STATISTICS 3N03 & 3J04
ASSIGNMENT 5 SOLUTIONS
2006-11-24
Updated 2006-12-12

Question 1

(a) You need at least 29 degrees of freedom.

```
> for (df in 25:30) print(c(df,qchisq(.995,df)/qchisq(.005,df)))
[1] 25.000000  4.460974
[1] 26.000000  4.326958
[1] 27.000000  4.204493
[1] 28.000000  4.092128
[1] 29.000000  3.988646
[1] 30.000000  3.893019
```

(b) Here, $\alpha = 0.01$, $\beta = 0.10$, $\delta = 0.5$ and $\sigma = 1.3$, so the required sample size n is given by

```
> ((qnorm(1-(0.01)/2) + qnorm(1-0.1))*1.3/0.5)^2
[1] 100.5847
```

or, using table values from the text

```
> ((2.576 + 1.282)*1.3/0.5)^2
[1] 100.6169
```

so 101 observations would be required.

The probability of a Type II error is computed by text formula (9-17); when n = 10 it gives

```
> pnorm(qnorm(1-(0.01)/2)-0.5*sqrt(10)/1.3)
 + pnorm(-qnorm(1-(0.01)/2)-0.5*sqrt(10)/1.3)
[1] 0.9130915
```

and this probability is much too high for the test to be useful.

(c) *One-sided test:*

```
> ((qnorm(.95)+qnorm(.95))*2/(11.5-12))^2
[1] 173.1548
> 1 - pnorm(-qnorm(.95)-(11.5-12)*sqrt(50)/2)
[1] 0.4510879
```

The number of paint samples required to reduce the Type II error rate to 5% is 174; if you could only test 50 samples, the Type II error rate would be 45% which is too high for the test to be useful.

Two-sided test:

```
> ((qnorm(.975)+qnorm(.95))*2/(11.5-12))^2
[1] 207.9154
> pnorm(qnorm(.975)-(11.5-12)*sqrt(50)/2)
 - pnorm(-qnorm(.975)-(11.5-12)*sqrt(50)/2)
```

```
[1] 0.5761095
```

The number of paint samples required to reduce the Type II error rate to 5% is 208; if you could only test 50 samples, the Type II error rate would be 57.6% which is too high for the test to be useful. ***(10 marks for <u>either</u> 1-sided <u>or</u> 2-sided answer)***

(d) Assuming that the flaws occur independently, at random, at a constant average rate over the windshield (i.e. as a Poisson process), the number of flaws per windshield will follow a Poisson distribution. The probability it was produced at Plant A given that it has 3 flaws is found by Bayes' Theorem to be 20.8%. ***(11 marks)***

Let A be the event it was produced at Plant A, let X be the number of flaws.

P(A|X=3) = P(X=3|A)P(A)/[P(X=3|A)P(A) + P(X=3|B)P(B)]

```
> dpois(3,2.1)*0.2/(dpois(3,2.1)*0.2 + dpois(3,4.3)*0.8)
[1] 0.2081147
```

Question 2

(*a*) Acceptable analyses: Paired t-test, 2-factor ANOVA without interaction, sign test, simple linear regression.
Acceptable graphs: dot, box or stem-leaf plot of differences; interaction plot; scatter plot with fitted line. ***(Graph 2, suitable analysis 2, correct calculation 4, assumptions 2, conclusions 2)***

Paired t-test

Assumptions: Differences independent (can't test), normal (looks OK on dot plot and stem & leaf plot).

Conclusions: There is no evidence (P = 0.37) from these data of a difference in mean time between the two processors.

```
> stem(procspeed$diff)

  The decimal point is 1 digit(s) to the right of the |

  -0 | 8
  -0 | 310
   0 | 03

> procspeed
  code procA procB diff
1    1  27.2  24.1  3.1
2    2  18.1  19.3 -1.2
3    3  27.2  26.8  0.4
4    4  19.7  20.1 -0.4
5    5  24.5  27.6 -3.1
6    6  22.1  29.8 -7.7
> t.test(procspeed$procA, procspeed$procB, pair=T)

        Paired t-test

data:  procspeed$procA and procspeed$procB
t = -0.9921, df = 5, p-value = 0.3667
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -5.326857  2.360190
sample estimates:
mean of the differences
              -1.483333
```

