Lösungen Testat STOC SW11

Daniel Winz

14. Juni 2013 19:43

Inhaltsverzeichnis

1 Aufgabe 1

```
> messga=c(120,265,157,187,219,288,156,205,163);
> messgb=c(127,281,160,185,220,298,167,203,171);
> messgdiff=messga-messgb;
```

1.1 a

gepaarte Stichproben

1.2 b

1. Modell
   $X \sim N(\mu, {\sigma_x}^2), \sigma_x$wird durch $\hat{\sigma_x}$ geschätzt.

2. Nullhypothese
   $H_0 : \mu = \mu_0$
   $H_A : \mu < \mu_0$

3. Teststatistik

   $$T = \frac{\sqrt{n} \cdot (\bar{X_n} - \mu_0)}{\hat{\sigma_x}}$$

   Verteilung der Teststatistik unter $H_0 : T \sim t_{15}$

4. Signifikanzniveau
   $\alpha = 0.05$

5. Verwerfungsbereich der Teststatistik
   $K = [t_{n-1;1-\alpha}, \infty]$

   ```
   >           qt(p=0.95,df=length(messgdiff))

   [1] 1.833113
   ```

6. Testentscheid
   $T = \frac{\sqrt{16} \cdot (204.2 - 200)}{10} = 1.68$
   T liegt im Verwerfungsbereich $\rightarrow H_0$ kann verworfen werden.

1.3 c

Z ist binomial verteilt.

2 Aufgabe 2

2.1 a

Es handelt sich um ungepaarte Stichproben.

2.2 b

2.3 c

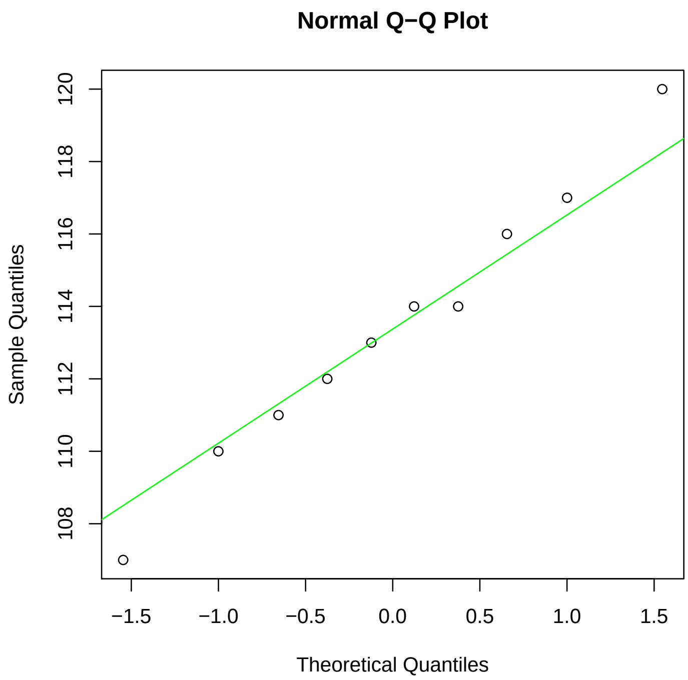
```
> # jackals=read.table("http://stat.ethz.ch/Teaching/Datasets/jackals.dat", header=TRUE)
> jackals=read.table("jackals.dat", header=TRUE)
> t.test(jackals[,'M'], jackals[,'W'], var.equal = TRUE)

        Two Sample t-test

data:  jackals[, "M"] and jackals[, "W"]
t = 3.4843, df = 18, p-value = 0.002647
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 1.905773 7.694227
sample estimates:
mean of x mean of y
    113.4     108.6
```

