

Homework #11

Justin Robinette

November 13, 2018

No collaborators for any problem

Problem #1: Consider the **alpha** dataset from the **coin** package. Compare the results when using **glht** and TukeyHSD (Refer to Chapter 5 review for TukeyHSD).

Results: Using both **glht** and **TukeyHSD**, we can compare the confidence intervals and significance of difference in allele length from the *alpha* dataset.

First, similarly to the lecture, we inspect the data to see the differences in the Expression Level by the allele length. This is shown by **Figure 1.1** below.

Next I used the Generalized Linear Hypothesis functionality from the **multcomp** library to get the confidence intervals and p-values of difference in allele length. **Figure 1.2** shows the confidence intervals by allele length. From this plot, we see very minimal difference between long and intermediate and short and intermediate alleles. We do see a marginal difference between long and short alleles from the plot. This difference can be further examined by **Figure 1.3** which shows that the p-value for the difference between long and short is statistically significant at a higher alpha of 0.1, but not at the alpha of 0.05.

For comparison, I used **TukeyHSD** to examine the effect that allele length has on the expression level. **Figure 1.4** shows a similar confidence interval plot to **Figure 1.2** where we see very little difference between the intermediate group and the other two groups when using TukeyHSD. **Figure 1.5** shows us that the p-value of the difference between the long allele group and the short allele group is again not statistically significant at an alpha of 0.05, but it is significant at an alpha of 0.1.

I see no real difference in the results of the two methods. Both show a marginally significant effect between long and short alleles on the expression level.

Only base R plots as plots were not requested by the exercise.

