

Homework 6

Maya Mileva

due date: Nov 14th, 2019

I did this homework by myself, with help from the book and the professor.

```
## Run these functions to get a clean test of homework code
#dev.off() # Clear the graph window
cat('\014') # Clear the console
```

```
rm(list=ls()) # Clear user objects from the environment
```

Exercises

1. The data sets package (installed in R by default) contains a data set called `InsectSprays` that shows the results of an experiment with six different kinds of insecticide. For each kind of insecticide, $n = 12$ observations were conducted. Each observation represented the count of insects killed by the spray. In this experiment, what is the dependent variable (outcome) and what is the independent variable? What is the total number of observations?

```
## Data exploration
str(InsectSprays) # 72 obs. of 2 variables
```

```
## 'data.frame': 72 obs. of 2 variables:
## $ count: num 10 7 20 14 14 12 10 23 17 20 ...
## $ spray: Factor w/ 6 levels "A","B","C","D",...: 1 1 1 1 1 1 1 1 1 1 ...
```

```
summary(InsectSprays)
```

```
##      count      spray
## Min.   : 0.00   A:12
## 1st Qu.: 3.00   B:12
## Median : 7.00   C:12
## Mean    : 9.50   D:12
## 3rd Qu.:14.25   E:12
## Max.    :26.00   F:12
```

```
head(InsectSprays, n=3)
```

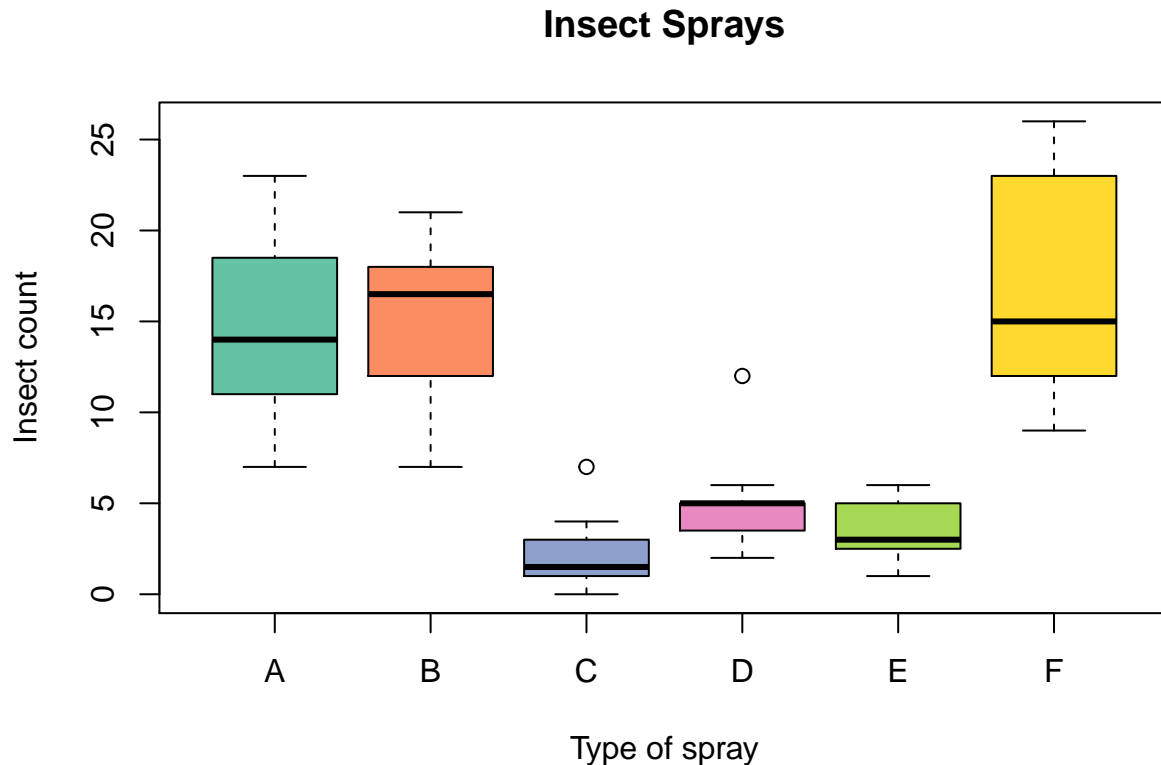
```
##    count spray
## 1     10     A
## 2      7     A
## 3     20     A
```

```
unique(InsectSprays$spray) # A B C D E F - kinds of sprays
```

```
## [1] A B C D E F
## Levels: A B C D E F
```

```
# Boxplots of the distribution of the data
```

```
boxplot(count ~ spray, data = InsectSprays,
        xlab = "Type of spray"
        , ylab = "Insect count"
        , main = "Insect Sprays"
        , varwidth = TRUE
        , col = brewer.pal(6, "Set2"))
```



Experiment with six different kinds of insecticide was performed. For each kind of insecticide, $n = 12$ observations were conducted. Each observation represented the count of insects killed by the spray. In this experiment the dependent variable(outcome) is the count of the insects killed and the independent variables(predictors) are the sprays used. The total number of observations is 72.

