige Wasserproben aus einem Fluss entnommen und jeweils deren Ammoniumkonzentration X_i (in µg $\mathrm{NH_4 - N/l}$) mit einem Messgerät bestimmt. Der Mittelwert der Proben ergab $\bar{x}_{16} = 204.2$.

Wir wollen nun wissen, ob mit diesem Experiment eine Überschreitung des Grenzwerts von 200 µg $\mathrm{NH_4 - N/l}$ nachgewiesen werden kann (auf dem 5 % Niveau).

a) Nehmen Sie an, die Standardabweichung der Messungen sei im Voraus aufgrund früherer Studien bekannt. Sie betrage 10 µg $\mathrm{NH_4 - N/l}$.

   Führen Sie unter dieser Annahme einen z-Test durch, um zu prüfen, ob eine Grenzwertüberschreitung nachgewiesen werden kann.

   Geben Sie die Modellannahmen, H_0, H_A, den Verwerfungsbereich, den Wert der Teststatistik und das Testergebnis explizit an.

b) Wie wahrscheinlich ist es, dass man mit 16 unabhängigen Wasserproben eine Grenzwertüberschreitung nachweisen kann, wenn die wahre Ammoniumkonzentration tatsächlich über dem Grenzwert, und zwar bei 205 µg $NH_4 - N/l$ liegt?

c) Wie wahrscheinlich ist es, dass man mit 16 unabhängigen Wasserproben fälschlicherweise eine Grenzwertüberschreitung nachweist, obwohl die wahre Ammoniumkonzentration bei 200 µg $NH_4 - N/l$ liegt und den Grenzwert somit genau einhält?

d) Nehmen Sie an, dass die Standardabweichung von 10 µg/l aus den 16 Proben geschätzt worden ist. Deshalb ist nun ein t-Test (zur Nullhypothese $\mu_0 = 200\,\mu g/l$) und nicht ein z-Test angebracht. Führen Sie den t-Test durch.

e) Welche Annahmen des t-Testes könnte verletzt sein und dazu führen, dass der t-Test schlechte Macht hat?

Aufgabe 11.3

Aufgrund langjähriger Untersuchungen ist bekannt, dass der Bleigehalt X von gewissen Bodenproben annähernd normal-verteilt ist

$$X \sim \mathcal{N}(\mu, \sigma^2)\,.$$

a) Es wurden in 10 Bodenproben der Bleigehalt X gemessen. Dabei wurde ein Mittelwert von $\overline{x}_{10} = 31$ ppb erhalten. Die Standardabweichung sei bekannt und beträgt 6 ppb. Geben Sie das zweiseitige 99 % Vertrauensintervall für den Mittelwert an.

b) Wie viele Beobachtungen sind nötig, um die Breite des in Teilaufgabe a) bestimmten zweiseitigen Vertrauensintervalls auf die Hälfte zu reduzieren? Wie viele (unabhängige) Bestimmungen des Bleigehalts müssen geplant werden, wenn der Bleigehalt mit einer Stichprobe „auf 1 ppb genau“ bestimmt werden soll, d. h., wenn die Breite des 99 % des Konfidenzintervalls nicht grösser als 1 ppb sein soll?

c) Normalerweise ist die Standardabweichung σ unbekannt. Um welchen Faktor verändert sich die Breite des zweiseitigen Vertrauensintervalls in Teilaufgabe a), wenn man die Standardabweichung aus den Daten geschätzt hat, also $\widehat{\sigma} = 6$?

Aufgabe 11.4

Im National Bureau of Standards (USA) wurden regelmässig Wägungen des 10-Gramm-Standardgewichtstücks durchgeführt. Bei 9 Wägungen erhielt man als durchschnittliche Differenz -403 Mikrogramm vom 10 Gramm-Sollgewicht und eine Standardabweichung von 3.127 Mikrogramm für eine einzelne Wägung.

a) Geben Sie das exakte, zweiseitige 95 %-Vertrauensintervall für die wahre Differenz an, unter der Annahme, dass die Messfehler normal-verteilt sind.

b) Könnte die wahre Differenz $-400.0\,\mu$g betragen? Entscheiden Sie aufgrund des Resultats in Aufgabe a). (Kurze Begründung)

Aufgabe 11.5

Eine Brücke soll aufgrund des höheren Verkehrsaufkommens renoviert werden. Im Bau wurden damals Schrauben mit einer mittleren Festigkeit von $500\,\text{N/mm}^2$ benutzt. Da dies für nicht mehr sicher genug gehalten wird, sollen diese nun durch Schrauben mit einer mittleren Festigkeit von mehr als $500\,\text{N/mm}^2$ ersetzt werden. Um diesen Anforderungen gerecht zu werden, hat der alte Schraubenlieferant ein neues Verfahren entwickelt. Zur Baustelle werden allerdings unbeschriftete Schrauben geliefert, aus denen nicht sofort hervorgeht, ob es sich um die alten 500er oder um die neuen verbesserten Schrauben handelt. Vor dem Verbau will der leitende Ingenieur zuerst sicherstellen, dass es sich um die besseren Schrauben handelt. Um dies herauszufinden, werden einige der Schrauben vermessen und ein statistischer Test durchgeführt. Je nach Ergebnis sollen die Schrauben verbaut oder zurückgeschickt werden.