2.4 d

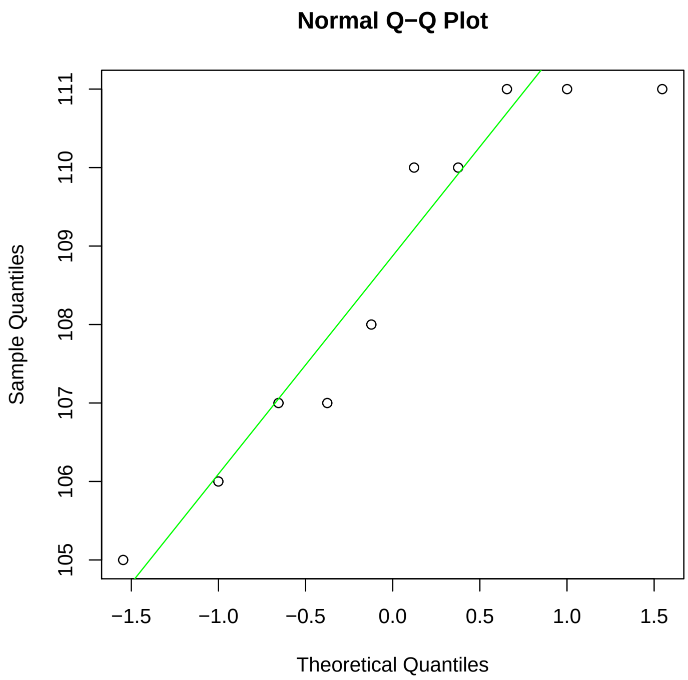
```
> wilcox.test(jackals[,'M'], jackals[,'W'], var.equal = TRUE)

        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
```

2.5 e

```
> qqnorm(jackals[,"M"])
> qqline(jackals[,"M"],col='green')
```




```
> qqnorm(jackals[,"W"])
> qqline(jackals[,"W"],col='green')
```



Aufgrund der qqplots kann von einer Normalverteilung ausgegangen werden. Daher ist der t-Test präziser.
Der Wilcoxon-Test wäre besser geeignet, wenn die Daten nicht normal verteilt wären.

3 Aufgabe 3

3.1 a

Es handelt sich um einen gepaarten Test, weil jeder Athlet mit beiden Schuhen getestet wird.

3.2 b

1. Modell
   $D \sim \mathcal{N}(\mu, \hat{\sigma}^2_{x-y})$

2. Nullhypothese
   $H_0 : \mu = \mu_0 = 0$
   Alternativhypothese:
   $H_A : \mu < \mu_0$

3. Teststatistik
   $T = \frac{\sqrt{n} \cdot (\bar{x} - \bar{y} - \mu_0)}{\hat{\sigma}_{x-y}}$
   $T = \frac{\sqrt{10} \cdot (-0.22 - 0)}{0.26} = -2.676$
   Verteilung der Teststatistik unter H_0
   $T \sim t_{n-1}$

4. Signifikanzniveau
   $\alpha = 0.05$

5. Verwerfungsbereich
   $K = (-\infty, -t_{n-1,1-\alpha}]$

   ```
   > qt(p=0.05,df=10-1)
   ```

   ```
   [1] -1.833113
   ```

   $K = (-\infty, -1.812461]$

6. Testentscheid
   T liegt im Verwerfungsbereich.

3.3 c

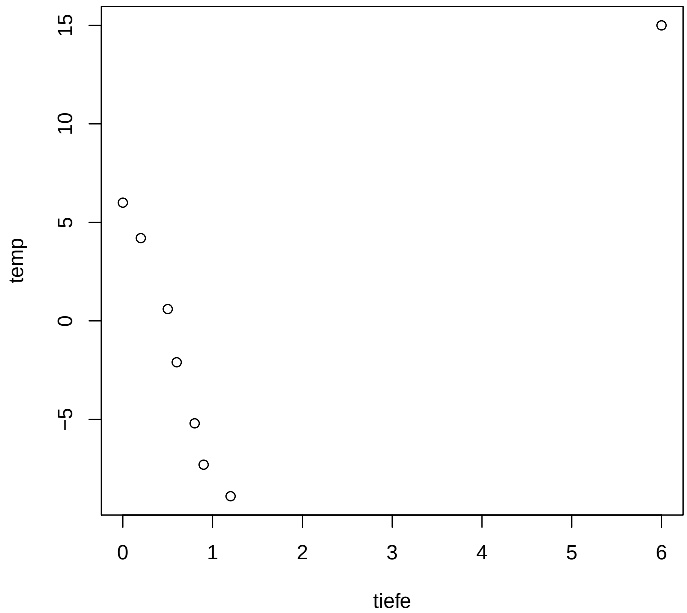
95 % Vertrauensintervall beinhaltet alle Werte $\mu_x - \mu_y$
$K = (-\infty, -t_{9,0.95}]$
$-t_{9,0.95} \leq \frac{\sqrt{n}(\bar{x}_n - \bar{x}_n - (\mu_x - \mu_y))}{\hat{\sigma}_{x-y}}$
$\Leftrightarrow \mu_x - \mu_y \leq \bar{x}_n - \bar{y}_n + \frac{\hat{\sigma}_{x-y} \cdot 9,0.95}{\sqrt{n}}$
$\Rightarrow (-\infty, -0.069]$