2-factor ANOVA

```
> procspeed2
   code time process
1     1 27.2       A
2     2 18.1       A
3     3 27.2       A
4     4 19.7       A
5     5 24.5       A
6     6 22.1       A
7     1 24.1       B
8     2 19.3       B
9     3 26.8       B
10    4 20.1       B
11    5 27.6       B
```

```
12    6 29.8       B

> anova(lm(time~code+process, procspeed2))
Analysis of Variance Table

Response: time
          Df  Sum Sq Mean Sq F value Pr(>F)
code       5 129.068  25.814  3.8488 0.0827 .
process    1   6.601   6.601  0.9842 0.3667
Residuals  5  33.534   6.707
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Sign test

Assumptions: Differences are independent (can't test).

Conclusions: There are 2 positive differences out of 6 non-zero differences, so P = 0.69 and there is no evidence from these data of a difference in the median time between the two processors.

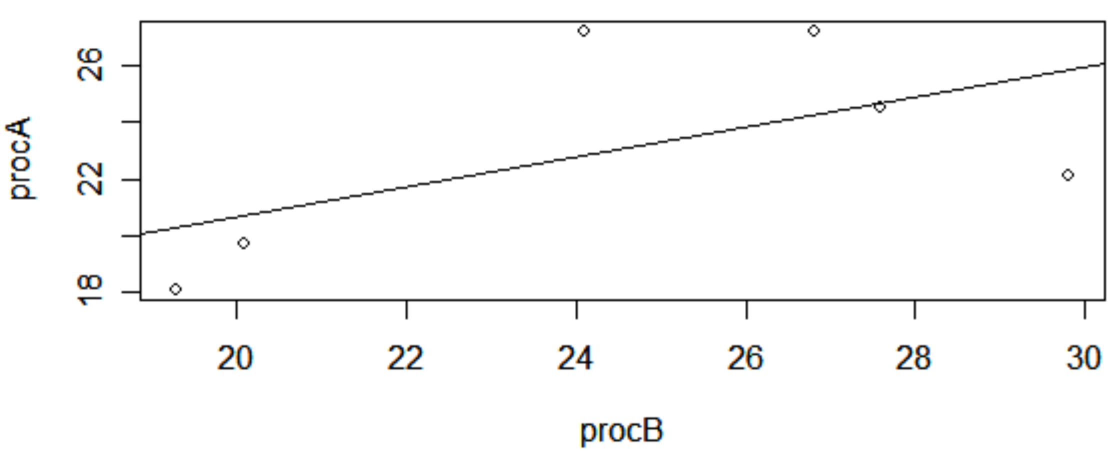
```
> 2*pbinom(2,6,.5)
[1] 0.6875
```

Simple Linear Regression

Assumptions: The time with Processor A is linearly related to the time with Processor B, data are normal and homoscedastic. We can't test any of these because the sample is small and there are no repeated x-values.

Conclusions: There is no evidence from these data of a linear relationship between Processor A and Processor B times.

```
> plot(procA~procB, procspeed)
```




```
> abline(lm(procA~procB, procspeed))
> anova(lm(procA~procB, procspeed))
Analysis of Variance Table
```

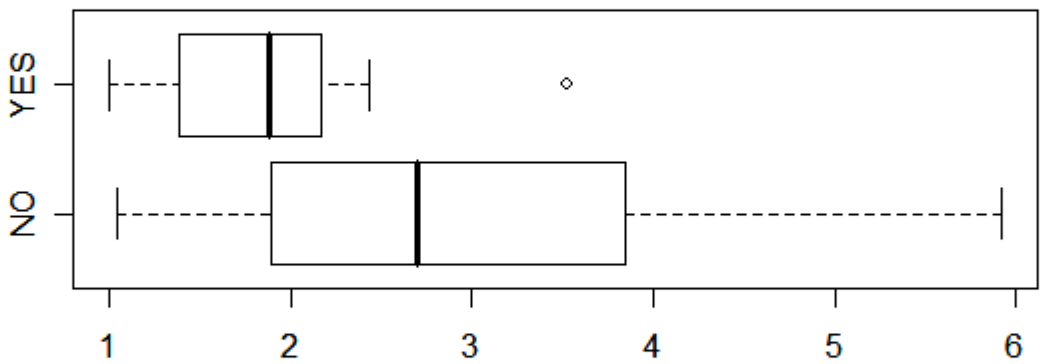
```
Response: procA
          Df Sum Sq Mean Sq F value Pr(>F)
procB      1 25.502  25.502  2.1417 0.2172
Residuals  4 47.631  11.908