Figure 1.1: Expression Level by Allele Length

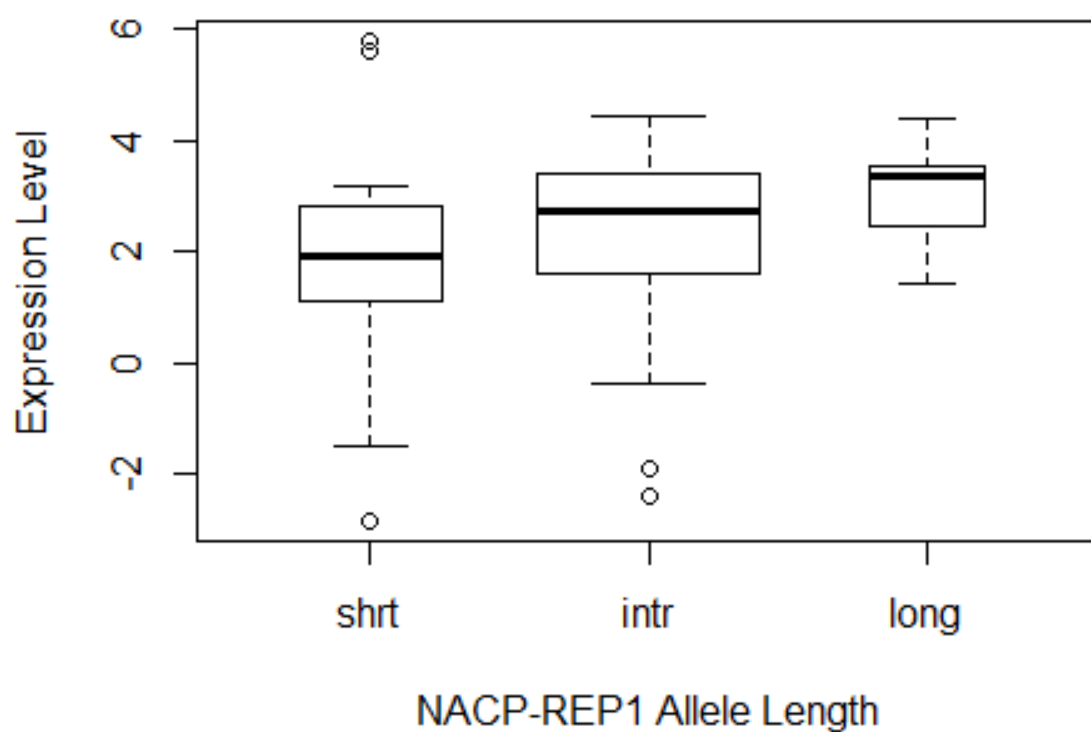


Figure 1.2: Generalized Linear Hypothesis CIs

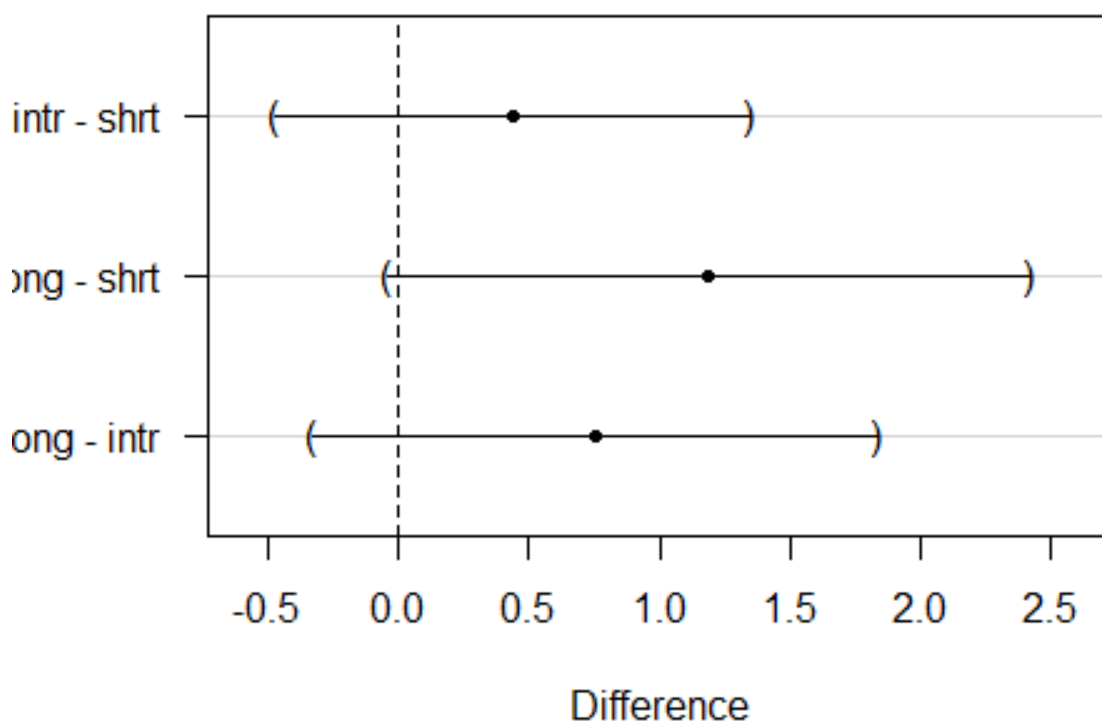


Figure 1.3

```
## Simultaneous Tests for General Linear Hypotheses
##
## Multiple Comparisons of Means: Tukey Contrasts
##
## Fit: aov(formula = elevel ~ alength, data = alpha)
##
## Linear Hypotheses:
##           Estimate Std. Error t value Pr(>|t|)
## intr - shrt == 0    0.4342     0.3836   1.132   0.4924
## long - shrt == 0    1.1888     0.5203   2.285   0.0614 .
## long - intr == 0    0.7546     0.4579   1.648   0.2271
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## (Adjusted p values reported -- single-step method)
```