4 Aufgabe 4

```
> tiefe=c(0,0.2,0.5,0.6,0.8,0.9,1.2,6)
> temp=c(6,4.2,0.6,-2.1,-5.2,-7.3,-8.9,15)
```

4.1 a

```
> plot(tiefe,temp)
```



Auffallend: Der letzte Datenpunkt weicht stark ab.
Interpretationen:

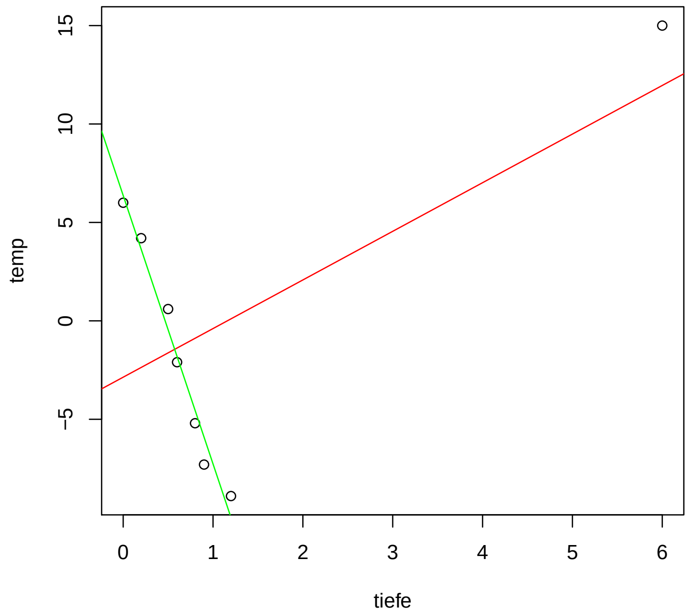
- Vorzeichenfehler beim letzten Datenpunkt
- Messfehler
- Erwärmung der Bohrung durch die lange dauernde Bohrung von 1.2 bis 6 m Tiefe

4.2 b

```
> tiefekorr=tiefe[1:length(tiefe)-1]
> tempkorr=temp[1:length(temp)-1]
> cor(tiefe,temp)

[1] 0.604961

> cor(tiefekorr,tempkorr)

[1] -0.9865329
```

4.3 c

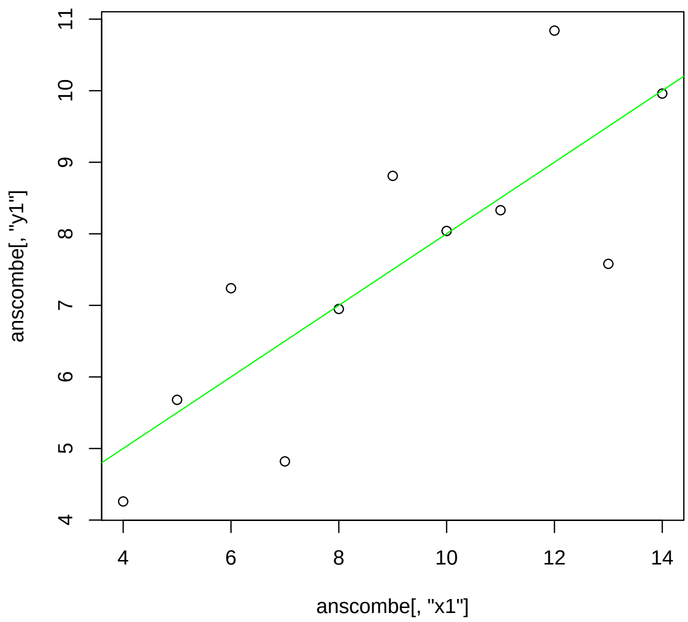
```
> plot(tiefe,temp)
> abline(lm(temp ~ tiefe),col='red')
> abline(lm(tempkorr ~ tiefekorr),col='green')
```