> anova(lm(procB~procA, procspeed))
Analysis of Variance Table

Response: procB
          Df Sum Sq Mean Sq F value Pr(>F)
procA      1 31.199  31.199  2.1417 0.2172
Residuals  4 58.270  14.567
```

(*b*) Acceptable analyses: Two-sample t-test or 1-factor ANOVA.
Acceptable graphs: Comparative dot, box or stem-leaf plots. ***(Graph 2, suitable analysis 2, correct calculation 3, F-test of variances 2, assumptions 2, conclusions 2)***

Assumptions: Independence (can't test), normality (OK by plot), homoscedasticity (graph looks heteroscedastic; F-test gives no evidence (P = 0.14) that the variances are not equal).

Conclusions: There is no evidence from these data that either the mean or the variance of resilient modulus differs between rutted and non-rutted pavement.

```
> ruts
   resmod rutted
1    1.48    YES
2    1.88    YES
3    1.90    YES
4    1.29    YES
5    3.53    YES
6    2.43    YES
7    1.00    YES
8    3.06     NO
9    2.58     NO
10   1.70     NO
11   5.76     NO
12   2.44     NO
13   2.03     NO
14   1.76     NO
15   4.63     NO
16   2.86     NO
17   2.82     NO
18   1.04     NO
19   5.92     NO
> boxplot(resmod~rutted, ruts, horizontal=T)
```




```
> t.test(resmod~rutted, ruts, var.eq=T)

        Two Sample t-test

data:  resmod by rutted
t = 1.7268, df = 17, p-value = 0.1023
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.2484140  2.4884140
sample estimates:
 mean in group NO mean in group YES
             3.05              1.93