Figure 1.4

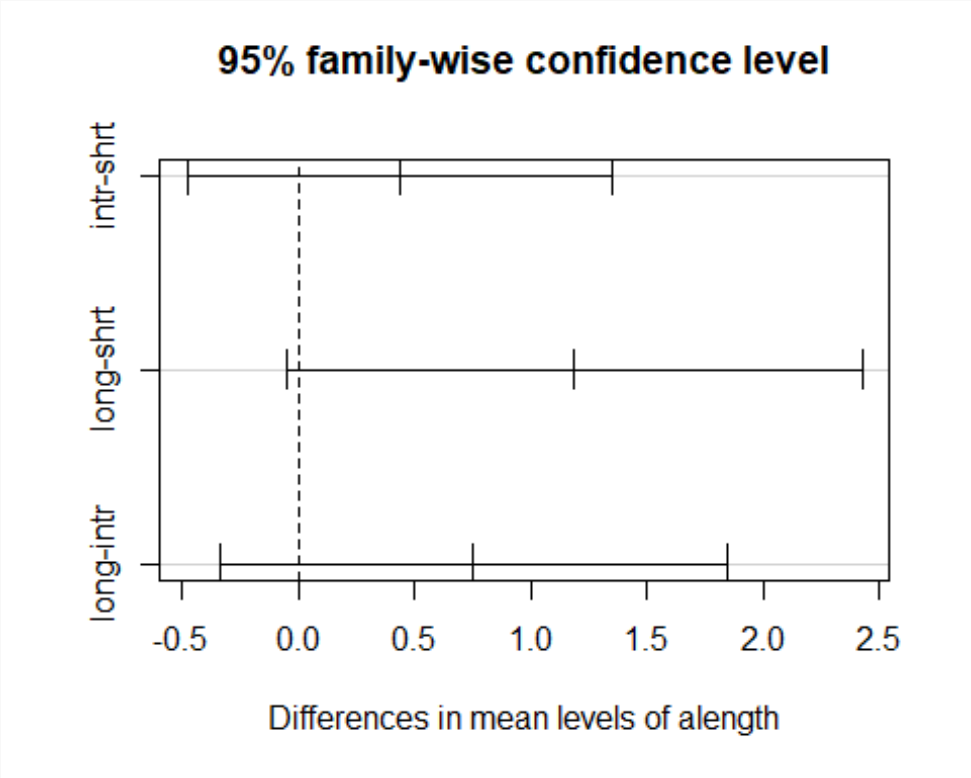


Figure 1.5

```
## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = elevel ~ alength, data = alpha)
##
## $alength
##           diff           lwr           upr           p adj
## intr-shrt 0.4341523 -0.47943766 1.347742 0.4970962
## long-shrt 1.1887500 -0.05017513 2.427675 0.0628589
## long-intr 0.7545977 -0.33575201 1.844947 0.2307995
```

Problem #2, Part A: Consider the **clouds** data from the **HSAUR3** package.

Read and write a report (no longer than one page) on the clouds data given in Chapter 15 section 15.3.3 from Handbook Ed 3.

Results: The clouds data looks at weather modification (cloud seeding), in conjunction with other contributory factors, and the corresponding effect on rainfall. This dataset, from 1975, contains a small number of observations - only 24 days. Because of this rather small sample size, we can examine the confidence band around an estimated regression line, with probability greater than or equal to $1 - \alpha$.

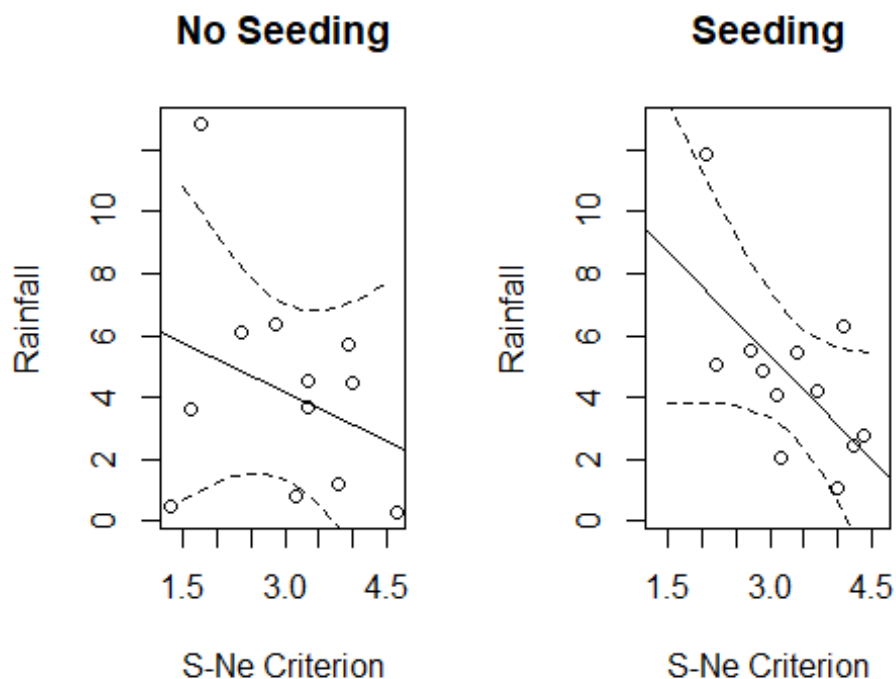
Although it is easy to calculate the pointwise confidence intervals, we want to make sure that we control the type I error, or the rejection of a true Null Hypothesis. To do so, we can multiply the $\beta_1 sne_1$ values by a matrix **K**.

