# HW6_Lin

## IST 772 Homework 6

**Due November 18, 2021 at 8:00AM EDT**

Homework 5 by Nora Lin: I produced the material below with no assistance.

### Excercise 1 p.117:

Indenpdent variables are used to predict the dependent variables.Therefore in the data set **InsectSprays**, the outcome variable(dependent variable) is the count of insects killed and the predictor varialbe(independent variable) is the type of spray used. There are **72** total observations, **6** sprays each with **12** observations.

### Excercise 2 p.117:

```
aov_SprayResults <- aov(count~spray, data=InsectSprays)
summary(aov_SprayResults)
```

Mean sq for sprays = **533.8** and mean sq for residuals is **15.4**.

```
##              Df Sum Sq Mean Sq F value Pr(>F)
## spray         5   2669   533.8    34.7 <2e-16 ***
## Residuals    66   1015    15.4
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The degrees of freedom is **5**. This means that among the sprays only **5** can vary freely which means that **66** degress of freedom within groups. The sum sq is a calculation of variability. The sm of squares for the between group is **2669** and that of the within group is **1015**. The **15.4** mean sq for Residuals represents the within group variance while **533.8** mean sq for spray represents between group variance. Between group variance shows the variance in the means while within group measures the spread from each data point from the raw data. The F-value a ratio of the mean of squares of the between group and the mean of squares of the within groups which is **34.7**. The Pr(>F) is the probability of the F-ratio.

### Excercise 3 p.117:

Since the F-ratio is the mean of squares of the between group overthe the mean of squuares of the within group. In this case, the F-ratio is **533.8/15.4= 34.662**. The probability of the F-ratio is the significance level. Since our Pr(>F) is really smaller than any alpha value we would reject the null hypothesis.

## Excercise 4 p.117:

Since there are six groups each with n=12 observations, that's **72** observations total. Among **6** groups only **5** can vary freely so that means that the degress ofo freedom is **5** and the residuals would be **72-5** is **67** but since we loose one degree of freedom for calculating the grand mean, the residuals degrees of freedomo is **66**.

## Excercise 5 p.117:

Our null hypothesis is that there is no difference between the sprays. The alternative hypothesis is that there is no difference between the sprays. Since our $Pr(>F)$ is smaller than any alpha value, we would rejec the null hypothesis.

## Excercise 6 p.117:

```r
#install.packages("BayesFactor")
#install.packages("BEST")
library("BayesFactor")
```

```
## Loading required package: coda
```

```
## Loading required package: Matrix
```

```
## ************
## Welcome to BayesFactor 0.9.12-4.2. If you have questions, please contact Richard Morey (richarddmorey
##
## Type BFManual() to open the manual.
## ************
```

```r
library("BEST")
```

```
## Loading required package: HDInterval
```

```r
sprayBayes <-  anovaBF(count~spray, data=InsectSprays)
sprayBayes
```

```
## Bayes factor analysis
## --------------
## [1] spray : 1.506706e+14 ±0%
##
## Against denominator:
##   Intercept only
## ---
## Bayes factor type: BFlinearModel, JZS
```

```r
sprayMCMC <- posterior(sprayBayes, iterations=10000)
summary(sprayMCMC)
```

```
##
## Iterations = 1:10000
## Thinning interval = 1
## Number of chains = 1
## Sample size per chain = 10000
##
## 1. Empirical mean and standard deviation for each variable,
##    plus standard error of the mean:
##
##            Mean     SD Naive SE Time-series SE
## mu        9.501 0.4756 0.004756       0.004756
## spray-A   4.824 1.0476 0.010476       0.010702
## spray-B   5.630 1.0573 0.010573       0.010573
## spray-C  -7.143 1.0538 0.010538       0.010880
## spray-D  -4.426 1.0483 0.010483       0.011203
## spray-E  -5.779 1.0499 0.010499       0.010499
## spray-F   6.895 1.0366 0.010366       0.010822
## sig2     16.182 2.9333 0.029333       0.034664
## g_spray   3.434 3.4598 0.034598       0.036147
##
## 2. Quantiles for each variable:
##
##             2.5%    25%    50%    75%  97.5%
## mu        8.5631  9.185  9.501  9.820 10.448
## spray-A   2.7743  4.126  4.816  5.519  6.900
## spray-B   3.4922  4.922  5.643  6.340  7.668
## spray-C  -9.2100 -7.839 -7.144 -6.442 -5.055
## spray-D  -6.4879 -5.124 -4.425 -3.730 -2.392
## spray-E  -7.8141 -6.467 -5.783 -5.089 -3.722
## spray-F   4.9021  6.166  6.892  7.601  8.954
## sig2     11.5109 14.090 15.827 17.894 22.860
## g_spray   0.8258  1.653  2.503  4.009 11.297
```