2. After running the `aov()` procedure on the `InsectSprays` data set, the “Mean Sq” for spray is 533.8 and the “Mean Sq” for Residuals is 15.4. Which one of these is the between groups variance and which one is the within- groups variance? Explain your answers briefly in your own words.

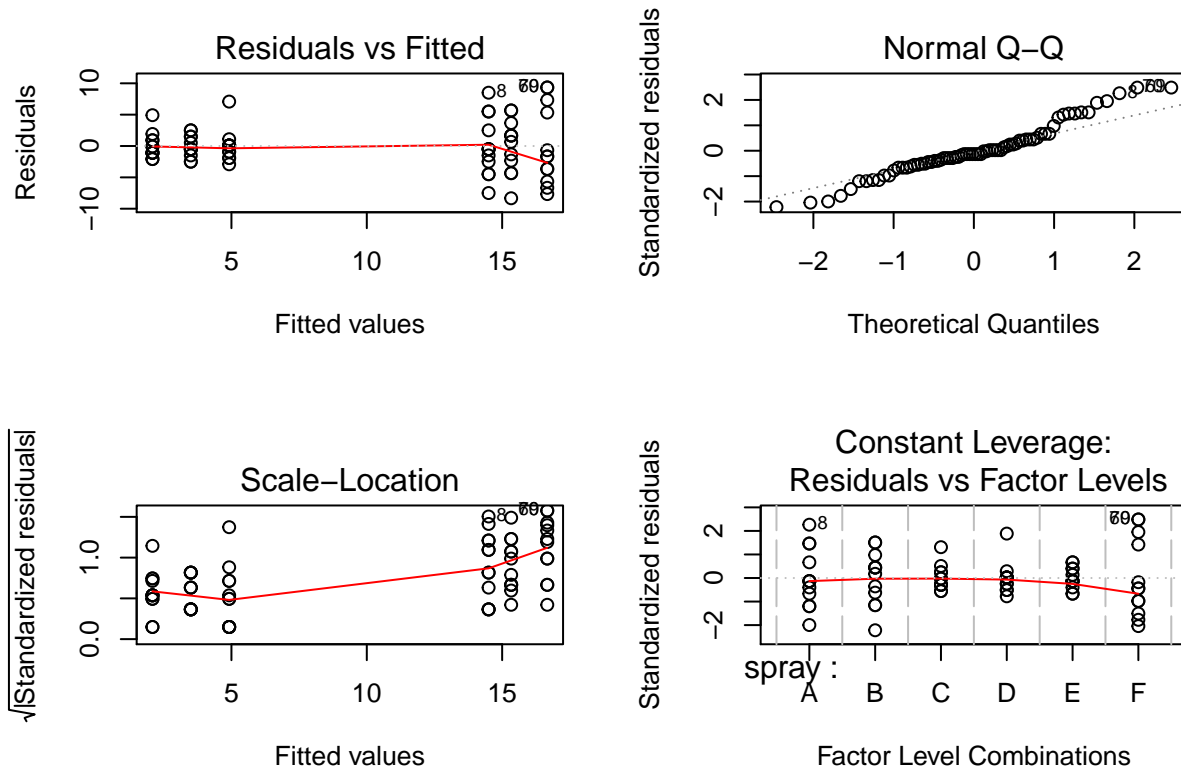
```
## Conventional ANOVA
```

```
aovOut <- aov(count ~ spray, data=InsectSprays)
summary(aovOut) # statistically significant at  $p < .05$ 
```

```
##           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
```

```
par(mfrow = c(2, 2))
plot(aovOut)
```



Mean Sq is the sum of the squares divided by the degree of freedom, aka variance.

Between the group variance is the variation among the group means (every element variance). In this case that is 533.8.

We obtain within group variance by mixing together all of the data from the samples and obtaining the variance among these scores(whole variance). In the above example the value of the within group variance is 15.4.

3. Based on the information in question 2 and your response to that question, calculate an F-ratio by hand or using a calculator. Given everything you have earned about F-ratios, what do you think of this one? Hint: If you had all the information you needed for a Null Hypothesis Significance Test, would you reject the null? Why or why not?

```
## Calculating F-ratio
533.8/15.4 # ~34.7
```

```
## [1] 34.66234
```

The F-ratio is a ratio of the mean squares from the first line (between groups) and the mean squares of the second line (within groups), that is 533.8 divided by 15.4. If we draw multiple groups from the same

population, the scaled between-group variance and the within group variance will be about equal to each other.

