Solutions to Exercise Sheet 9

1. a) **Paired samples:** Each blood platement observation before smoking corresponds to the blood platelet observation of the same person after smoking.

One-sided test: We do not want to know whether platelet aggregation has *changed* – we want to know whether it has *increased*.

H_0**:** Smoking has no effect on the aggregation of blood platelets. ($\mu_{\mathrm{R}} = \mu_{\mathrm{NR}}$)

H_A**:** Smoking increases the aggregation of blood platelets. ($\mu_{\mathrm{R}} > \mu_{\mathrm{NR}}$)

b) **Paired samples:** Each height of a self-pollinated seedling corresponds to the height of its cross-pollinated "partner".

One-sided test: We do not want to know whether the heights *differ*, but whether the cross-pollinated seedlings grow *higher* than the self-pollinated ones.

H_0**:** The heights do not differ. ($\mu_{\mathrm{f}} = \mu_{\mathrm{s}}$)

H_A**:** Cross-pollinated seedlings grow higher than self-pollinated ones. ($\mu_{\mathrm{f}} > \mu_{\mathrm{s}}$)

c) **Unpaired samples:** The group size varies between groups. A blood pressure measurement from the test group does not correspond to one from the control group.

Two-sided test: We only want to know whether calcium has an efffect on blood pressure, *in whatever direction.*

H_0**:** Calcium has no effect on blood pressure. ($\mu_{\mathrm{Kalz}} = \mu_{\mathrm{Kontr}}$)

H_A**:** Calcium has an effect on blood pressure. ($\mu_{\mathrm{Kalz}} \neq \mu_{\mathrm{Kontr}}$)

d) **Unpaired sample:** The group sizes do not necessarily have to be the same. One iron measuremment on an "Fe^{2+} mouse" does not correspond to any particular measurement on an "Fe^{3+} mouse".

Two-sided test: We only want to know about *differences* in the absorbtion of the different types of iron.

H_0**:** Iron absorption is independent of its type. ($\mu_2 = \mu_3$)

H_A**:** Iron absorbtion depends on iron type. ($\mu_2 \neq \mu_3$)

2. a) Es handelt sich um **verbundene** Stichproben. Am gleichen Ort wird es mit beide Geräte gemessen.

b)
- Modellannahme: Die Verteilung der Differenzen ist $X_i \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\mu, \sigma^2),\ i = 1, \dots n;$
- $H_0 : X_i \sim \mathcal{N}(\mu_0, \sigma^2),\ i = 1, \dots n;$ mit $\mu_0 = 0,\ \sigma^2$ unbekannt.
- $H_A : X_i \sim \mathcal{N}(\mu_A, \sigma^2),\ i = 1, \dots n;$ mit $\mu_A \neq 0,\ \sigma^2$ wie unter H_0.

- Teststatistik:
$$t = \frac{\bar{x} - 0}{sd_x/\sqrt{9}} = -2.8$$
- Verwerfungsbereich: $K = \{|T| \geq t_{9-1,0.975}\} = \{|T| \geq 2.31\}$
- Testergebnis: Der Wert t der Teststatistik liegt im Verwefungsbereich von H_0, d.h. eine neue Eichung der Geräte ist angezeigt.

c) Z wäre binomialverteilt mit Parametern $n = 9$ und $p = \frac{1}{2}$. Der Vorzeichentest nützt dieses Faktum aus.

3. a) Es handelt sich um **unverbundene** 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) Zweiseitiger t-Test:

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. $\mathcal{N}(\mu, \sigma^2)$, Y_j i.i.d. $\mathcal{N}(\mu, \sigma^2)$ unabhängig

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

Teststatistik: $T = (\overline{X} - \overline{Y})/s_{\overline{X}-\overline{Y}}$, wobei
$$s_{\overline{X}-\overline{Y}} = \sqrt{\left(\frac{1}{n} + \frac{1}{m}\right) \frac{1}{n+m-2} \cdot ((n-1)s_x^2 + (m-1)s_y^2)}.$$
Unter H_0 gilt: $T \sim t_{n+m-2}$, also hier $T \sim t_{18}$

Verwerfungsbereich: Tabelle: $t_{18,0.975} = 2.1$ (Test zweiseitig auf 5%-Niveau)
somit: $\mathcal{K} = \{|T| > t_{18,0.975}\} = \{|T| > 2.1\}$

Wert der Teststatistik: $s_{\overline{X}-\overline{Y}} = \sqrt{\frac{2}{10}\frac{1}{18} \cdot (9 \cdot 13.82 + 9 \cdot 5.16)} = 1.38$
$T = \frac{113.4-108.6}{1.38} = 3.48$

Entscheidung: Da $T \in \mathcal{K}$ (“T Element des Verwerfungsbereichs”), wird die Nullhypothese H_0 auf dem 5%-Niveau durch den t-Test **verworfen**.

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

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

        Welch Two Sample t-test

data:  jackals[, "M"] and jackals[, "W"]
t = 3.48, df = 14.9, p-value = 0.00336
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 1.86 7.74
sample estimates:
mean of x mean of y
      113       109
```

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

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

```
> 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 p-Wert ist $0.0048 < 0.05$, 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 normalverteilt sind und wir diese Voraussetzung in keiner Weise überprüft haben. Allerdings ist die stark unterschiedliche Standardabweichung in den zwei Gruppen problematisch für beide Tests.