```
plot(sprayMCMC[,"mu"], main=NULL)
```
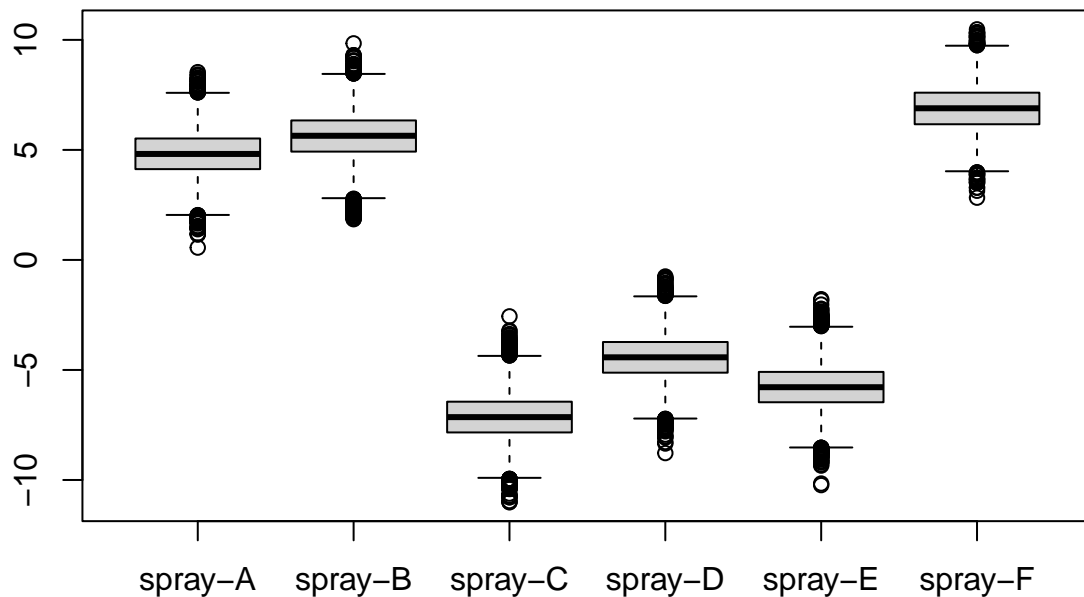
## Trace of var1
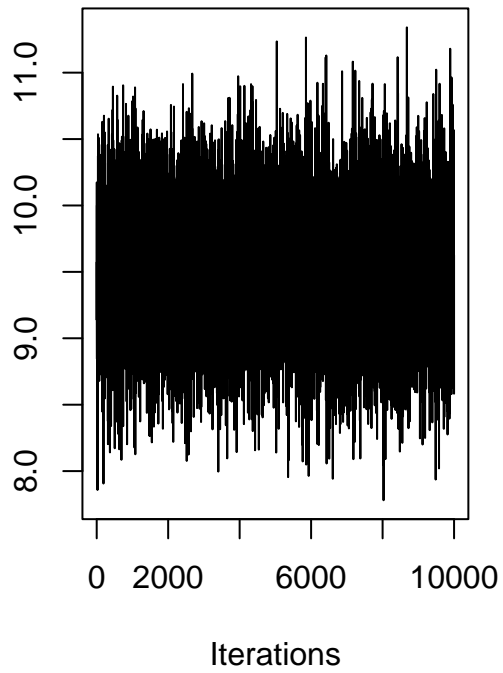


## density.default(x = y, width = widt
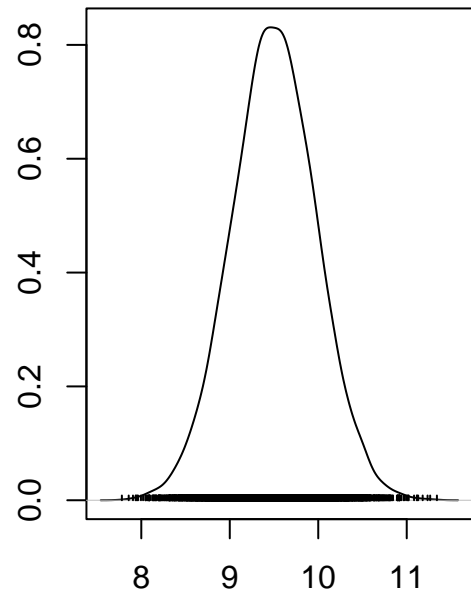