For ANOVA result to be statistically significant, F must substantially exceed 1 (34.7 in this case). The $\Pr(>F)$ is the significance level, the probability of finding value of F larger than the observed value under the assumption that all of the data were samples from the same population (the null hypothesis). For F to be significant, according to the logic of the null hypothesis significance test, the value of $\Pr(>F)$ **must be less than alpha level** set (usually .05, .01), before conducting the test. In this case $p=2e-16$ is a really small number, smaller than .05 so we can reject the null hypothesis, where the null hypothesis was that all 6 groups were samples from the same population (have equal means). The samples were from different underlying population - F-value was large and p-value small.

4. Continuing with the InsectSprays example, there are six groups where each one has n = 12 observations. Calculate the degrees of freedom between groups and the degrees of freedom within groups. Explain why the sum of these two values adds up to one less than the total number of observations in the data set.

In the original data set there are 72 observations. Then we lost 1 df when we calculated the overall mean of the data set which leave us with $df=71$. Out of that total we borrow 5 degrees of freedom to represent the between group variance. We have 6 groups and need only 5 degrees of freedom, because if we consider that we know the grand mean and the mean of the 5 groups, the sixth mean is no longer free to vary (can be easily calculated). If we have k groups, the degree of freedom between groups is $k-1$. With total of 71, and 5 used for between-groups, the remaining degrees of freedom are allocated to the within-group variance($df=66$).

Together, between-groups df and within-groups df always add up to the total $df(66+5 = 71)$. The shape of F depends on both of those df.

5. Use R or R- Studio to run the aov() command on the InsectSprays data set. You will have to specify the model correctly using the “~” character to separate the dependent variable from the independent variable. Place the results of the aov() command into a new object called insectResults. Run the summary() command on insectResults and interpret the results briefly in your own words. As a matter of good practice, you should state the null hypothesis, the alternative hypothesis, and what the results of the null hypothesis significance test lead you to conclude.

```
## Conventional ANOVA
insectResults <- aov(count ~ spray, data=InsectSprays)
summary(insectResults) # statistically significant at  $p<.05$ 
```

```
##           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
```

Df: degree of freedom - elements of a set that are free to vary once some statistics have been calculated; from data set of 72 we lose 1 degree of freedom for calculating the grand mean; among 6 groups means only 5 can vary freely; this leaves 66 df within groups

Sum Sq: sum of squares - a raw initial calculation of variability; the first line is the “between groups” sum of squares; the second line is the “within groups” sum of squares

Mean Sq: variance - the first line is the “between groups” variance as discussed above, the second line is the “within groups” variance “between groups” variance - the variability among the group means, adjusted for sample size “within groups” variance - the variability of all observations from all groups, relative to the grand mean

F value: the F-Ratio - ratio of the mean squares from the first line (between groups) and the mean squares of the second line (within groups) - (533.8/15.4)

Pr(>F): the probability of larger F-ratio - when examining the distribution of F-Ratios for the degree of freedom appearing in this table, this is the probability of finding an F value at least this high(in this example 34.7); The F-distribution has only one positive tail, so to reject the null hypothesis, we must look for extreme values of F that appear in the tail of the distribution

ANOVA should show no meaningful differences among means and Null Hypothesis Significance test— should be “fail to reject” the null.

The p-value from the table is less than .05 ($p=2e-16$ ***) so we reject the null, therefore is statistically significant. We can report that as $F(5,66)=34.7$, $p<.05$. According to the null hypothesis the mean outcome is the same across all groups; all 6 groups were samples from the same population such that any variation among means was attribute to sampling error. The Alternative hypothesis is that at least one pair of means are different from each other. When the null hypothesis of F is statistically significant(in our example $p<.05$), this result says nothing about which means are actually different from one another. We can calculate eta-squared effect size ($2669/(2669+1015) = 0.72$) suggesting that spray type explained about 72% of the variance in count.

If we want to know exactly which one, we have to conduct a “post hoc” test.

6. Load the BayesFactor package and run the anovaBF() command on the InsectSprays data set. You will have to specify the model correctly using the “~” character to separate the dependent variable from the independent variable. Produce posterior distributions with the posterior() command and display the resulting HDIs. Interpret the results briefly in your own words, including an interpretation of the BayesFactor produced by the grouping variable. As a matter of good practice, you should state the two hypotheses that are being compared. Using the rules of thumb offered by Kass and Raftery (1995), what is the strength of this result?

```
## Bayesian ANOVA
set.seed(1234)

## Calc Bayes Factors
bayesOut <- anovaBF(count ~ spray, data=InsectSprays)

summary(bayesOut)
```

