ANOVA

MATH 2441, BCIT

Technical Mathematics for Food Technology

May 7, 2018

Analysis of Variance

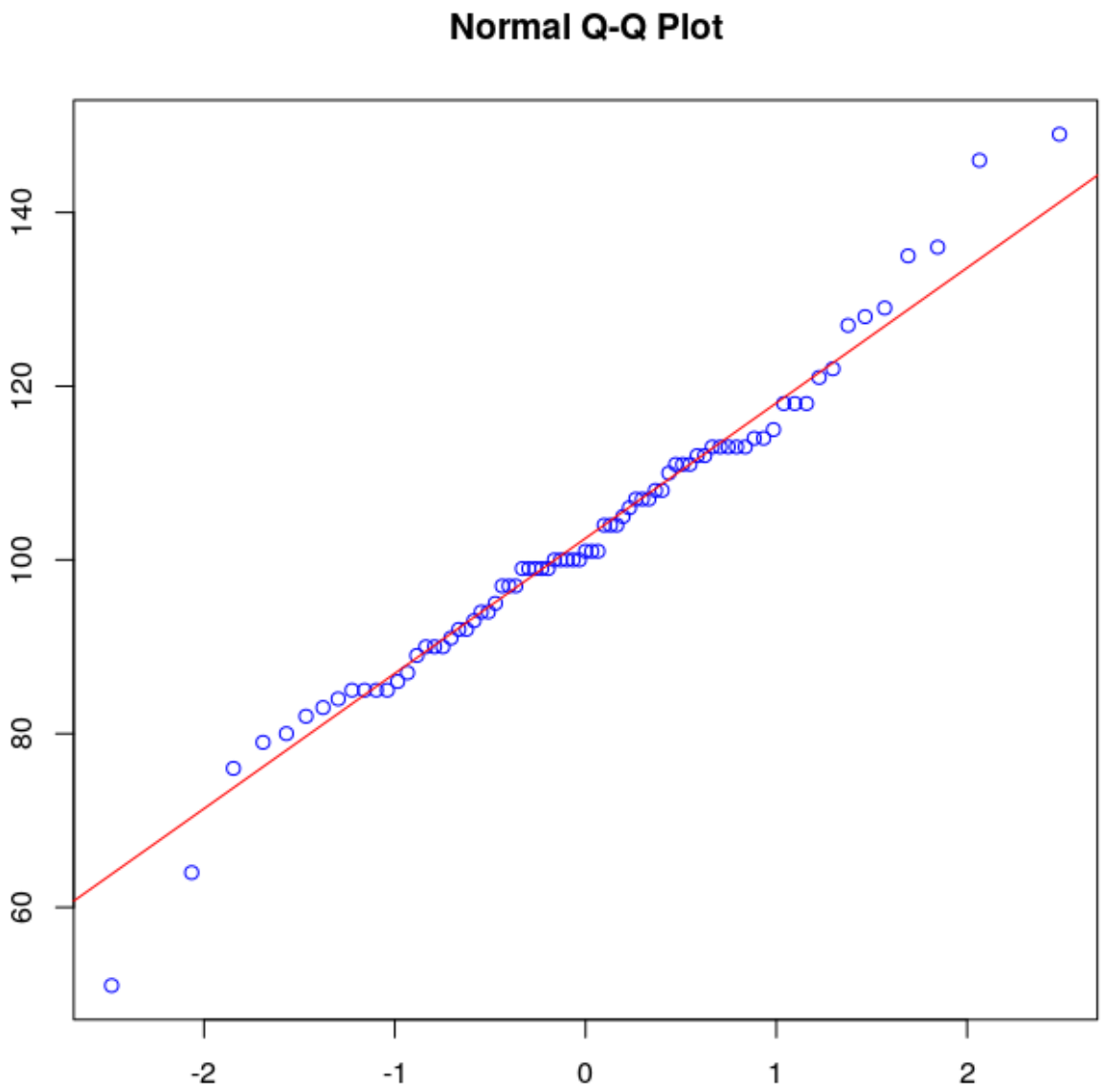
One-way analysis of variance (ANOVA) is a method of testing the equality of three or more population means by analyzing sample variances. One-way analysis of variance is used with data categorized with one factor (or treatment), so there is one characteristic used to separate the sample data into the different categories.

Consider the R Statistics code on the following page. It corresponds to the narrative in Triola, page 562. The result is a test statistic following the F-distribution and a p-value that can be compared to the significance level. This type of ANOVA is always a right-hand one-tailed test.

