Math 3080 § 1.	**Loblolly Example:**	Name: Example
Treibergs	**Correlation**	Feb. 26, 2014

This note explores estimating correlation of a sample taken from a bivariate normal distribution.

This data is taken from VPI Department of Forestry study as quoted in Walpole, Myers & Myers, *Probability and Statistics for Engineers and Scientists*,6th ed, Prentice Hall, Upper Saddle River NJ, 1998. Researcher wanted to compare anatomy and mechanical properties of trees. They measured 29 Loblolly Pines and recorded specific gravity (gm/cm^3) and the modulus of rupture (kilopascals).

Data Set Used in this Analysis :

```
# Math 3082        Loblolly Data    Feb. 24, 2014
#
# Data from VPI Department of Forestry study as quoted in Walpole, Myers &
# Myers, Probability and Statistics for Engineers and Scientists, 6th ed,
# Prentice Hall, Upper Saddle River NJ, 1998.
#
# Researcher wanted to compare anatomy and mechanical properties of trees.
# They measured 29 Loblolly Pines and recorded specific gravity (gm/cm^3)
# and the modulus of rupture (kilopascals)
"SpecGrav" "ModulusRupture"
.414 29186
.383 29266
.399 26215
.402 30162
.442 38867
.422 37831
.466 44576
.500 46097
.514 59698
.530 67705
.569 66088
.558 78486
.577 89869
.572 77369
.548 67095
.581 85156
.557 69571
.550 84160
.531 73466
.550 78610
.556 67657
.523 74017
.602 87291
.569 86836
.544 82540
.557 81699
.530 82096
.547 75657
.585 80490
```

