Math 3080 § 1. | **Concrete Example:** | Name: Example
Treibergs | **CI for Apriori Contrast** | January 25, 2016

This **R**© program explores testing for an apriori contrast. This data was taken from Walpole, Myers, Myers, *Probability and Statistics for Engineers and Scientists, 6th ed.*, Prentice Hall, 1998. Six samples of five different concrete aggregates were exposed to moisture for 48 hours and the absorbed water was measured. The question of interest today is whether the fourth aggregate differs from the others. This is an example of an apriori contrast. In this experimental design, multiple tests are not performed on the data which increase the chances of a type I error for one of them.

A statistic to indicate that the fourth aggregate performs differently than the others is

$$\theta = -\mu_1 - \mu_2 - \mu_3 + 4\mu_4 - \mu_5 = \sum_{i=1}^{I} c_i \mu_i$$

where the weights $\sum_{i=1}^{I} c_i = 0$. In a balanced design, where the same number J of replicates are tested for each factor level, the estimator is

$$\hat{\theta} = -\overline{x_{1\cdot}} - \overline{x_{2\cdot}} - \overline{x_{3\cdot}} + 4\overline{x_{4\cdot}} - \overline{x_{5\cdot}} = \sum_{i=1}^{I} c_i \overline{x_{i\cdot}}$$

Because

$$V(\hat{\theta}) = V\left(\sum_{i=1}^{I} c_i \overline{x_{i\cdot}}\right) = \sum_{i=1}^{I} c_i^2 V(\overline{x_{i\cdot}}) = \frac{\sigma^2}{J}\sum_{i=1}^{I} c_i^2$$

the α significance CI is given by

$$\sum_{i=1}^{I} c_i \overline{x_{i\cdot}} \pm t_{\alpha/2, I(J-1)} \sqrt{\frac{MSE \sum_{i=1}^{I} c_i^2}{J}}.$$

The data shows that there is significant evidence that $\theta \neq 0$.

The same contrast may be tested using the f-statistic. In this case, the contrast gives a sum of squares. The statistic

$$\frac{(\hat{\theta})^2}{\sigma_{\hat{\theta}}^2} = \frac{J\left(\sum_{i=1}^{I} c_i \overline{x_{i\cdot}}\right)^2}{\sigma^2 \sum_{i=1}^{I} c_i^2}$$

is a χ^2-square random variable with one degree of freedom. Since $s^2 = MSE$ estimates the variance,

$$\frac{J(\hat{\theta})^2}{s_{\hat{\theta}}^2} = \frac{J\left(\sum_{i=1}^{I} c_i \overline{x_{i\cdot}}\right)^2}{MSE \sum_{i=1}^{I} c_i^2}$$

is f-distributed random variable with 1 and $I(J-1)$ degrees of freedom. It is equivalent to the t-statistic.

In entering the contrast, observe that the matrix $C = \text{contrast}(ag)$ has the desired contrast as the first column, and $I - 2 = 3$ other orthononormal contrasts inserted in the other columns. Indeed, we check that $C^T C$ is the identity matrix except the upper left corner, which gives $\sum_{i=1}^{I} c_i^2$. To access the contrast sum square, the summary code for linear models is invoked instead of the summary for anova. The two tests are equivalent. They give the same p-value.

Finally, we plot the results of anova. This gives four diagnostic plots. The qq-normal plot shows points lining up along the 45°-line which shows that the standardized residuals seem to be normally distributed.