> F0 <- var(ruts$resmod[ruts$rutted=="NO"])/var(ruts$resmod[ruts$rutted=="YES"])
> F0
[1] 3.474133
> 2*(1-pf(F0, 11, 6))
[1] 0.1387276
```

Question 3

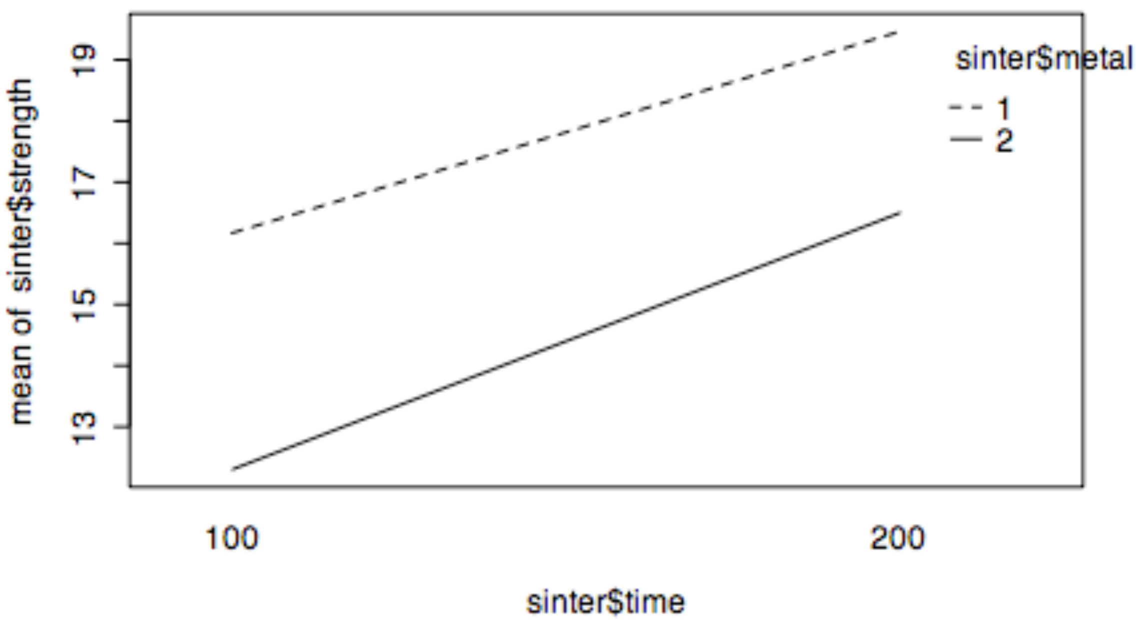
The only acceptable analysis is a 2-factor ANOVA with a test for interaction.

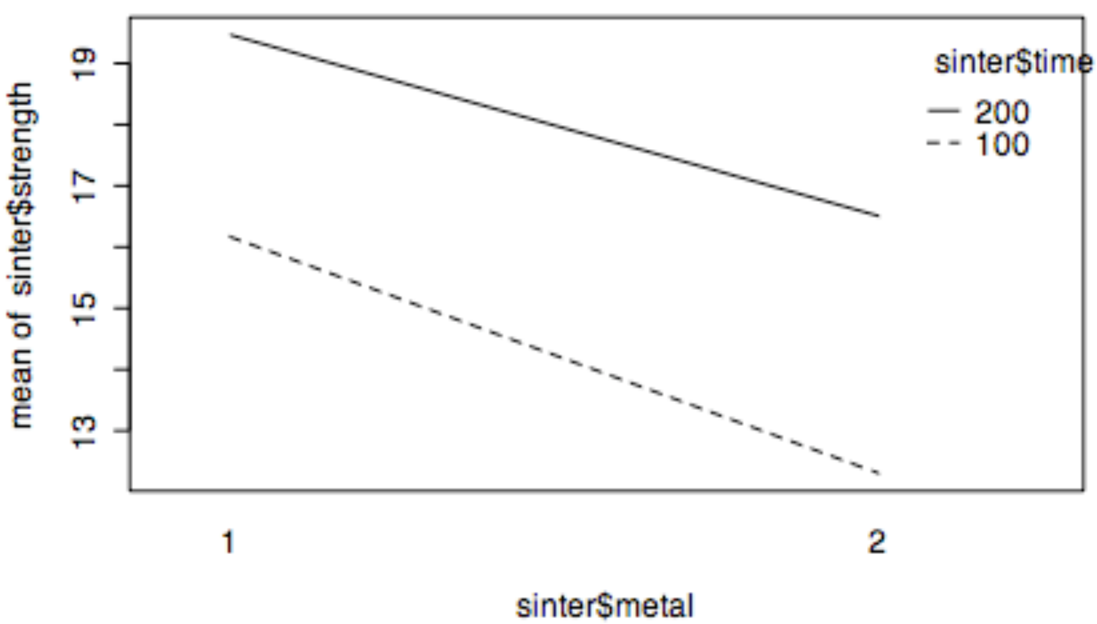
Assumptions: Normality, independence, homoscedasticity.
Conclusions: There is no evidence (P = 0.49) from these data of an interaction between metal type and sintering time. The interaction plots confirm this as the lines are parallel. There is strong evidence that the mean compressive strength is different for the different sintering times (P = 0.0003) and for the different metals (P = 0.0006).

The 95% confidence interval for the residual variance is (0.522, 4.199).

(Correct calculation with P-values 9, assumptions 2, conclusions 3, either interaction plot 3, CI for residual variance 3)

```
> sinter
   strength metal  time
1      17.1     1  100
2      16.5     1  100
3      14.9     1  100
4      12.3     2  100
5      13.8     2  100
6      10.8     2  100
7      19.4     1  200
8      18.9     1  200
9      20.1     1  200
10     15.6     2  200
11     17.2     2  200
12     16.7     2  200
> anova(lm(strength~metal*time, sinter))
Analysis of Variance Table

Response: strength
           Df Sum Sq Mean Sq F value    Pr(>F)
metal       1 35.021  35.021  30.608 0.0005522 ***
time        1 42.187  42.187  36.872 0.0002985 ***
metal:time  1  0.608   0.608   0.531 0.4869859
Residuals   8  9.153   1.144
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> mse <- anova(lm(strength~metal*time, sinter))["Residuals","Mean Sq"]
> mse
[1] 1.144167
> mse/(qchisq(c(.975,.025),8)/8)
 [1] 0.5220171 4.1992954