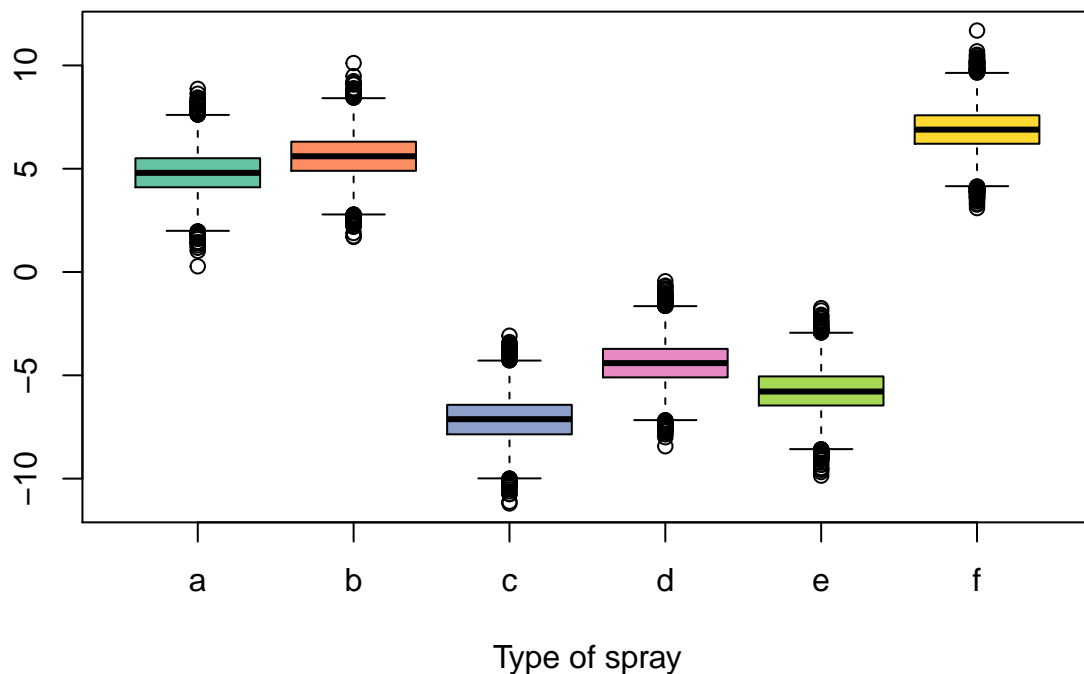
```
## Bayes factor analysis
## -----
## [1] spray : 1.506706e+14 ±0%
##
## Against denominator:
##   Intercept only
## ---
## Bayes factor type: BFlinearModel, JZS
```

```
## Run mcmc iterations
mcmcOut <- posterior(bayesOut, iterations=10000)
```

Yields a Bayes Factor of 1.506706e:1 in favor of an effect of spray, which is really strong. Technically, odds ratio represents comparison between alternative hypothesis and null. Using the rules of thumb offered by Kass and Raftery (1995), any odd ratio in excess of 150:1 is considered very strong evidence. This result confirms the previous evidence of suggesting support for an alternative hypothesis of credible difference among these group means. According to the null hypothesis the mean outcome is the same across all groups;

all 6 groups were samples from the same population such that any variation among means was attribute to sampling error.

```
## Box plots are informal methods of comparing HDIs
## Note that these values are relative to the grand mean
boxplot(as.matrix(mcmcOut[,2:7]) # Boxplot the posteriors for the grs
, xaxt = "n"
, col = brewer.pal(6, "Set2")
, main=NULL
, xlab = "Type of spray")
axis(1, at=1:6, labels=letters[1:6])
```



The boxplots above shows posterior distributions for deviations of each group from the grand mean (which is about 9.48, ranges from 8.6 to 10.4). The grand mean is represented on this figure as 0 on the y-axis. Where there is no overlap, group means are substantially different. The boxplots are giving us a direct way of comparing the groups. The F group looks like the best one - has the highest mean of 6.7 and is not overlapping with some of the boxplots at the bottom. Another good one is B (mean of B is slightly lower than the mean of F). F and B are overlapping with each other, which means that they are similar, there is no credible difference between them. We can see that despite that they are definitely better than C for example. Spray A is overlapping with B, but it is also one the positive site and doesn't overlap with the mean.