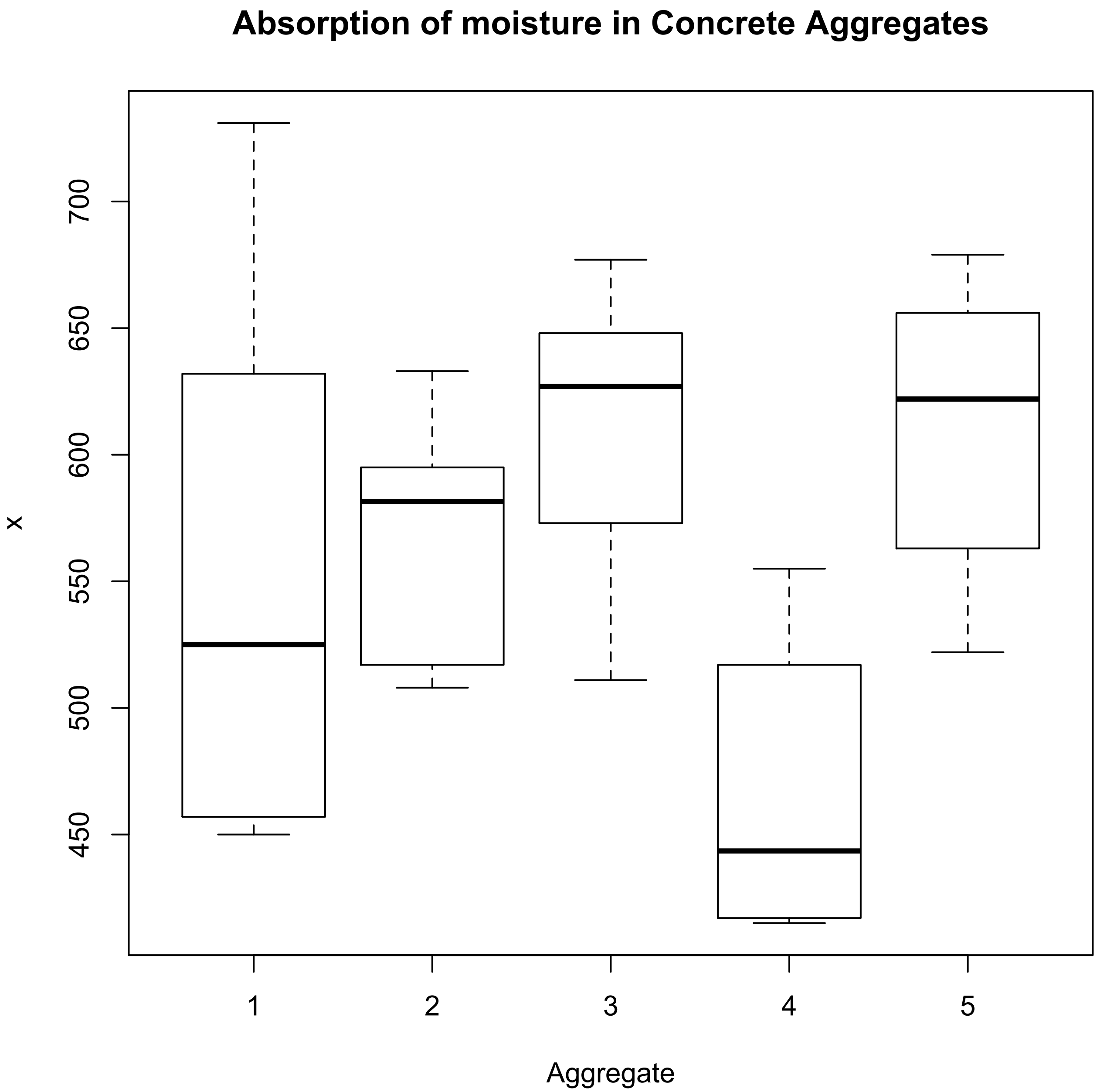
R Session:

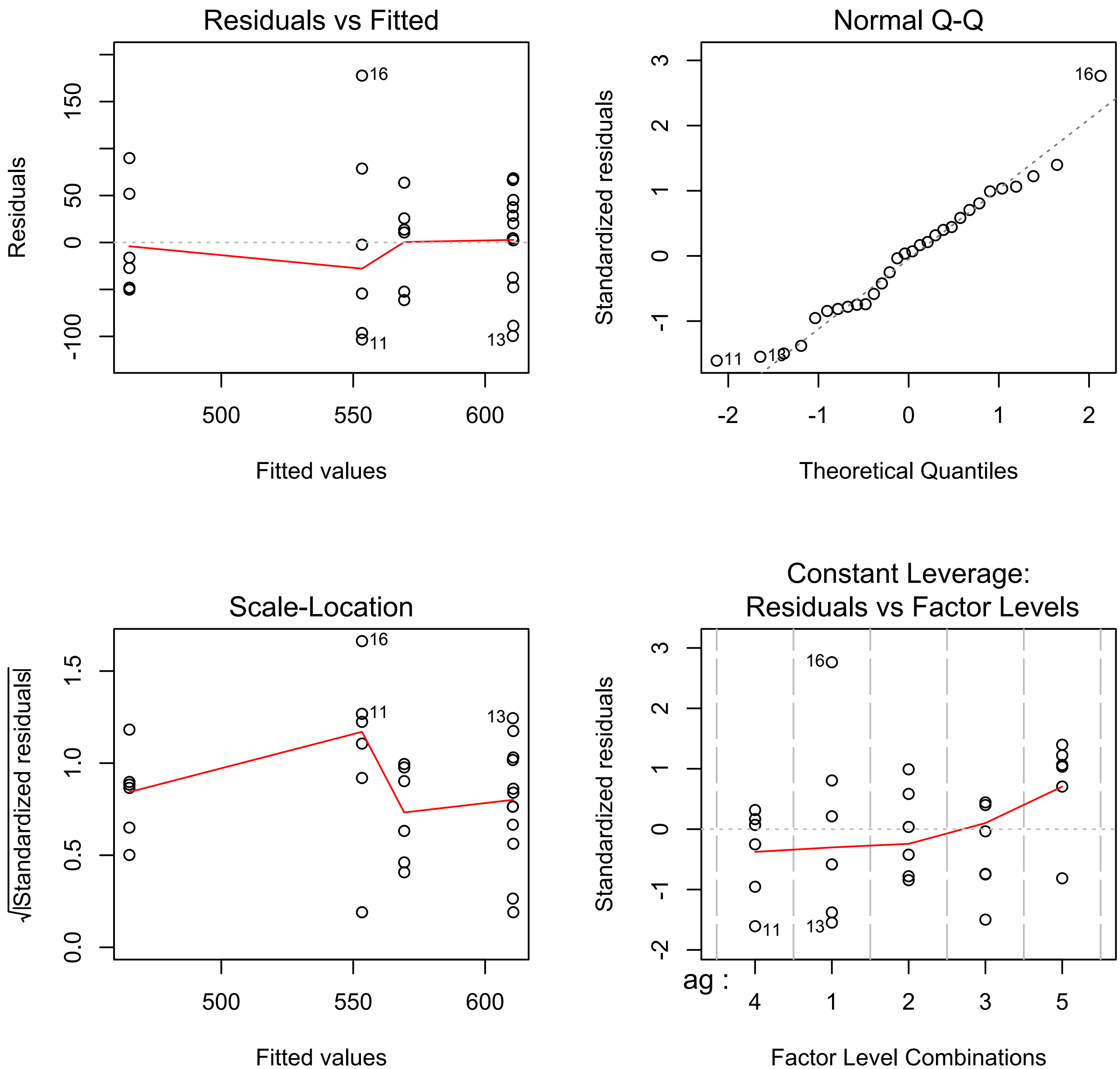
```
R version 2.13.1 (2011-07-08)
Copyright (C) 2011 The R Foundation for Statistical Computing
ISBN 3-900051-07-0
Platform: i386-apple-darwin9.8.0/i386 (32-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

  Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

[R.app GUI 1.41 (5874) i386-apple-darwin9.8.0]

[Workspace restored from /Users/andrejstreibergs/.RData]
[History restored from /Users/andrejstreibergs/.Rapp.history]

> # Absorption data
> x=scan()
1: 551
2:
Read 1 item
> x=scan()
1: 551 595 639 417 563   457 580 615 449 631   450 508 511 517 522
16: 731 583 573 438 613   499 633 648 415 656   632 517 677 555 679
31:
Read 30 items
> matrix(x,ncol=5,byrow=T)
     [,1] [,2] [,3] [,4] [,5]
[1,]  551  595  639  417  563
[2,]  457  580  615  449  631
[3,]  450  508  511  517  522
[4,]  731  583  573  438  613
[5,]  499  633  648  415  656
[6,]  632  517  677  555  679

> Ag=rep(1:5,6);Ag
 [1] 1 2 3 4 5 1 2 3 4 5 1 2 3 4 5 1 2 3 4 5 1 2 3 4 5 1 2 3 4 5
> ag=ordered(Ag)
```

```
> tapply(x,ag,summary)
$`1`
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
  450.0   467.5   525.0   553.3   611.8   731.0 

$`2`
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
  508.0   532.8   581.5   569.3   592.0   633.0 

$`3`
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
  511.0   583.5   627.0   610.5   645.8   677.0 

$`4`
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
  415.0   422.2   443.5   465.2   500.0   555.0 

$`5`
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
  522.0   575.5   622.0   610.7   649.8   679.0 

> plot(x~ag,main="Absorption of moisture in Concrete Aggregates",xlab="Aggregate")
> 
> ######### ANALYSIS OV VARIANCE
> 
> aov1 <- aov(x~ag);summary(aov1);anova(aov1)
            Df Sum Sq Mean Sq F value   Pr(>F)   
ag           4  85356 21339.1  4.3015 0.008752 **
Residuals   25 124020  4960.8                    
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1 
Analysis of Variance Table

Response: x
          Df Sum Sq Mean Sq F value   Pr(>F)   
ag         4  85356 21339.1  4.3015 0.008752 **
Residuals 25 124020  4960.8                    
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1 

> 
> #######  PLOT CANNNED DIAGNOSTICS (FOUR PLOTS ON SAME PAGE)
> 
>  layout(matrix(c(1,3,2,4),ncol=2))
>  plot(aov1)
> 
> 
```

```
> ### APRIORI CONTRAST: IS 4 SIG. DIFFERENT FROM THE OTHERS?
>
> contrasts(ag) <- cbind(c(-1,-1,-1,4,-1))
> contrasts(ag)
  [,1]          [,2]          [,3]          [,4]
1   -1 -4.172882e-01 -6.338305e-01 -4.172882e-01
2   -1 -2.834576e-01  7.676611e-01 -2.834576e-01
3   -1  8.503729e-01 -6.691527e-02 -1.496271e-01
4    4  2.775558e-17 -1.665335e-16  5.551115e-17
5   -1 -1.496271e-01 -6.691527e-02  8.503729e-01

> t(contrasts(ag)) %*% contrasts(ag)