> interaction.plot(sinter$time,sinter$metal,sinter$strength)
> interaction.plot(sinter$metal,sinter$time,sinter$strength)
```


mean of sinter$strength
13
15
17
19
100
200
sinter$time
sinter$metal
1
2


mean of sinter$strength
13
15
17
19
1
2
sinter$metal
sinter$time
200
100

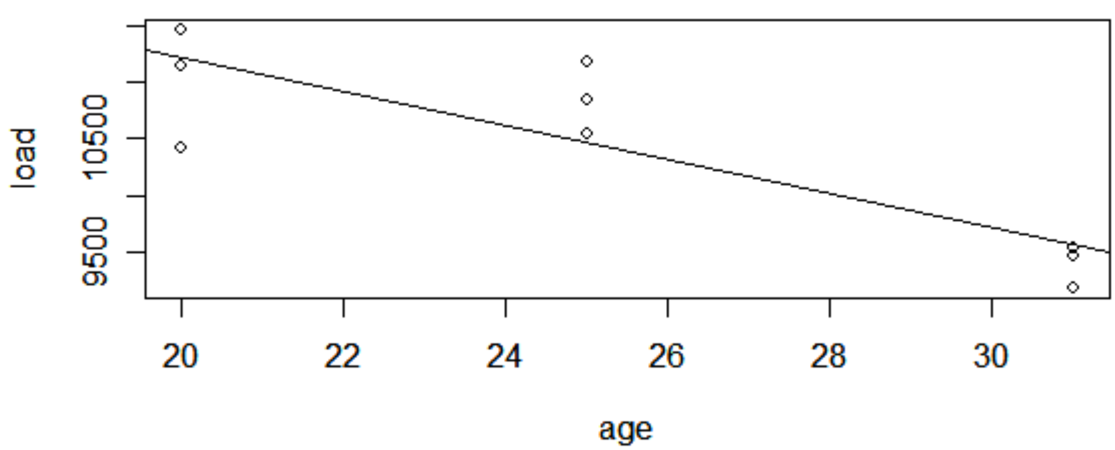
Question 4

Assumptions: Independence, normality, homoscedasticity. The test of the slope also assumes linearity. Linearity is tested by the lack of fit F-test. Normality and Homoscedasticity can't be tested in such a small sample but look OK on the graph.

Conclusions: There is some evidence (P = 0.07) from these data that the relationship is non-linear. However, if we choose to ignore that and test the slope, there is strong evidence (P = 0.002) that the slope is not zero.

(Graph with line 4, regression analysis 7, lack of fit test 5, assumptions 3, conclusions 3, 99% confidence interval for MSE <u>or</u> MSPE 3)

```
> concrete
   load age
1 11450  20
2 10420  20
3 11142  20
4 10840  25
5 11170  25
6 10540  25
7  9470  31
8  9190  31
9  9540  31
> plot(load~age, concrete)
> abline(lm(load~age, concrete))
```




```
> coef(lm(load~age, concrete))
(Intercept)         age
 14192.1099   -148.9780

> anova(lm(load~age, concrete))
Analysis of Variance Table

Response: load
          Df  Sum Sq Mean Sq F value   Pr(>F)
age        1 4039390 4039390   19.03 0.003305 **
Residuals  7 1485858  212265
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> mse <- anova(lm(load~age, concrete))["Residuals", "Mean Sq"]
> mse
[1] 212265.4
> mse/(qchisq(c(.995,.005),7)/7)
[1]   73275.32 1501995.83