The HDI table below give us more accurate numeric results.

```
## Review the results from MCMC
summary(mcmcOut) # Show the HDIs
```

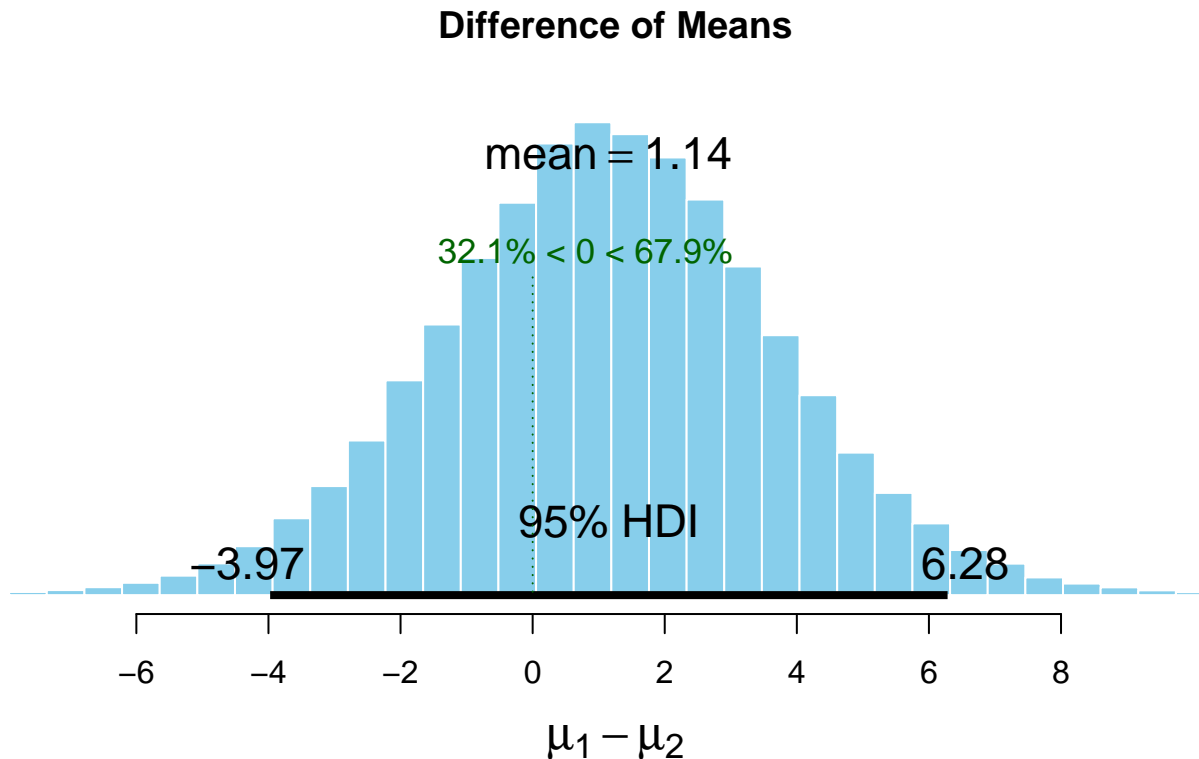
```
##
## Iterations = 1:10000
## Thinning interval = 1
## Number of chains = 1
## Sample size per chain = 10000
##
## 1. Empirical mean and standard deviation for each variable,
##    plus standard error of the mean:
##
##           Mean      SD Naive SE Time-series SE
## mu          9.494 0.4737 0.004737      0.004737
## spray-A     4.812 1.0488 0.010488      0.010743
## spray-B     5.602 1.0584 0.010584      0.010584
## spray-C    -7.135 1.0599 0.010599      0.010907
## spray-D    -4.412 1.0353 0.010353      0.010353
## spray-E    -5.766 1.0528 0.010528      0.010747
## spray-F     6.899 1.0441 0.010441      0.010402
## sig2       16.124 2.8851 0.028851      0.033547
## g_spray     3.430 3.4915 0.034915      0.037066
##
## 2. Quantiles for each variable:
##
##           2.5%    25%    50%    75%   97.5%
## mu          8.5607  9.182  9.489  9.812 10.439
## spray-A     2.7741  4.102  4.801  5.509  6.873
## spray-B     3.5066  4.899  5.602  6.309  7.697
## spray-C    -9.1770 -7.856 -7.124 -6.428 -5.060
## spray-D    -6.4704 -5.100 -4.410 -3.720 -2.349
## spray-E    -7.8440 -6.461 -5.785 -5.053 -3.670
## spray-F     4.8417  6.210  6.894  7.587  8.992
## sig2       11.4054 14.099 15.773 17.790 22.773
## g_spray     0.8307  1.658  2.512  3.998 11.414
```

The detailed numeric output from `anovaBF()` provides the specific boundaries of the 95% highest density intervals (HDI) around each group deviation. All the spray types are listed on the left, starting with the grand mean (the HDI for `mu` ranges from 8.6 to 10.4). So we can assert with some confidence that the population mean value of the spray groups falls somewhere in that range of 8.6 to 10.4. We can also look at the mean of each group - start from 6.89 for spray F as the highest and ends with the mean of spray C (-7.12). These are deviations away from the mean - some are higher, and some are lower. We're mainly concerned with comparisons of each group of spray with the others and if their HDI's overlap. C, D and E overlap with each other so there is no credible difference between them. The same with spray F which is the best but overlap with A and B so there is no credible difference between them too. Winner are F, C(top 2), and losers are C and E sprays.