We can write a function to get confidence intervals for all the parameters of interest to form a confidence band of the estimated regression line:

```
confband <- function(subset, main) {  
  mod <- lm(rainfall ~ sne, data = clouds, subset = subset)  
  sne_grid <- seq(from = 1.5, to = 4.5, by = 0.25)  
  K <- cbind(1, sne_grid)  
  sne_ci <- confint(glht(mod, linfct = K))  
  plot(rainfall ~ sne, data = clouds, subset = subset,  
       xlab = "S-Ne Criterion", main = main, ylab = "Rainfall",  
       xlim = range(clouds$sne),  
       ylim = range(clouds$rainfall))  
  abline(mod)  
  lines(sne_grid, sne_ci$confint[,2], lty = 2)  
  lines(sne_grid, sne_ci$confint[,3], lty = 2)  
}
```

The above function, **confband** fits a linear model to a subset of the clouds dataset. Next, we can use this function to produce plots showing subsets of the data based on whether or not seeding action occurred.

As we can see, there is larger variability observed when seeding did not occur than there is when seeding occurred. *Only base R plots as plots were not requested by the exercise.*



Problem #2, Part B: Consider the linear model fitted to the clouds data as summarized in Chapter 6, Figure 6.5. Set up a matrix K corresponding to the global null hypothesis that all interaction terms presented in the model are zero. Test both the global hypothesis and all hypotheses corresponding to each of the interaction terms.

Results: First, I used the clouds_formula summarized in Chapter 6 to fit a linear model for the clouds dataset. Next, I set up a matrix K with all interaction terms equal to 0.

I then used the **glht** function with 'linfct = K' to test both the global hypothesis and each hypothesis corresponding to the individual interaction terms.

Figure 2.1 shows the p-value for the global hypothesis test as significant at an alpha of 0.05.

Figure 2.2 shows the p-values for each individual interaction term within the model, again leaving the terms set equal to zero. Here we see that most of the interaction terms are not significant. Seeding = yes interaction with 'sne' is marginally significant at an alpha of 0.1.

Figure 2.1: P-Value of Model with All Interaction Terms = 0

Global Hypothesis Test P-Value
0.0243093

Figure 2.2: P-Value Corresponding to Each Interaction Term = 0

attribute	pvalue
seedingyes == 0	0.0293579
time == 0	0.5009460
seedingno:sne == 0	0.9992449
seedingyes:sne == 0	0.0769518
seedingno:cloudcover == 0	0.5095466
seedingyes:cloudcover == 0	0.9656074
seedingno:prewetness == 0	0.8855053
seedingyes:prewetness == 0	0.9978002
seedingno:echomotionstationary == 0	0.6028928
seedingyes:echomotionstationary == 0	0.7332756

Problem #2, Part C: How does adjustment for multiple testing change which interactions are significant?

Results: **Figure 2.3** shows the p-values for the attributes with the interactions without adjusting for multiple testing. Here we see that two attributes, from the formula provided in Chapter 6 (per the instructions), are significant at an alpha of 0.05 - denoted by the blue circle.

Figure 2.4 shows the p-values for the attributes with the interaction terms after adjusting for multiple testing. Here we see that there is only one significant term from the formula provided in Chapter 6. Additionally we see that every point appears to have a larger p-value than it did in the model provided in Chapter 6.

From these plots, we can deduce that adjusting for multiple testing lessens the significance of the attributes.