> anova(lm(load~age+as.factor(age), concrete))
Analysis of Variance Table

Response: load
               Df  Sum Sq Mean Sq F value   Pr(>F)
age             1 4039390 4039390 29.3341 0.001639 **
as.factor(age)  1  659642  659642  4.7903 0.071204 .
Residuals       6  826216  137703
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> mspe <- anova(lm(load~age+as.factor(age), concrete))["Residuals", "Mean Sq"]
> mspe
[1] 137702.7
> mspe/(qchisq(c(.995,.005),6)/6)
[1]   44545.75 1222707.21
```

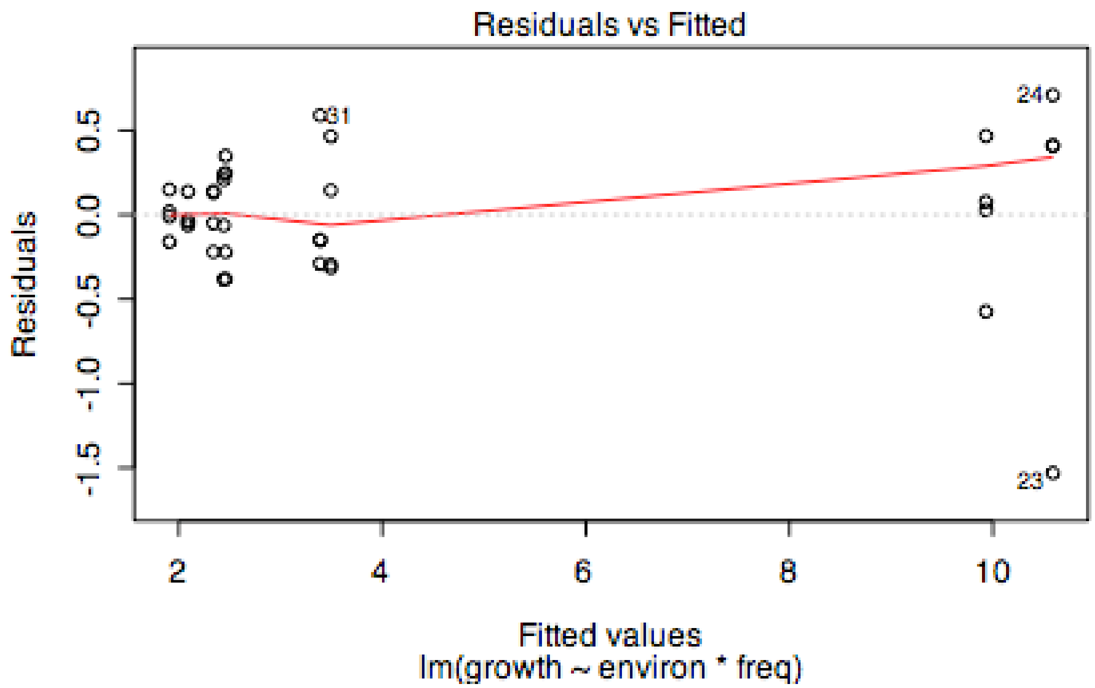
Question 5

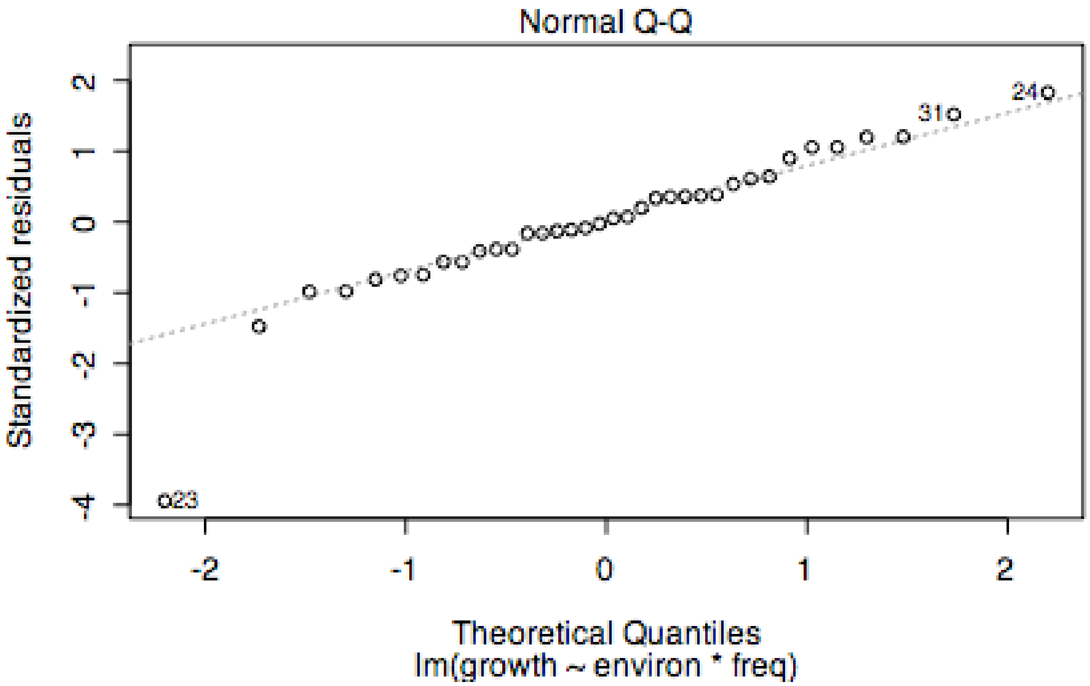
The analyses in original units and on a log scale give very similar results and lead to the same conclusion: the interaction is significant at the 5% level (or, better to say, $P << 0.001$ so there is very strong evidence of an interaction between frequency and environment). That means that both frequency and environment affect the crack growth rate, but the effect of the environment is different at different frequencies; the higher the frequency, the less difference the environment makes. Because the interaction is significant, we do not test the main effects.

