

## Exercise Sheet 9

1. You will find several examples of two-sample comparisons detailed below. For each of these examples, please give a brief answer to the following questions:

- Are the samples paired or unpaired? Give reasons for your answer.
  - Is a one-sided or two-sided test adequate for this situation? Again, give reasons for your answer.
  - How can the null hypothesis be formulated in words?
  - How can the alternative hypothesis be formulated in words?
- a) An experiment is performed to investigate the effect of smoking cigarettes on the aggregation of blood platelets. To this end, blood samples were taken from 11 subjects before and after they smoked a cigarette. The point of interest is whether or not smoking causes increased aggregation of blood platelets.
- b) The following data are taken from a study by Charles Darwin into self- and cross-pollination. 15 pairs of seedlings were bred, each pair consisting of one produced by self-pollination and one produced by cross-pollination. In each pair, both plants had practically the same conditions. The goal was to observe whether the cross-pollinated plants grew taller than the self-pollinated ones. For this purpose, the height of each plant was measured after a fixed period of time.
- c) Does the calcium content of food have an influence on systolic blood pressure? To investigate this question, a test group of 10 men were given calcium supplements over a period of 12 weeks. A control group of 11 men was given a placebo.
- d) An experiment was carried out as to whether mice absorb the two iron ions ( $\text{Fe}^{2+}$  and  $\text{Fe}^{3+}$ ) equally well. For this, 36 mice were divided into 2 groups of 18 each; one group were “fed”  $\text{Fe}^{2+}$ , and the other  $\text{Fe}^{3+}$ . Since the iron had a radioactive marker, its initial and subsequent concentration could be measured. In this way, the proportion of iron absorbed by each mouse could be computed.

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?

2. 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?