Lösungen zur Übung 3

1.
```
> lyso <- read.table("/afs/ethz.ch/users/s/sfs/datasets/lyso.dat",header=T)
> attach(lyso)

> t.test(lysozym[gruppe==0],lysozym[gruppe==1])

         Standard Two-Sample t-Test
data:  lysozym[gruppe == 0] and lysozym[gruppe == 1]
t = 2.0569, df = 57, p-value = 0.0443
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
  0.1754851 13.0785379
sample estimates:
 mean of x mean of y
  14.31034  7.683333

> t.test(log(lysozym[gruppe==0]),log(lysozym[gruppe==1]))

         Standard Two-Sample t-Test
data:  log(lysozym[gruppe == 0]) and log(lysozym[gruppe == 1])
t = 1.4088, df = 57, p-value = 0.1643
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.216398  1.243467
sample estimates:
 mean of x mean of y
  1.921094  1.407559

> wilcox.test(lysozym[gruppe==0],lysozym[gruppe==1])

         Wilcoxon rank-sum test
data:  lysozym[gruppe == 0] and lysozym[gruppe == 1]
rank-sum normal statistic with correction Z = 1.5998, p-value = 0.1096
```

a) Signifikanz auf dem 5%- Niveau

b) c) nicht signifikant

d) Boxplots zeigen deutlich unterschiedliche Varianzen. Normalplot zeigt Nicht-Normalität der untransformierten Daten. Nach dem Logarithmieren sieht die Situation besser aus. Der 2-Stichproben-t-Test ist jetzt einigermassen legitim. Ein nichtparametrischer Test ist die sicherste Variante.

2. **a)** Nach einem Jahr sind von den Haustierbesitzern 5.7% gestorben, von den Nicht-Haustierbesitzern 28.2%, also deutlich mehr.

b) H_0 : Es existiert kein Zusammenhang zwischen Überleben und Haustierbesitz. H_A : Es existiert ein Zusammenhang zwischen Überleben und Haustierbesitz.

c) `chisq.test(matrix(c(28,11,50,3),ncol=2))` liefert `X-squared = 7.1899, df = 1, p-value = 0.0073`

d) Ein Zusammenhang scheint zu bestehen, aber das heisst natürlich nicht, dass das Haustier selbst eine positive Wirkung auf die Gesundheit seines Herrchens hat. Vielleicht haben auch gesündere Leute eher ein Haustier,
oder reichere Leute, oder ....

3.
```
> beob <- c(3,12,17,25,14,5,4,0)
> tage <- 0:7
> sum(tage*beob)/(7*80)
[1] 0.4035714
```

a) Die Wahrscheinlichkeit, dass an einem Tag Regen fällt, ist gleich 0.40.

```
> nieder <- data.frame(80*dbinom(0:7,7,226/560),beob)
> names(nieder)[1] <- "erw"
> nieder
         erw beob
1  2.1478259    3
2 10.1732350   12
3 20.6510579   17
4 23.2891172   25
5 15.7585045   14
6  6.3977641    5
7  1.4430087    4
8  0.1394867    0

> sum((nieder[,2]-nieder[,1])^2/nieder[,1])
[1] 6.609371
> qchisq(0.95,6)
[1] 12.59159
```

b) Der χ^2-Quadrat-Test ist nicht signifikant. Die Daten sind mit der Binomialverteilung vereinbar.