

Übung 9

1. 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 anhäufeten. 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.

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 x_i	-7	-16	-3	2	-1	-10	-11	2	-8

Kennzahlen für die Differenz: \bar{x} beträgt -5.78, die Standardabweichung $sd_x = 6.2$.

Es soll die Frage untersucht werden: Ist eine neue Eichung der Geräte angezeigt oder messen sie noch, bis auf Zufallsschwankungen, das Gleiche?

- a) Handelt es sich um verbundene (gepaarte) oder um unabhängige Stichproben?
- b) Führe einen t-Test auf dem Niveau $\alpha = 0.05$ durch. Formuliere explizit: Modellannahmen, Nullhypothese, Alternative, Teststatistik, Verwerfungsbereich und Testergebnis.
Hinweis: Eine Tabelle mit den Perzentilen der t -Verteilung befindet sich am Ende der Übung 8.
- 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? Welcher Test nützt dieses Faktum aus?

3. 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: $\bar{x} = 113.4$, $\bar{y} = 108.6$, $s_x^2 = 13.82$, $s_y^2 = 5.16$

- a) Handelt es sich um gepaarte oder ungepaarte Stichproben? Begründe!
- b) Unterscheiden sich die Kieferlängen von Männchen und Weibchen signifikant? Führe von Hand einen vollständigen Zwei-Stichproben t-Test durch. Gib Null- und Alternativhypothese, Teststatistik, kritischen Wert resp. Verwerfungsbereich sowie Testentscheid explizit an! Muss der Test ein- oder zweiseitig durchgeführt werden?
Hinweis: Eine Tabelle mit den Perzentilen der t -Verteilung befindet sich am Ende der Übung 8.
- c) Führe den t-Test nun noch mit Hilfe von **R** durch. Gib den resultierenden p-Wert sowie den daraus folgenden Testentscheid an.

```
> jackals <- read.table("http://stat.ethz.ch/Teaching/Datasets/jackals.dat",
+                       header=TRUE) # Datensatz einlesen
> jackals # Datensatz anschauen
> t.test(jackals[, "M"], jackals[, "W"]) # t-Test durchfuehren
```

Bemerkung: Die Anzahl Freiheitsgrade, die im **R**-Output angegeben werden, ist anders, als die von Dir in b) verwendete. Dies liegt daran, dass **R** nicht davon ausgeht, dass die Varianzen in beiden Gruppen gleich gross sind und deshalb der t-Test leicht anders gerechnet wird als in der Vorlesung gelernt. Wenn **R** dasselbe machen soll wie Du von Hand, dann muss noch ein Argument gesetzt werden:

```
> t.test(jackals[, "M"], jackals[, "W"], var.equal = TRUE)
```

- d) Führe mit Hilfe von **R** einen Wilcoxon-Test durch. Gib wiederum p-Wert und Testentscheid an.
- ```
> wilcox.test(jackals[, "M"], jackals[, "W"],) # Wilcoxon-Test durchfuehren
```
- e) Falls die Resultate der beiden Tests unterschiedlich ausgefallen wären, welchem würdest Du eher vertrauen? Weshalb?