	1	2	3	4	5
Schraubenfestigkeit (X_i)	520	512	499	524	505

Für den empirischen Mittelwert und empirische Varianz ergeben sich bei obiger Stichprobe $\bar{x}_5 = 512$ und $s_x^2 = 106.5$.

Wir modellieren die Daten mit einer Normalverteilung, d. h. X_i i.i.d. $\sim \mathcal{N}(\mu, \sigma^2)$.

a) Stellen Sie die geeigneten Null- und Alternativhypothesen auf und begründen Sie Ihre Wahl.

b) Sie führen nun einen einseitigen t-Test zum Niveau $\alpha = 0.05$ durch (unabhängig von Ihrer obigen Antwort).

Stellen Sie die Teststatistik T auf und berechnen Sie deren Wert. Geben Sie die Verteilung der Teststatistik T unter H_0 und den Verwerfungsbereich des Tests an. Was ist der Testentscheid?

c) Berechnen Sie ein (zweiseitiges) 95 %-Vertrauensintervall für μ. Wie würde das entsprechende Vertrauensintervall aussehen, wenn wir die Streuung als bekannt voraussetzen würden (mit dem gleichen Wert wie der beobachtete)?

d) Betrachten Sie (unabhängig von dem oben aufgeführten Beispiel) einen einseitigen t-Test von $H_0 : \mu = 0$ gegen $H_A : \mu > 0$ zum Niveau 0.05. Obwohl die beobachteten n Datenpunkte einen empirischen Mittelwert grösser Null haben, ergeben die Berechnungen, dass die Nullhypothese nicht verworfen wird. Entscheiden Sie, ob folgende Aussagen wahr oder falsch sind.

	Richtig	**Falsch**
Man verwirft H_0 für kein Niveau $\alpha < 0.05$	◯	◯
Es gibt ein Niveau $\alpha < 1$, bei dem man H_0 verwirft.	◯	◯
Der p-Wert ist strikt kleiner als 0.5.	◯	◯
Führt man statt eines einseitigen einen zweiseitigen Test zum Niveau 0.05 durch, verwirft man H_0 nicht.	◯	◯
Wenn man die Daten immer öfter kopiert (d. h., man betrachtet jeden Datenpunkt k-Mal, so dass man insgesamt $k \cdot n$ Datenpunkte erhält), verwirft man H_0 für ein grosses k beim Niveau 0.05.	◯	◯

Aufgabe 11.6

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?

b) Wie lautet das Ergebnis mit dem Wilcoxon-Test?

R-Hinweis:

```
wilcox.test(x, mu = ...)
```

Aufgabe 11.7

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 11.8

Zwei Tiefen-Messgeräte messen für die Tiefe einer Gesteins-Schicht an 9 verschiedenen Orten die folgenden Werte:

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

Kennzahlen für die Differenz: $\bar{d}_n$ beträgt -5.78, die 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 11.9

In der folgenden Tabelle sind die Kieferlängen von 10 männlichen und 10 weiblichen Goldschakalen eingetragen:

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

Einige Kennzahlen: $\bar{x}_n = 113.4$, $\bar{y}_n = 108.6$, $\hat{\sigma}_x^2 = 13.82$, $\hat{\sigma}_y^2 = 5.16$

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 R durch. Geben Sie den resultierenden p-Wert sowie den daraus folgenden Testentscheid an.

```
jackals <- read.table(./Daten/jackals.txt, header = TRUE)  # Datensatz einlesen
jackals  # Datensatz anschauen
t.test(jackals[, "M"], jackals[, "W"])  # t-Test durchfuehren
```

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

```
# Wilcoxon-Test durchfuehren
wilcox.test(jackals[, "M"], jackals[, "W"], )
```

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

Aufgabe 11.10

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 augeführt - die Werte finden Sie in der Datei *mdma.txt* im Verzeichnis **Daten** auf dem Desktop. Die Angaben sind in mg pro 1000 Einwohner pro Tag.

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

Nehmen Sie an, dass die täglichen Differenzen 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 **R** auf dem Signifikanzniveau 5% durch, unter der Annahme, dass die Daten normalverteilt sind. Wie lautet die Teststatistik und wie ist diese unter der Nullhypothese verteilt?

e) Geben Sie das (einseitiges) 95%-Vertrauensintervall für die Differenzen D_i an (mit Hilfe von **R**). Wie interpretieren Sie dieses Vertrauensintervall?

f) Führen Sie nun einen statistischen Test mit Hilfe von **R** auf dem Signifikanzniveau 5% durch, unter der Annahme, dass die Daten nicht normalverteilt sind.

Kurzlösungen vereinzelter Aufgaben

A 11.2:

a) $K =]1.64, \infty[$

b) $P[\overline{X}_n > 204.1] = 0.6406$

d) $K =]204.38, \infty[$

A 11.3:

a) $[26.1, 35.9]$

b) 959

c) $[24.8, 37.2]$

A 11.4:

a) $[-405.4, -400.6]$

A 11.6: a) $K = [0, 2]$ b) p-Wert: 0.6838

A 11.8: a) gepaart b) $K = (-\infty, -1.86]$

A 11.9: a) ungepaart c) p-Wert: 0.0034

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

d) p-Wert: 0.0048