             [,1]          [,2]          [,3]          [,4]
[1,]  2.000000e+01  3.330669e-16 -9.436896e-16  2.220446e-16
[2,]  3.330669e-16  1.000000e+00 -3.989864e-17  0.000000e+00
[3,] -9.436896e-16 -3.989864e-17  1.000000e+00 -7.632783e-17
[4,]  2.220446e-16  0.000000e+00 -7.632783e-17  1.000000e+00
> aov2 <- aov(x ~ ag);summary.lm(aov2)

Call:
aov(formula = x ~ ag)

Residuals:
     Min       1Q   Median       3Q      Max
-103.333  -49.667    3.417   43.375  177.667

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  561.800     12.859  43.688  < 2e-16 ***
ag1          -24.158      6.430  -3.757 0.000921 ***
ag2           35.499     28.754   1.235 0.228469
ag3            4.621     28.754   0.161 0.873622
ag4           35.666     28.754   1.240 0.226355
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

Residual standard error: 70.43 on 25 degrees of freedom
Multiple R-squared: 0.4077,Adjusted R-squared: 0.3129
F-statistic: 4.302 on 4 and 25 DF,  p-value: 0.008752

> anova(aov2)
Analysis of Variance Table

Response: x
          Df Sum Sq Mean Sq F value   Pr(>F)
ag         4  85356 21339.1  4.3015 0.008752 **
Residuals 25 124020  4960.8
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1
>
```

```
> #######  COMPUTE THE CI FOR THE CONTRAST FOR SEVERAL alpha's
>
> xbar=tapply(x,ag,mean);xbar
       1        2        3        4        5
553.3333 569.3333 610.5000 465.1667 610.6667
> contvec=c(-1,-1,-1,4,-1);contvec
[1] -1 -1 -1  4 -1
> SSc=sum(contvec^2);SSc
[1] 20
> thetahat=sum(contvec*xbar);thetahat
[1] -483.1667
> I=length(ag);I
[1] 30
> I=5;J=6;tcrit=qt(.025,I*(J-1),lower.tail=F);tcrit
[1] 2.059539
> MSE=mean(tapply(x,ag,var));MSE
[1] 4960.813
> c(thetahat-tcrit*sqrt(MSE*SSc/J),thetahat+tcrit*sqrt(MSE*SSc/J))
[1] -748.0080 -218.3253
> tcrit=qt(.005,I*(J-1),lower.tail=F);tcrit
[1] 2.787436
> c(thetahat-tcrit*sqrt(MSE*SSc/J),thetahat+tcrit*sqrt(MSE*SSc/J))
[1] -841.6102 -124.7232
> tcrit=qt(.0025,I*(J-1),lower.tail=F);tcrit
[1] 3.078199
> c(thetahat-tcrit*sqrt(MSE*SSc/J),thetahat+tcrit*sqrt(MSE*SSc/J))
[1] -879.00019  -87.33314
> tcrit=qt(.0005,I*(J-1),lower.tail=F);tcrit
[1] 3.725144
> c(thetahat-tcrit*sqrt(MSE*SSc/J),thetahat+tcrit*sqrt(MSE*SSc/J))
[1] -962.192434   -4.140899

>
>##########  COMPUTE THE  f-TEST FOR THIS CONTRAST
>
> SStheta=J*thetahat^2/(SSc);SStheta
[1] 70035.01
> f=SStheta/MSE;f
[1] 14.11765
> pvalue=pf(f,1,25,lower.tail=F);pvalue
[1] 0.0009214009
> fcrit=qf(.001,1,25,lower.tail=F);fcrit
[1] 13.8767
>
```


Absorption of moisture in Concrete Aggregates
x
450
500
550
600
650
700
1
2
3
4
5
Aggregate


Residuals vs Fitted
Residuals
Fitted values
Normal Q-Q
Standardized residuals
Theoretical Quantiles
Scale-Location
√|Standardized residuals|
Fitted values
Constant Leverage:
Residuals vs Factor Levels
Standardized residuals
ag :
Factor Level Combinations