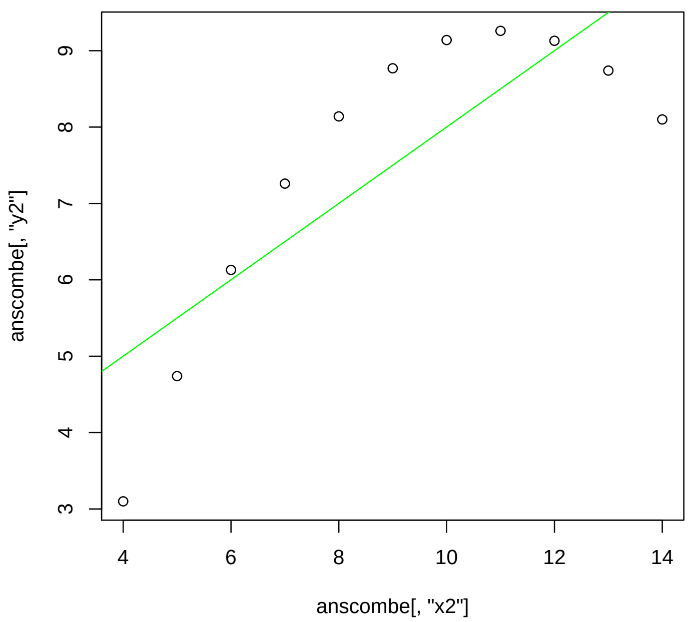
5 Aufgabe 5

5.1 a

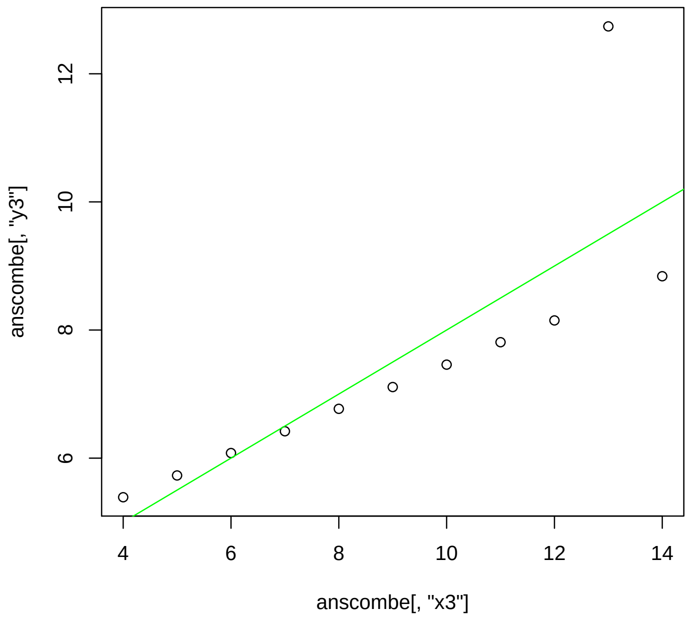
```
> plot(anscombe[,'x1'],anscombe[,'y1'])
> abline(lm(anscombe[,'y1']~anscombe[,'x1']),col='green')
```