The residual plots show that the residuals from the log-scale analysis follow a normal distribution more closely than residuals from the original-scale analysis. In the original scale, the 23rd observation is an outlier with a large negative residual.

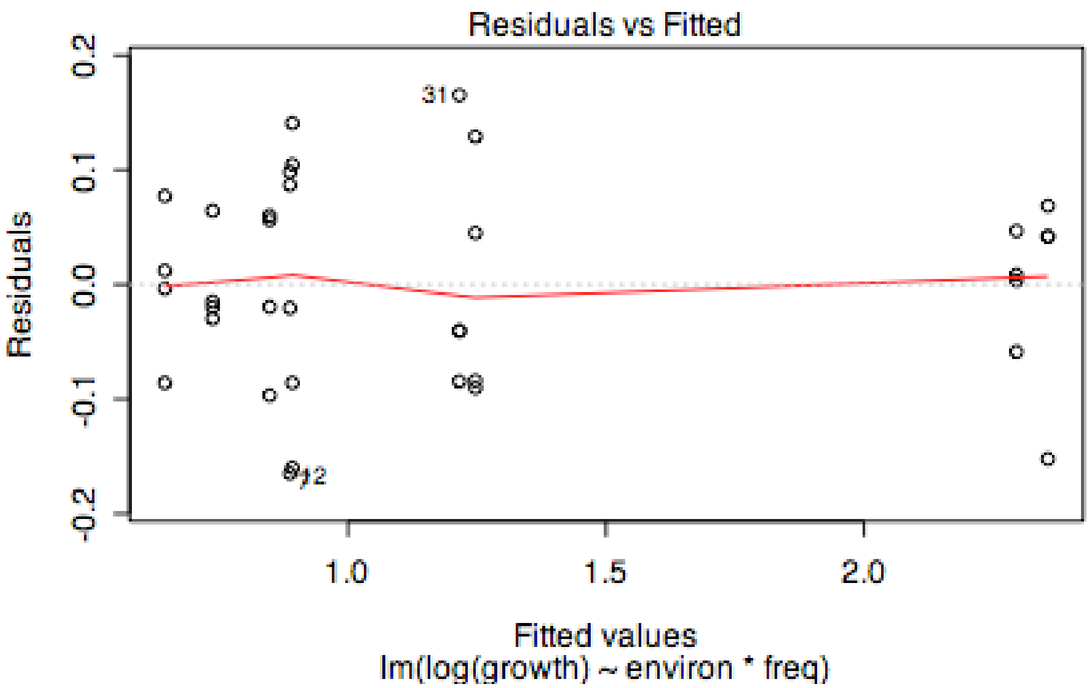
```
> cracks
   growth   environ freq
1    2.29       Air   10
2    2.47       Air   10
3    2.48       Air   10
4    2.12       Air   10
5    2.65       Air    1
6    2.68       Air    1
7    2.06       Air    1
8    2.38       Air    1
9    2.24       Air  0.1
10   2.71       Air  0.1
11   2.81       Air  0.1
12   2.08       Air  0.1
13   2.06     Water   10
14   2.05     Water   10
15   2.23     Water   10
16   2.03     Water   10
17   3.20     Water    1
18   3.18     Water    1
19   3.96     Water    1
20   3.64     Water    1
21  11.00     Water  0.1
22  11.00     Water  0.1
23   9.06     Water  0.1
24  11.30     Water  0.1
25   1.90 Saltwater   10
26   1.93 Saltwater   10
27   1.75 Saltwater   10
28   2.06 Saltwater   10
29   3.10 Saltwater    1
30   3.24 Saltwater    1
31   3.98 Saltwater    1
32   3.24 Saltwater    1
33   9.96 Saltwater  0.1
34  10.01 Saltwater  0.1
35   9.36 Saltwater  0.1
36  10.40 Saltwater  0.1
```

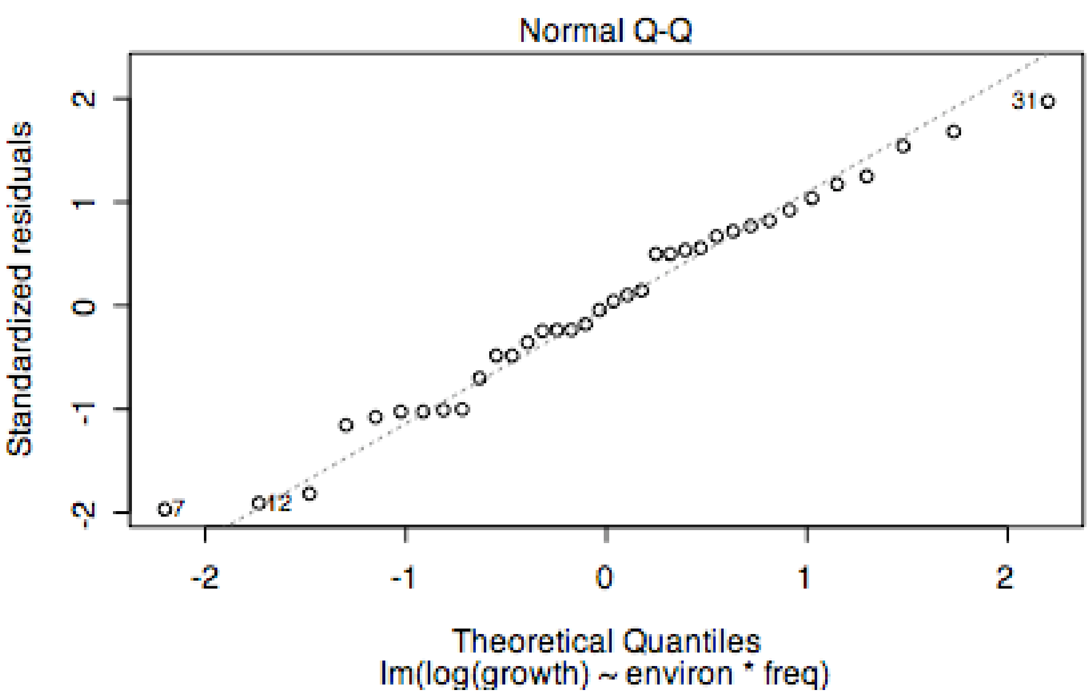
```
> anova(lm(growth~environ*freq, cracks))
Analysis of Variance Table

