Stat 5309 Lab 1

Tom Wilson

January 18, 2019

1 Data: Bacteria with Packages

a

Set up th dataframe

```r
bacteria <- data.frame(package = rep(c("plastic","vacuum","1%CO","100%CO2"),each=3),
                       log_count = c(7.66,6.98,7.80,
                                     5.26,5.44,5.80,
                                     7.41,7.33,7.04,
                                     3.51,2.91,3.66))


bacteria %>% kable()
```

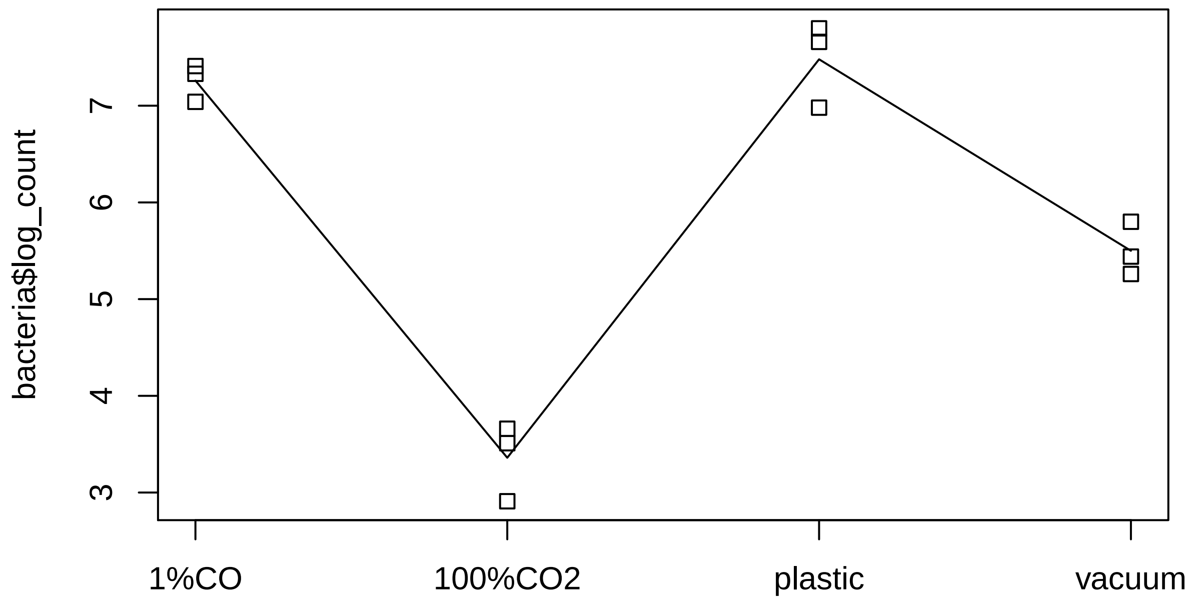
package	log_count
plastic	7.66
plastic	6.98
plastic	7.80
vacuum	5.26
vacuum	5.44
vacuum	5.80
1%CO	7.41
1%CO	7.33
1%CO	7.04
100%CO2	3.51
100%CO2	2.91
100%CO2	3.66

b

perform a stripchart with line connecting means of logcount vs package

```r
stripchart(bacteria$log_count~bacteria$package,vertical = TRUE)

lines(tapply(bacteria$log_count,bacteria$package , mean))
```



c

build a linear model using aov() response as logcount. Do a summary.lm() and summary.aov()

```r
fit <- lm(data=bacteria,formula = log_count ~ package)
summary_of_fit <- summary(fit)
anova_of_fit <- anova(fit)
anova_of_fit
```

```
## Analysis of Variance Table
##
## Response: log_count
##           Df Sum Sq Mean Sq F value    Pr(>F)
## package    3 32.873 10.9576  94.584 1.376e-06 ***
## Residuals  8  0.927  0.1159
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
aov(fit)
```

```
## Call:
##    aov(formula = fit)
##
## Terms:
##                 package Residuals
## Sum of Squares  32.8728    0.9268
```

```
## Deg. of Freedom       3         8
##
## Residual standard error: 0.3403674
## Estimated effects may be unbalanced
```