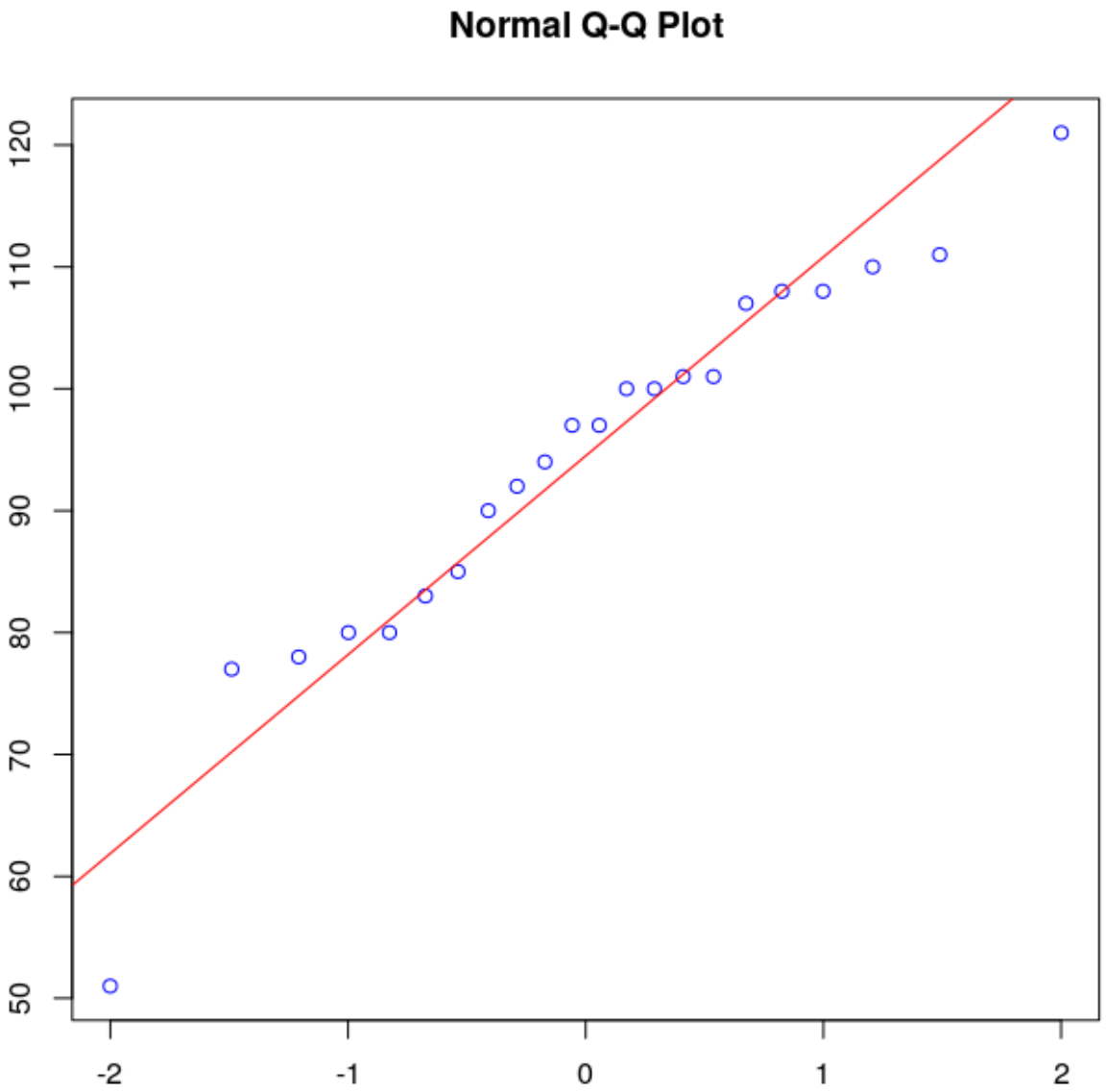
Analysis of Variance

```
l<-c(85,900,107,85,100,97,101,64,111,100,76,136,100,90,135,104,149,99,1
  07,99,113,104,101,111,118,99,122,87,118,113,128,121,111,104,51,100,113,
  82,146,107,83,108,93,114,113,94,106,92,79,129,114,99,110,90,85,94,127,
  101,99,113,80,115,85,112,112,92,97,97,91,105,84,95,108,118,86,89,100)
m<-c(78,97,107,80,90,83,101,121,108,100,110,111,97,51,94,80,101,92,100,
  77,108,85)
h<-c(93,100,97,79,97,71,111,99,85,99,97,111,104,93,90,107,108,78,95,78,
  86)
n<-c(length(l),length(m),length(h))
group<-rep(1:3,n)
y<-c(l,m,h)
data<-data.frame(y=y,group=factor(group))
fit<-lm(y~group,data)
anova(fit)
```