Conclusion

An experiment was conducted with six different kinds of insecticide. For each kind of insecticide, $n = 12$ observations were conducted. Results of a conventional ANOVA showed a significant difference among spray counts measured $a(F(5,66)=34.7, p<.001)$. A Bayesian test confirmed this result with a Bayes factor of $2e-16$ ***:1 in favor of a mean differences model as compared to an intercept-only model. A Bayesian analysis of the group means showed that both spray C and spray B provided superior counts with an estimated 6.9 and 5.6 respective increases over the overall mean of 9.5. Highest density intervals of the posterior distributions for these means overlapped with spray A, but had no overlap (and therefore were superior to) spray C, D and E.

```
## Comparing only 2 groups to find difference
sprayBest <-BESTmcmc(InsectSprays[spray == "F",1],
  InsectSprays[spray == "B",1])
plot(sprayBest)
```



Overlap with 0, no credible difference. The two sprays look very similar.

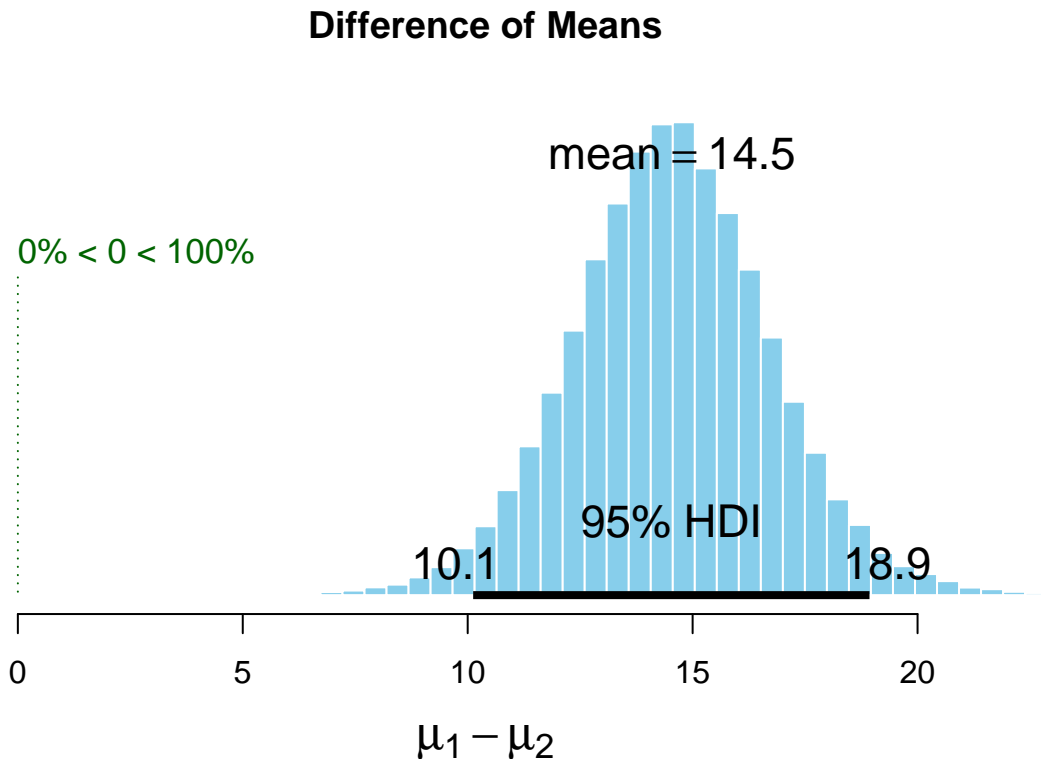
7. In situations where the alternative hypothesis for an ANOVA is supported and there are more than two groups, it is possible to do post-hoc testing to uncover which pairs of groups are substantially different from one another. Using the `InsectSprays` data, conduct a t-test to compare groups C and F (preferably a Bayesian t-test). Interpret the results of this t-test.

```
t.test(InsectSprays[spray == "F",1],
  InsectSprays[spray == "C",1])
```

```
##
## Welch Two Sample t-test
##
## data: InsectSprays[spray == "F", 1] and InsectSprays[spray == "C", 1]
## t = 7.7484, df = 13.201, p-value = 2.876e-06
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## 10.52358 18.64308
## sample estimates:
## mean of x mean of y
## 16.66667 2.083333
```



```
## Comparing only 2 groups to find difference
sprayBEST2 <-BESTmcmc(InsectSprays[InsectSprays$spray == "F",1],
  InsectSprays[InsectSprays$spray == "C",1])
plot(sprayBEST2)
```



The Bayesian t-test examined whether there was a mean difference in count of insects killed by the spray-F and spray-C. Mean insects killed for spray-F was 17 and for spray-C only 2. The Bayesian t-test produced a distribution of estimates for the mean difference. The center of this distribution was a difference of 15 insects. The population mean difference probably lies somewhere close to this value. The 95% highest density interval for this distribution of estimates ranged from 10 to 19. Among all the estimates generated, none of the suggested a mean difference of zero, thereby providing a strong evidence for a credible difference in means. The HDI doesn't overlap with 0 and is 100% positive. Spray-C is definitely the better spray.