d

perform a bartlett test of equal variances

```r
bartlett.test(bacteria$log_count~bacteria$package)
```

```
##
##  Bartlett test of homogeneity of variances
##
## data:  bacteria$log_count by bacteria$package
## Bartlett's K-squared = 1.2079, df = 3, p-value = 0.7511
```

e

perform a multiple comparison of treatment mean using TukeyHSD()

```r
TukeyHSD(aov(fit), conf.level=0.95)
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = fit)
##
## $package
##                   diff       lwr       upr     p adj
## 100%CO2-1%CO    -3.90 -4.789962 -3.010038 0.0000031
## plastic-1%CO     0.22 -0.669962  1.109962 0.8563618
## vacuum-1%CO     -1.76 -2.649962 -0.870038 0.0010160
## plastic-100%CO2  4.12  3.230038  5.009962 0.0000020
## vacuum-100%CO2   2.14  1.250038  3.029962 0.0002639
## vacuum-plastic  -1.98 -2.869962 -1.090038 0.0004549
```

2 Data: Tensile strength of Portland Cement

a

Set up a data frame with variables mixing and strength

```r
cement <- data.frame(mixing = rep(c(1,2,3,4),each=4),
                     strength = c(3129,3000,2865,2890,
                                  3200,3300,2975,3150,
                                  2800,2900,2985,3050,
                                  2600,2700,2600,2765))

cement %>% kable()
```

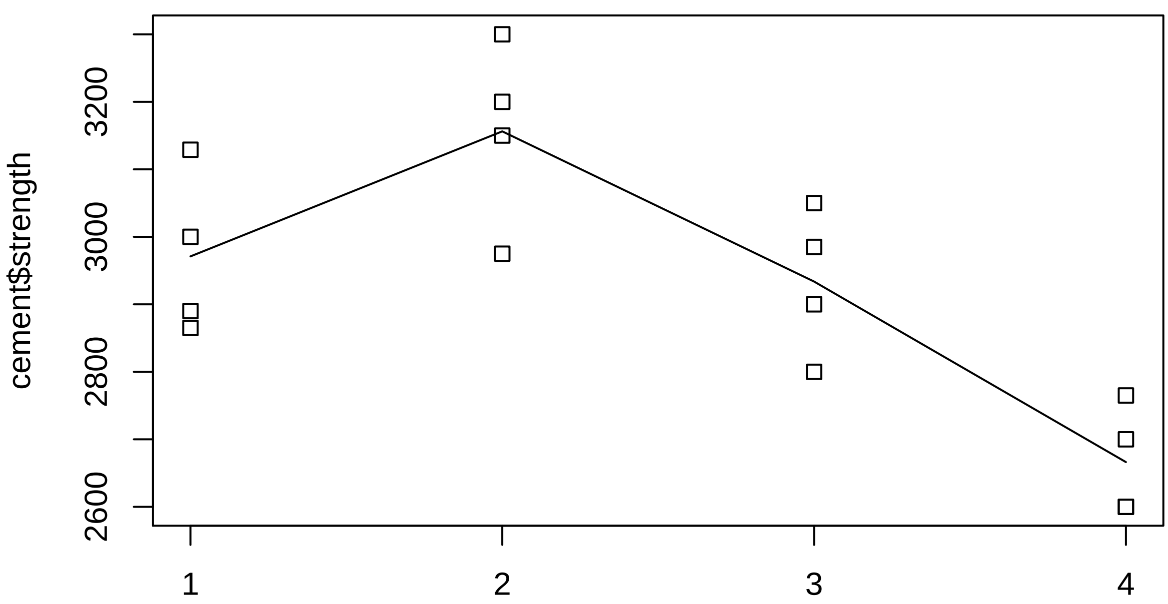
mixing	strength
1	3129
1	3000
1	2865
1	2890
2	3200
2	3300
2	2975
2	3150
3	2800
3	2900
3	2985
3	3050
4	2600
4	2700
4	2600
4	2765