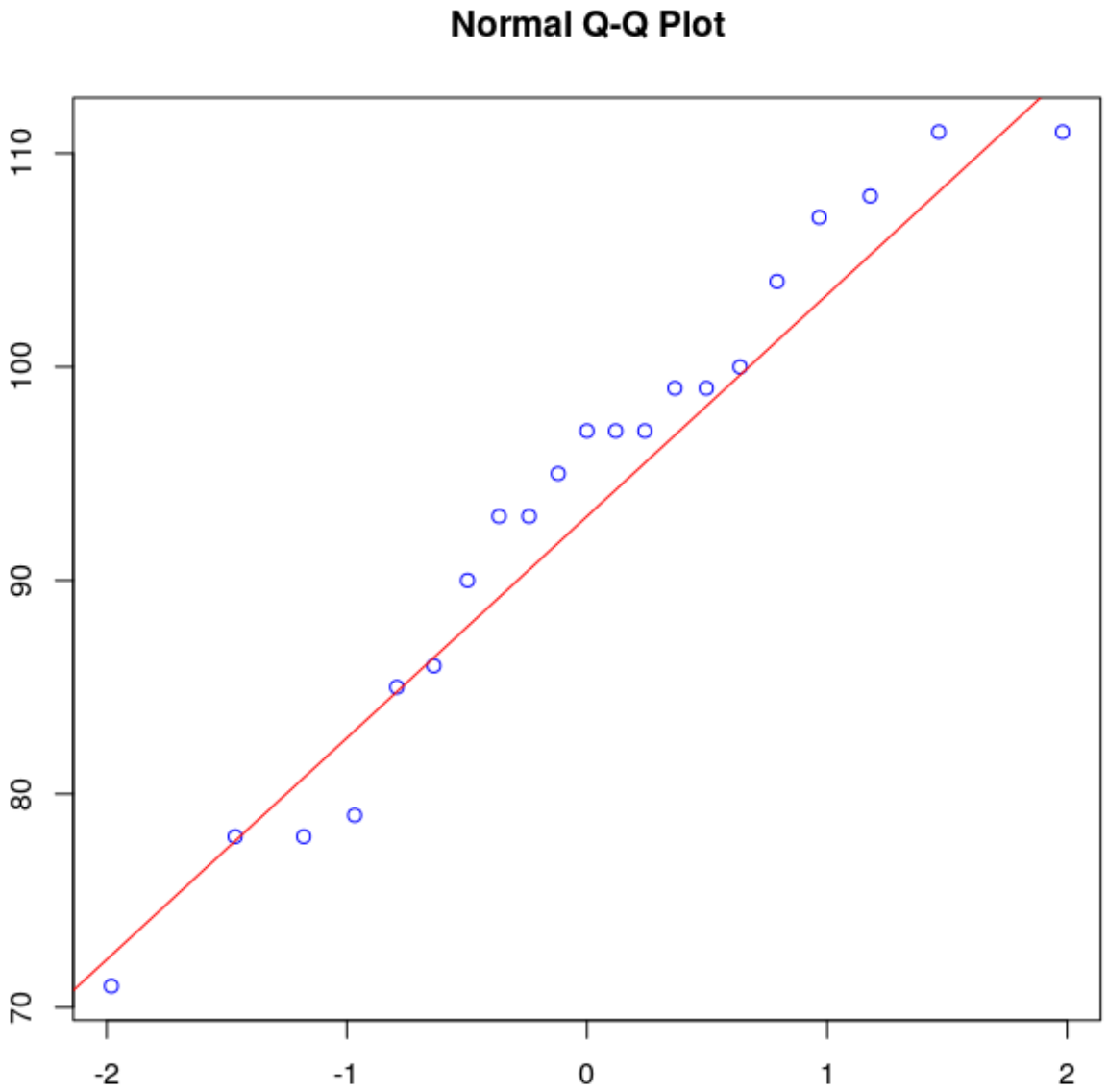
Anova: Lead and Intelligence Quotients (low)