```
boxplot(as.matrix(sprayMCMC[, 2:7]))
```

4

#### In the trace of var1, we can see that most of the steps converge towards the center. This appears to be a symmetrical rather than bell curve distribution. From the boxplot, we can see that Spray A, B and F seem to be superior than the spray C, D, and E. This is because they result in a higher in y-axis value. This shows that we should expect there to be a difference in variance between our sprays and thus demonostrates BayesFactor.There are also no overlapping boxplots.

## Excercise 7 p.117:

```
C_observations <- InsectSprays[InsectSprays$spray=='C',1]
F_observations <- InsectSprays[InsectSprays$spray=='F',1]
plot(BESTmcmc(C_observations,F_observations))
```

```
## Waiting for parallel processing to complete...done.
```

# Difference of Means

mean = −14.5

100% < 0 < 0%

95% HDI

−18.9   −10

$$\mu_1 - \mu_2$$
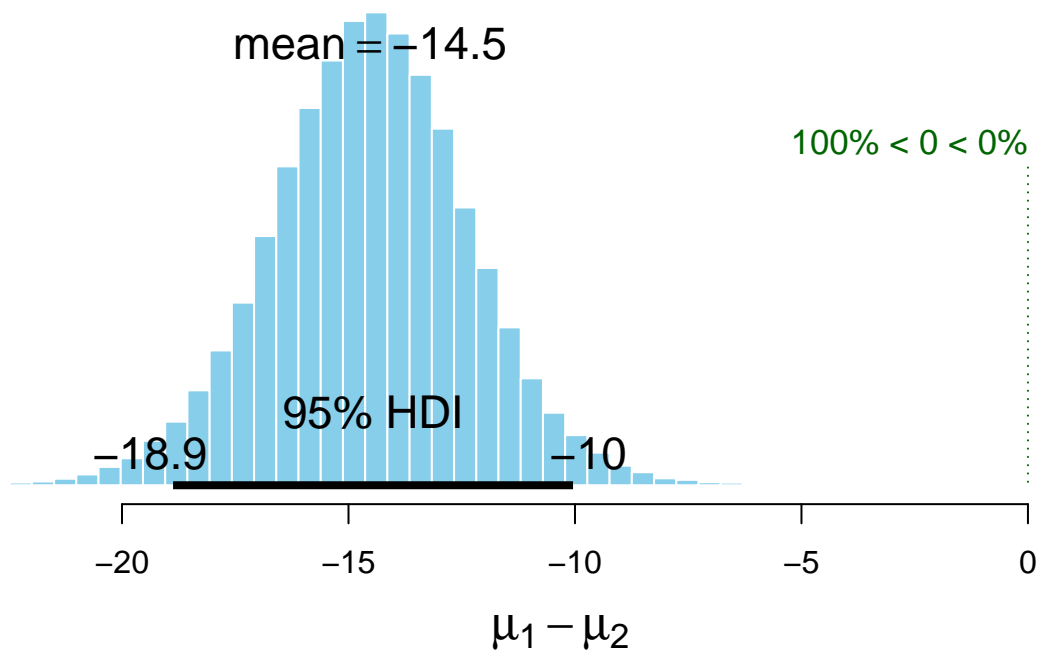
```
C_and_F <- InsectSprays[InsectSprays$spray == 'C' | InsectSprays$spray == "F",]
spray_BFOut <- anovaBF(count ~ spray, data=C_and_F)
summary(spray_BFOut)
```

```
## Bayes factor analysis
## --------------
## [1] spray : 90005.78 ±0%
##
## Against denominator:
##   Intercept only
## ---
## Bayes factor type: BFlinearModel, JZS
```