b

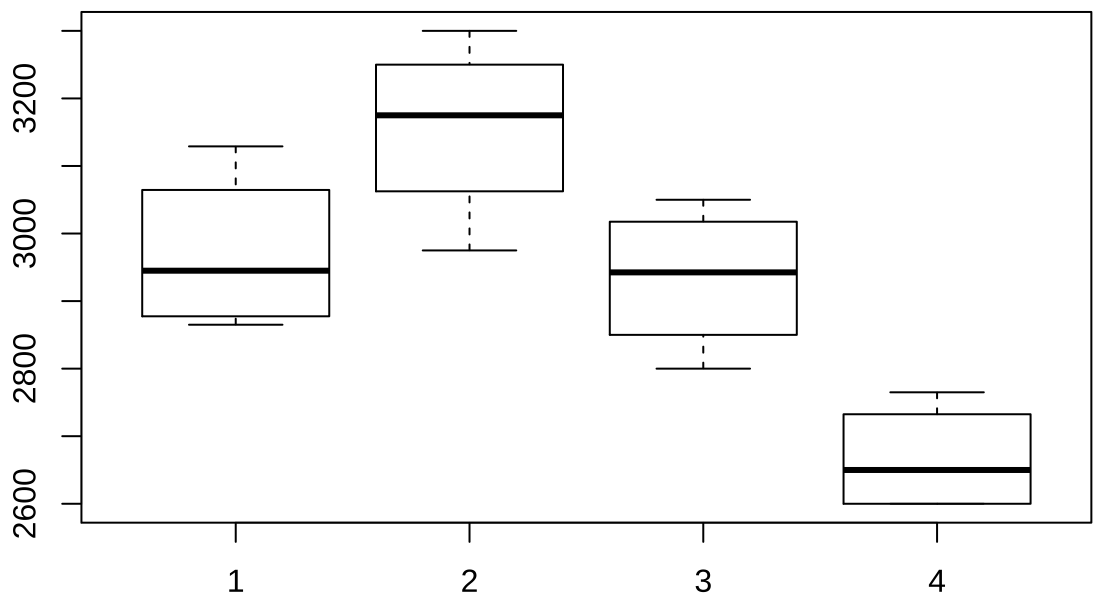
Perform a stripchart and boxplot

```
stripchart(cement$strength~cement$mixing,vertical = TRUE)

lines(tapply(cement$strength,cement$mixing , mean))
```



```r
boxplot(cement$strength~cement$mixing)
```



c

use Fisher Least Significant Difference (LSD) with $\alpha = 0.05$ to make a comparison

```r
cement_fit <- lm(cement$strength~cement$mixing)
cement_fit_anova <- anova(cement_fit)
MSerror <- cement_fit_anova$`Mean Sq`[2]
#LSD.test(g, "trt", MSerror)
#LSD.test(g, "trt", MSerror, console=T)
```

d

Test the hypothesis that mixing techniques affect the strength of the cement. Use $\alpha = 0.05$. What test do you use? Conclusion?

3

A manufacture of television sets is interested in the effect on tube conducivity of four different types of ocutaing for color picture tubes. the following conductivity data are obtained.

```r
television <- data.frame(coating_type = rep(c(1,2,3,4),each=4),
                         conductivity = c(143,141,150,146,
                                          152,149,137,143,
                                          134,136,132,127,
                                          129,127,132,129))

television %>% kable()
```

coating_type	conductivity
1	143
1	141
1	150
1	146
2	152
2	149
2	137
2	143
3	134
3	136
3	132
3	127
4	129
4	127
4	132
4	129

a

Is there a difference in conductivity due to coating type? Use $\alpha = 0.05$

b

Estimate the overal mean and the treatment effects.

c

Compute a 95% confidence interval estimate of th mean of coating type 4.

Compute a 99% confidence interval estimate of the mean difference between coating types 1 and 4.

d

Test all pairs of means using the Fisher LSD method with $\alpha = 0.05$

e

Use the graphical method discussed in Section 3-5.3 to compare the means. Which coating type produces the highest conductivity?

f

Assuming that coating type is currently in use, what are your recommendations to the manufacturer? We wish to minimize conductivity.