Another way to compare them is Tukey's honest significance test, or Tukey's HSD (honestly significant difference) test. The test is a single-step multiple comparison procedure and statistical test. It can be used to find means that are significantly different from each other.

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

```
summary(aovOut)
```

```
##           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
```

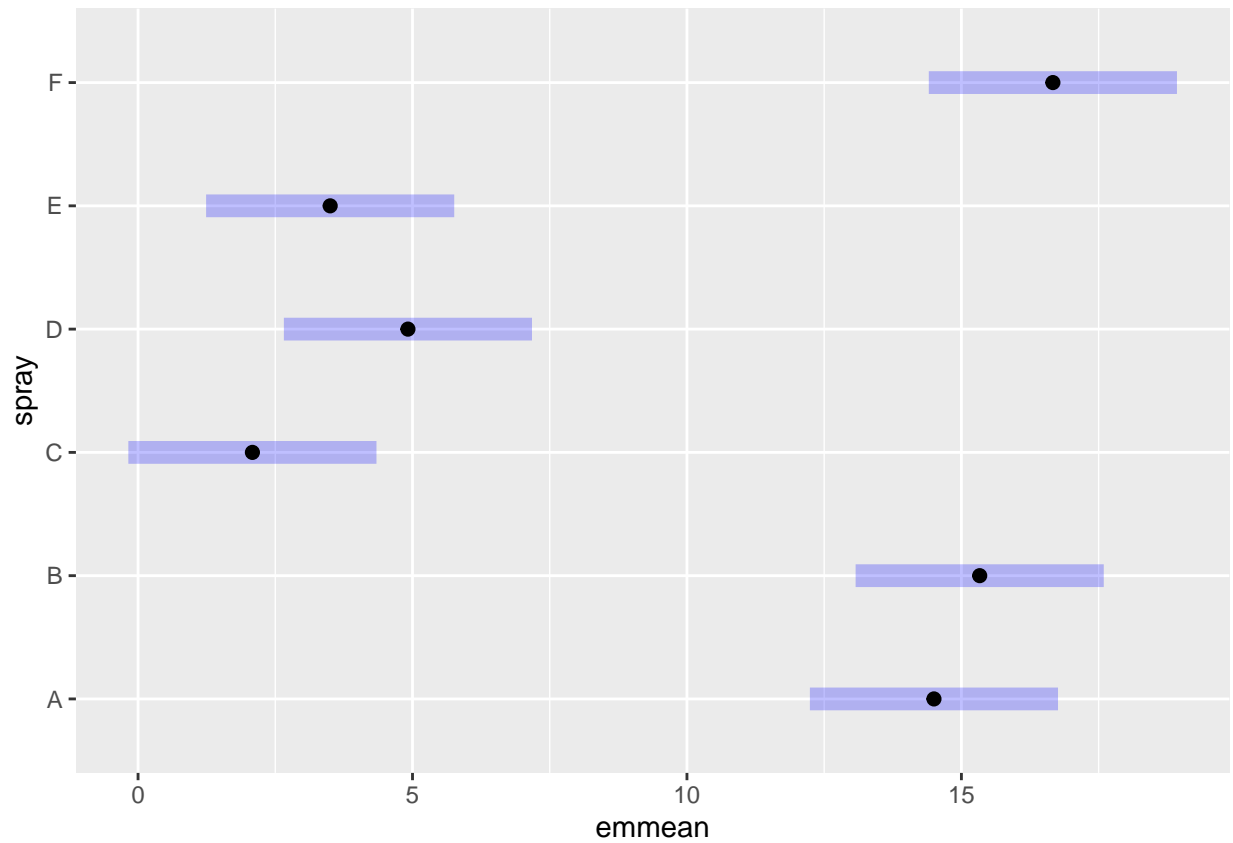
```
TukeyHSD(aovOut) # Create a Tukey table
```

```
## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = count ~ spray, data = InsectSprays)
##
## $spray
##      diff      lwr      upr    p adj
## B-A  0.8333333 -3.866075  5.532742 0.9951810
## C-A -12.4166667 -17.116075 -7.717258 0.0000000
## D-A  -9.5833333 -14.282742 -4.883925 0.0000014
## E-A -11.0000000 -15.699409 -6.300591 0.0000000
## F-A   2.1666667  -2.532742  6.866075 0.7542147
## C-B -13.2500000 -17.949409 -8.550591 0.0000000
## D-B -10.4166667 -15.116075 -5.717258 0.0000002
## E-B -11.8333333 -16.532742 -7.133925 0.0000000
## F-B   1.3333333  -3.366075  6.032742 0.9603075
## D-C   2.8333333  -1.866075  7.532742 0.4920707
## E-C   1.4166667  -3.282742  6.116075 0.9488669
## F-C  14.5833333   9.883925 19.282742 0.0000000
## E-D  -1.4166667  -6.116075  3.282742 0.9488669
## F-D  11.7500000   7.050591 16.449409 0.0000000
## F-E  13.1666667   8.467258 17.866075 0.0000000
```