Response: growth
             Df  Sum Sq Mean Sq F value    Pr(>F)    
environ       2  64.252  32.126  159.92 1.076e-15 ***
freq          2 209.893 104.946  522.40 < 2.2e-16 ***
environ:freq  4 101.966  25.491  126.89 < 2.2e-16 ***
Residuals    27   5.424   0.201                      
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> plot(lm(growth~environ*freq, cracks))
```






```
> anova(lm(log(growth)~environ*freq, cracks))
Analysis of Variance Table

Response: log(growth)
             Df Sum Sq Mean Sq F value    Pr(>F)    
environ       2 2.3576  1.1788 125.849 2.061e-14 ***
freq          2 7.5702  3.7851 404.095 < 2.2e-16 ***
environ:freq  4 3.5284  0.8821  94.172 1.885e-15 ***
Residuals    27 0.2529  0.0094                      
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

> plot(lm(log(growth)~environ*freq, cracks))
```


Residuals vs Fitted
Residuals
0.2
0.1
0.0
-0.1
-0.2
31
12
1.0
1.5
2.0
Fitted values
lm(log(growth) ~ environ * freq)


Normal Q-Q
Standardized residuals
2
1
0
-1
-2
31
12
7
-2
-1
0
1
2
Theoretical Quantiles
lm(log(growth) ~ environ * freq)