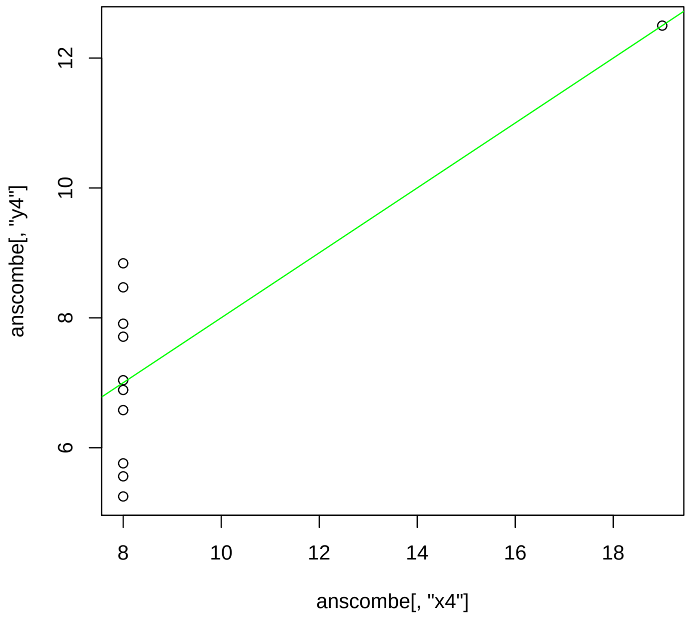

```
> plot(anscombe[,'x2'],anscombe[,'y2'])
> abline(lm(anscombe[,'y2']~anscombe[,'x2']),col='green')
```



```
> plot(anscombe[,'x3'],anscombe[,'y3'])
> abline(lm(anscombe[,'y3']~anscombe[,'x3']),col='green')
```




```
> plot(anscombe[,'x4'],anscombe[,'y4'])
> abline(lm(anscombe[,'y4']~anscombe[,'x4']),col='green')
```



5.2 b

```
> summary(lm(anscombe[,'y1']~anscombe[,'x1']))

Call:
lm(formula = anscombe[, "y1"] ~ anscombe[, "x1"])

Residuals:
     Min       1Q   Median       3Q      Max
-1.92127 -0.45577 -0.04136  0.70941  1.83882

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
(Intercept)        3.0001     1.1247   2.667  0.02573 *
anscombe[, "x1"]   0.5001     0.1179   4.241  0.00217 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.237 on 9 degrees of freedom
Multiple R-squared: 0.6665,     Adjusted R-squared: 0.6295
F-statistic: 17.99 on 1 and 9 DF,  p-value: 0.00217

> summary(lm(anscombe[,'y2']~anscombe[,'x2']))

Call:
lm(formula = anscombe[, "y2"] ~ anscombe[, "x2"])
```

```
Residuals:
    Min      1Q  Median      3Q     Max 
-1.9009 -0.7609  0.1291  0.9491  1.2691 

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)   
(Intercept)         3.001      1.125   2.667  0.02576 * 
anscombe[, "x2"]    0.500      0.118   4.239  0.00218 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1 

Residual standard error: 1.237 on 9 degrees of freedom
Multiple R-squared: 0.6662,	Adjusted R-squared: 0.6292 
F-statistic: 17.97 on 1 and 9 DF,  p-value: 0.002179 

> summary(lm(anscombe[,'y3']~anscombe[,'x3']))

Call:
lm(formula = anscombe[, "y3"] ~ anscombe[, "x3"])

Residuals:
    Min      1Q  Median      3Q     Max 
-1.1586 -0.6146 -0.2303  0.1540  3.2411 

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)   
(Intercept)        3.0025     1.1245   2.670  0.02562 * 
anscombe[, "x3"]   0.4997     0.1179   4.239  0.00218 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1 

Residual standard error: 1.236 on 9 degrees of freedom
Multiple R-squared: 0.6663,	Adjusted R-squared: 0.6292 
F-statistic: 17.97 on 1 and 9 DF,  p-value: 0.002176 

> summary(lm(anscombe[,'y4']~anscombe[,'x4']))

Call:
lm(formula = anscombe[, "y4"] ~ anscombe[, "x4"])

Residuals:
   Min     1Q Median     3Q    Max 
-1.751 -0.831  0.000  0.809  1.839 

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)   
(Intercept)        3.0017     1.1239   2.671  0.02559 * 
anscombe[, "x4"]   0.4999     0.1178   4.243  0.00216 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1 
```

```
Residual standard error: 1.236 on 9 degrees of freedom
Multiple R-squared: 0.6667,       Adjusted R-squared: 0.6297
F-statistic:    18 on 1 and 9 DF,  p-value: 0.002165
```