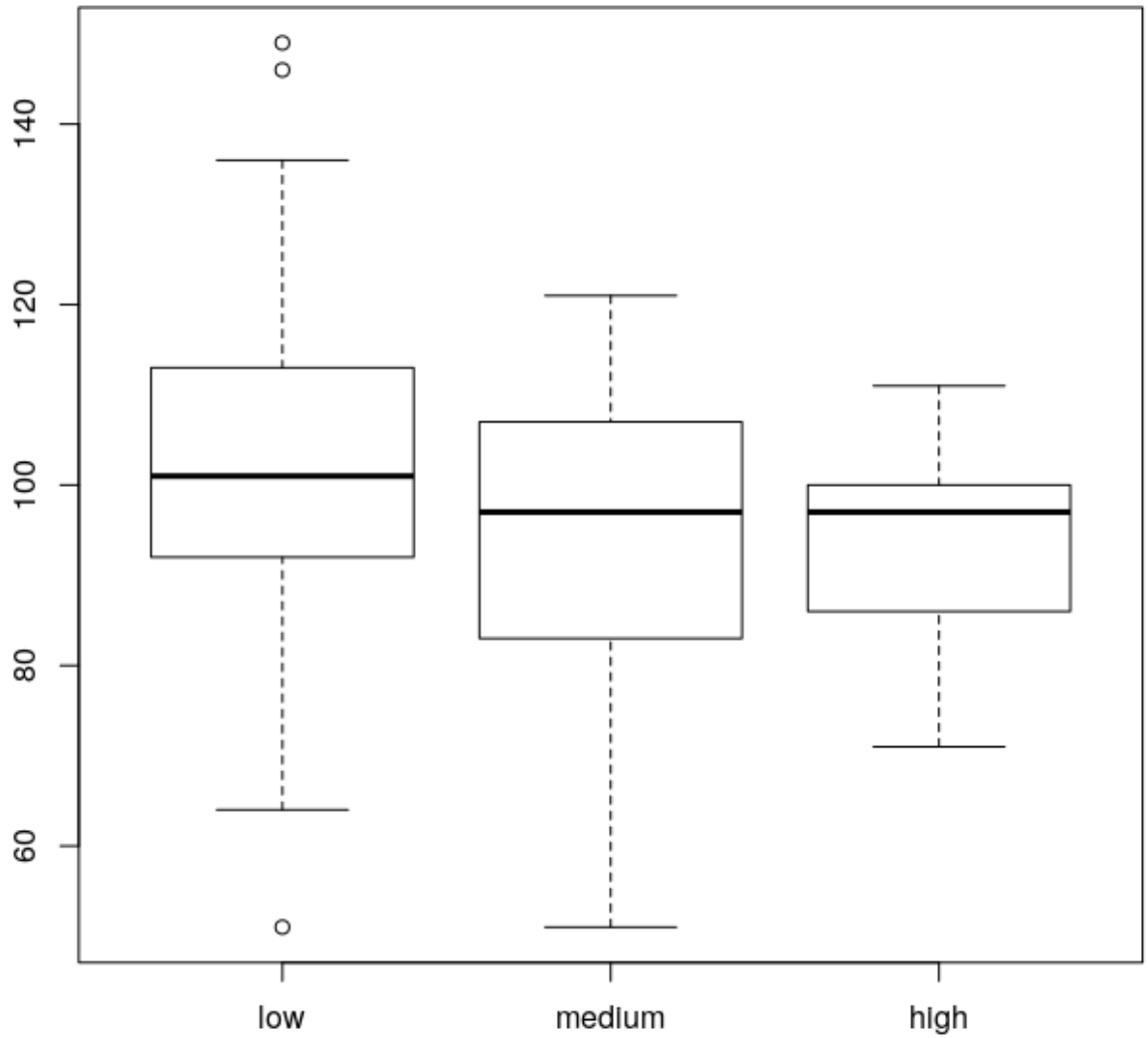
Anova: Lead and Intelligence Quotients (medium)



Anova: Lead and Intelligence Quotients (high)



Anova: Lead and Intelligence Quotients (boxplots)



Anova: Lead and Intelligence Quotients

R Statistics yields the following output:

```
Analysis of Variance Table

Response: y
           Df  Sum Sq Mean Sq F value Pr(>F)
group       2  1920.9  960.45  3.8646 0.0237 *
Residuals 117 29077.1  248.52
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1
```