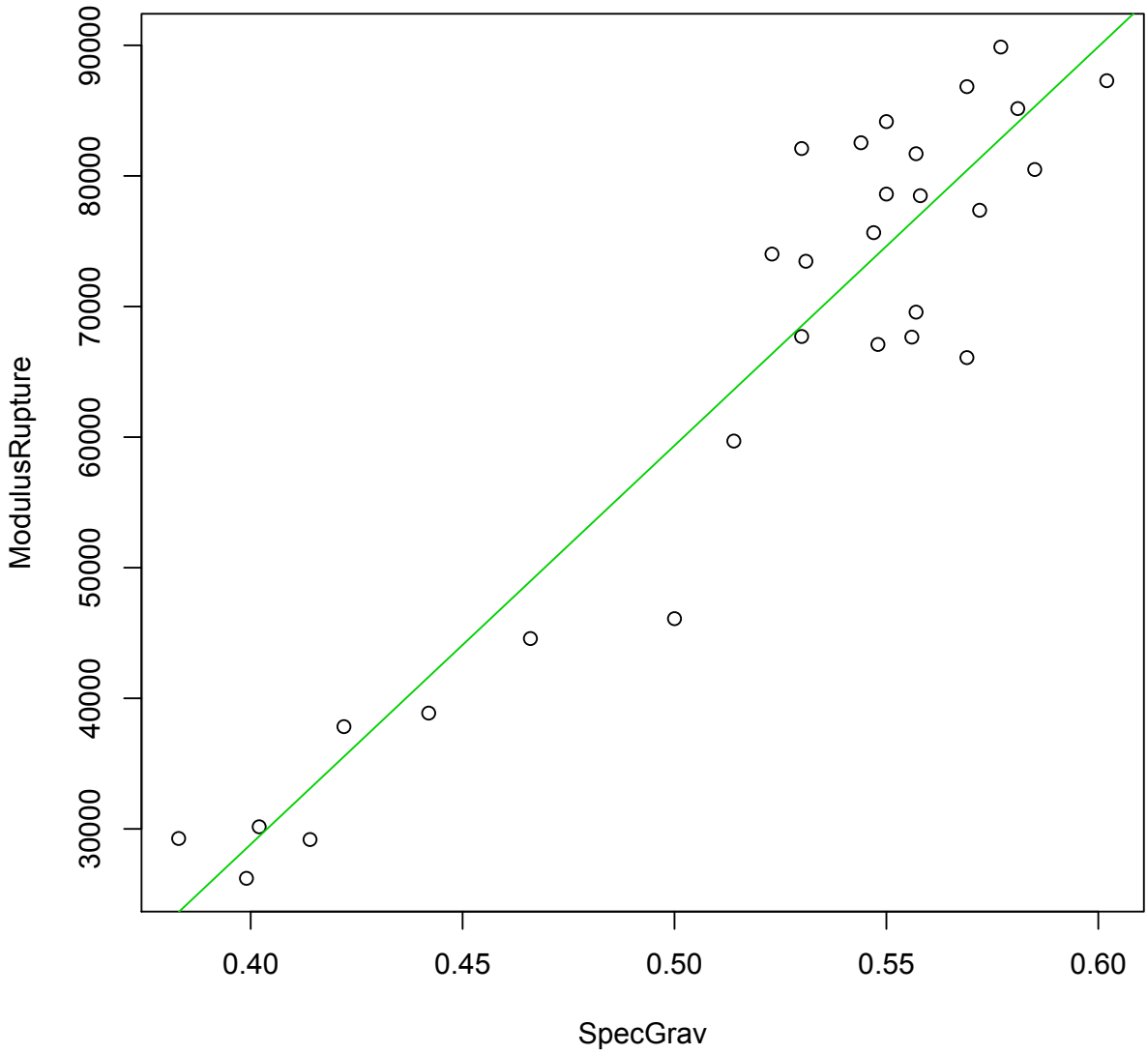
R Session:

```
R version 2.10.1 (2009-12-14)
Copyright (C) 2009 The R Foundation for Statistical Computing
ISBN 3-900051-07-0

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 1.31 (5538) powerpc-apple-darwin8.11.1]

[Workspace restored from /Users/andrejstreibergs/.RData]

> tt=read.table("M3082DataLoblolly.txt",header=T)
Error in file(file, "rt") : cannot open the connection
In addition: Warning message:
In file(file, "rt") :
  cannot open file 'M3082DataLoblolly.txt': No such file or directory
> tt=read.table("M3082DataLoblolly.txt",header=T)
> attach(tt)
> tt
   SpecGrav ModulusRupture
1     0.414          29186
2     0.383          29266
3     0.399          26215
4     0.402          30162
5     0.442          38867
6     0.422          37831
7     0.466          44576
8     0.500          46097
9     0.514          59698
10    0.530          67705
11    0.569          66088
12    0.558          78486
13    0.577          89869
14    0.572          77369
15    0.548          67095
16    0.581          85156
17    0.557          69571
18    0.550          84160
```

```
19    0.531          73466
20    0.550          78610
21    0.556          67657
22    0.523          74017
23    0.602          87291
24    0.569          86836
25    0.544          82540
26    0.557          81699
27    0.530          82096
28    0.547          75657
29    0.585          80490

> ############  RUN ANOVA ON DATA ################################
> f1=lm(ModulusRupture~SpecGrav)
> summary(f1); anova(f1)

Call:
lm(formula = ModulusRupture ~ SpecGrav)

Residuals:
   Min     1Q Median     3Q    Max
-14335  -3970    732   4940  13582

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   -93319      10808  -8.634 3.00e-09 ***
SpecGrav      305346      20640  14.794 1.80e-14 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

Residual standard error: 6930 on 27 degrees of freedom
Multiple R-squared: 0.8902,Adjusted R-squared: 0.8861
F-statistic: 218.9 on 1 and 27 DF,  p-value: 1.797e-14

Analysis of Variance Table

Response: ModulusRupture
          Df     Sum Sq    Mean Sq F value    Pr(>F)
SpecGrav   1 1.0511e+10 1.0511e+10  218.87 1.797e-14 ***
Residuals 27 1.2966e+09 4.8022e+07
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

> ###############  SCATTER PLOT WITH REGRESSION LINE  ###############
> plot(ModulusRupture~SpecGrav)
> abline(f1,col=3)
```

```
> ################  TEST H0: RHO=0 VS Ha:  RHO NE 0  ################
> r=cor(ModulusRupture,SpecGrav); r; r^2
[1] 0.943497
[1] 0.8901864

> cor.test(ModulusRupture,SpecGrav)

Pearson's product-moment correlation

data:  ModulusRupture and SpecGrav
t = 14.7943, df = 27, p-value = 1.799e-14
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.8819758 0.9734028
sample estimates:
     cor
0.943497

> ###################  REDO BY HAND  ################################
> # cor test by hand
> cov(ModulusRupture,SpecGrav)
[1] 1229.367
> xbar=mean(SpecGrav);xbar
[1] 0.519931
> ybar=mean(ModulusRupture);ybar
[1] 65439.86
> n=length(SpecGrav); n
[1] 29
> Sxx=sum(SpecGrav^2)-xbar^2/n;Sxx
[1] 7.94293
> Sxx=sum(SpecGrav^2)-n*xbar^2;Sxx
[1] 0.1127319
> Sxy=sum(SpecGrav*ModulusRupture)-n*xbar*ybar;Sxy
[1] 34422.28
> Syy=sum(ModulusRupture^2)-n*ybar^2;Syy
[1] 11807324805
> Sxy/sqrt(Sxx*Syy)
[1] 0.943497
```

```
> ###############  t - TEST FOR CORRELATION  ########################

> t=r/sqrt((1-r^2)/(n-2));t
[1] 14.79430
> qt(.025,n-2,lower.tail=F)
[1] 2.051831
> 2*pt(t,n-2,lower.tail=F)
[1] 1.796901e-14

> ############### Z-test H0: rho=.9 Ha: rho>.9  ####################

> r0=.9
> z=sqrt(n-3)*(log((1+r)/(1-r))-log((1+r0)/(1-r0)))/2;z
[1] 1.513159
> qnorm(.05,lower.tail=F)
[1] 1.644854
> pnorm(z,lower.tail=F)
[1] 0.0651196

> ###################### 2-sided CI on rho  ##########################

> za=qnorm(.025,lower.tail=F);za
[1] 1.959964
> v=log((1+r)/(1-r))/2;v
[1] 1.768974
> cl=v-za/sqrt(n-3);cu=v+za/sqrt(n-3);c(cl,cu)
[1] 1.384594 2.153355
> (exp(2*cl)-1)/(exp(2*cl)+1);tanh(cl)
[1] 0.8819758
[1] 0.8819758
> c(tanh(cl),tanh(cu))
[1] 0.8819758 0.9734028
>
```


ModulusRupture
30000
40000
50000
60000
70000
80000
90000
0.40
0.45
0.50
0.55
0.60
SpecGrav