No base R plots provided since the question does not ask for plotting.

Figure 2.3: P-Values without Multiple Testing

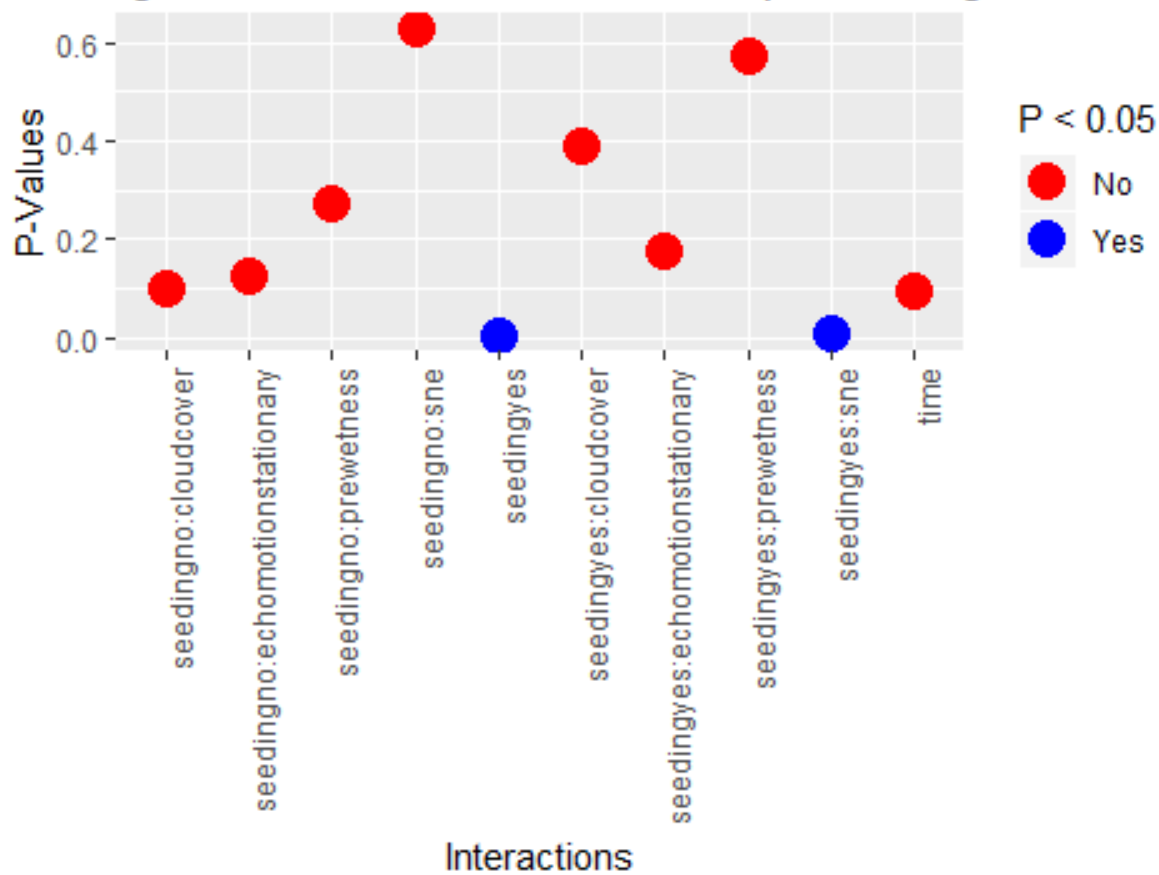
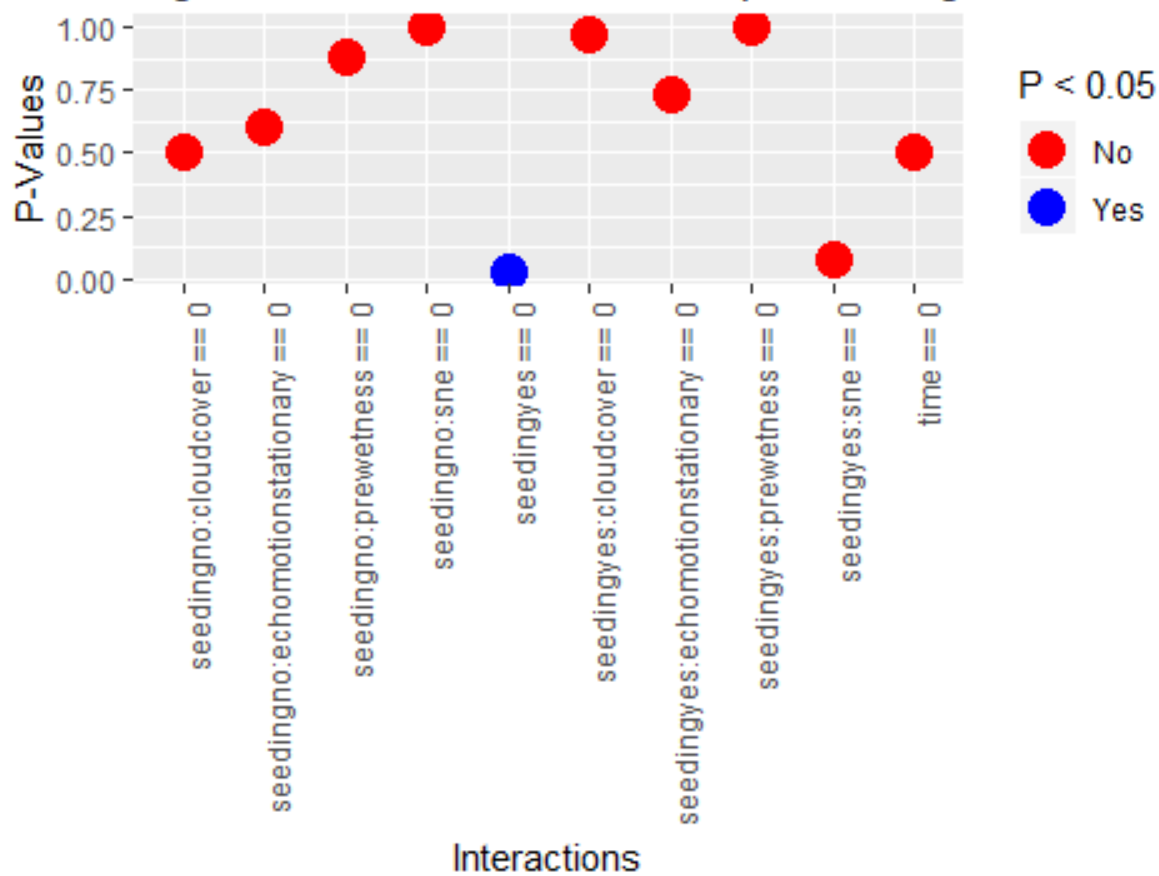