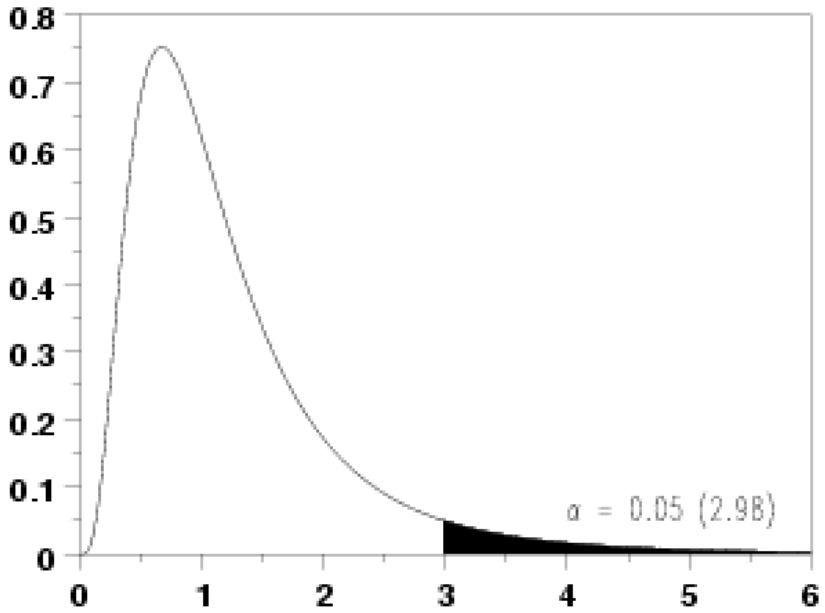
What is important to us is the test statistic, whose distribution is the F-distribution, 3.8646; and the p-value 0.0237. Since ANOVA is always one-tailed, area to the right, all you need to do is compare the p-value to the significance level. "If p is low, the NULL must go."

F-Distribution

The F-distribution depends on two different degrees of freedom, which makes using a table of critical values awkward. We will use p-values provided by technology instead. A table with critical values is here:

`http://www.itl.nist.gov/div898/handbook/eda/section3/eda3673.htm`.

The shape of the F-distribution is similar to the shape of the χ^2-distribution.



One-Way ANOVA in Microsoft Excel

For instructions, see `http://www.excel-easy.com/examples/anova.html`.

low	medium	high
85	78	93
90	97	100
107	107	97
85	80	79
100	90	97
97	83	71
101	101	111
64	121	99
111	108	85
100	100	99
76	110	97
136	111	111
100	97	104
90	51	93
135	94	90
104	80	107
149	101	108
99	92	78
107	100	95
99	77	78
113	108	86
104	85	
101		
111		
118		

Anova: Single Factor

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
Column 1	77	7893	102.5065	282.3848
Column 2	22	2071	94.13636	239.5519
Column 3	21	1978	94.19048	129.2619

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	1920.891	2	960.4455	3.864629	0.023699	3.073763
Within Groups	29077.08	117	248.522			
Total	30997.97	119				

ANOVA Exercise

Susan predicts that students will learn most effectively with a constant background sound, as opposed to an unpredictable sound or no sound at all. She randomly divides twenty-four students into three groups of eight. All students study a passage of text for 30 minutes. Those in group 1 study with background sound at a constant volume in the background. Those in group 2 study with noise that changes volume periodically. Those in group 3 study with no sound at all. After studying, all students take a 10 point multiple choice test over the material. Test the appropriate null hypothesis using one-way ANOVA at a 0.05 significance level.

constant sound	7	4	6	8	6	6	2	9
random sound	5	5	3	4	4	7	2	2
no sound	2	4	7	1	2	1	5	5

Final Exam Practice

Exercise 1: At a gas station, 40% of customers fill their tanks. Of those who fill their tanks, 80% pay with a credit card.

1. What is the probability that a customer fills their tank and pays with a credit card?
2. What is the probability that either three or four out of ten customers fill their tank and pay by credit card?
3. What is the probability that more than half of eight customers fill their tank and pay by credit card?

Final Exam Practice

Exercise 2: At a certain time in the afternoon, London Heathrow sees on average 2 planes landing per minute.

1. What is the probability of four or more planes landing in one minute?
2. What is the probability that no plane will land in a particular minute?

End of Lesson

Next Lesson: End of Term! Have a Happy Holiday!