Significant comparisons: C-A; D-A; E-A, C-B; D-B; **F-C**; F-D; F-E

That confirms the significant difference between spray-F and spray-C.

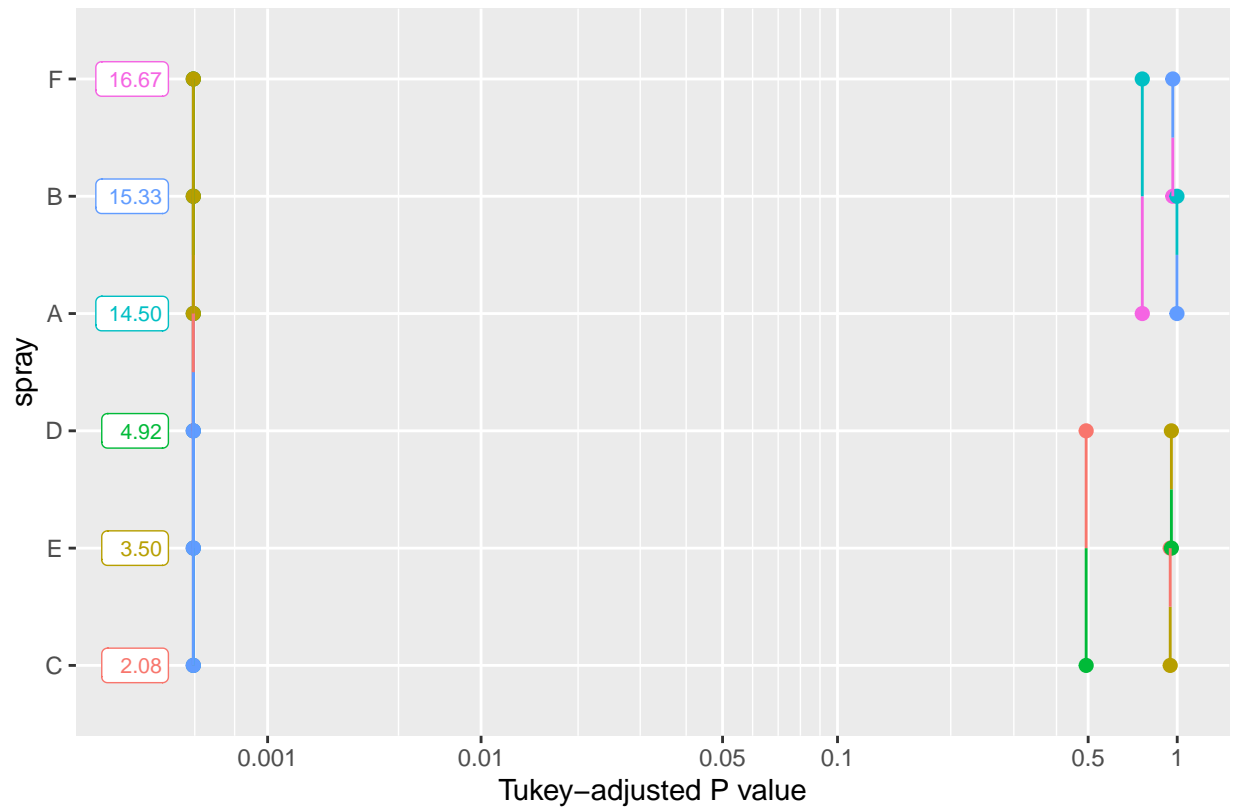
```
emOut <- emmeans(aovOut, "spray") # Estimate marginal means
plot(emOut)
```



```
summary(emOut)
```

```
##  spray emmean   SE df lower.CL upper.CL
##  A      14.50 1.13 66   12.240   16.76
##  B      15.33 1.13 66   13.073   17.59
##  C       2.08 1.13 66    -0.177    4.34
##  D       4.92 1.13 66     2.656    7.18
##  E       3.50 1.13 66     1.240    5.76
##  F      16.67 1.13 66    14.406   18.93
##
## Confidence level used: 0.95
```

```
pwpp(emOut) # Show a Tukey plot based on em means results
```



We can clearly see where they overlap.