Figure 2.4: P-Values with Multiple Testing



Problem #3: For the logistic regression model presented in Chapter 7 in Figure 7.7, perform a multiplicity adjusted test on all regression coefficients (except for the intercept) being zero. Do the conclusions drawn in Chapter 7 remain valid?

Results: First, I recreated the model from Figure 7.7, including the interaction between gender and education as a predictor of agree / disagree. Next I created a model with each coefficient being zero.

My next step was to compare the p-values. The conclusion from Figure 7.7, per page 131 of the *Handbook* is that “the **gender** and **education** interaction term is seen to be highly significant...”. Examining this conclusion with the **glht** function, we can see from **Figure 3.1** that the p-values remain similar and statistically significant when the regression coefficients (except for the intercept) are zero.

In the glm model from Chapter 7, the interaction term was significant at a level of 0.01. In the glht model, that is used for comparison in this exercise, the p-value is slightly greater than 0.01.

Therefore, I conclude that the original conclusion from Chapter 7 remains valid at an alpha of 0.05 but would not be valid at a smaller alpha of 0.01.

Figure 3.1: Comparison of P-Values by Model

attributes	glm p-values	attributes	ht p-values
genderFemale	0.0119813	genderFemale ==0	0.0245315
education	0.0000000	education ==0	0.0000000
genderFemale:education	0.0088588	genderFemale:education ==0	0.0174179