36-618 HW6

Daniel Nason

5/1/2022

```
resp <- read.table('C:/Users/Owner/CMU/Spring/36-618/HW/HW6/respiratory_exchange_ratio.dat', header = T)
weld <- read.table('C:/Users/Owner/CMU/Spring/36-618/HW/HW6/weld_strength.dat', header = T)
```

Question 1

There are some potential issues with the economist's proposed analysis about the effect of the new stove on air quality and indoor air pollution. While looking at a village in the same area implies similarities between geography and climate, other information in the data description suggests that there are issues with random assignment and confounders that could influence outcomes and taint the results. There are issues with randomization in that the new stoves are not randomly assigned between the villagers. The people with more wealth could be the ones purchasing the stove, and these people could live in homes with better air quality or ventilation systems that impact the health outcome of interest but are not related to the stove.

Another issue would be judging the health outcomes as a measure of whether or not air quality has improved. The demographics differences of people who bought new stoves versus ones who didn't also comes into play, as wealthier people typically have healthier lifestyles, diets, and habits as well as better access to medical care. As a result, differences in health outcomes may be driven by these demographics and not by installation of the new stove. Additionally, people with larger families or different demographics of families (i.e. 3 or more generations in a single house) could use the stove more frequently, which as a greater impact on air quality in the house. Older generations and people with pre-existing conditions may also be more susceptible to worse health outcomes, and poorer households are more likely to have these generations in their homes. One final thing to note is the time frame of the analysis may not be appropriate to determine the efficacy of the stoves, as people may use the stoves to both cook and keep the home more often in colder seasons of the year. This increased usage could have a greater impact on air quality, and the new stove could have diminished benefits from repeated usage during these months. As a result, the shortened time frame could lead to inappropriate conclusions about the benefits of the new stove.

In order to better gauge the effect stove has on air quality and health outcomes, we would first need to randomly assign the new stoves throughout the community. Ideally, the assignment would be amongst people of similar demographics such as wealth, home type, access to health care, family size and composition (grandparents, parents, children, etc.). We would then need to observe the health outcomes and make sure that the family members did not have any pre-existing conditions that could potentially influence measures of health outcome. Finally, we would need to make sure that the data are over the same periods of time for when the analysis is performed. After including random assignment and controlling for potential confounders, we could then make valid inferences about the new stove's effect on indoor air quality.

Question 2

a)

It is important to randomize the order of treatments within each block since one treatment may have an impact on the outcome measure of the next treatment. Without randomizing, we would not get an accurate measure of the effects of the treatments since order could potentially be a confounding variable. Randomization ensures that each person gets an equal likelihood of receiving a given sequence of treatments.

b)

$$\text{Two-Way ANOVA: } ratio_{hi} = \mu + subject_h + protocol_i + \epsilon_{hi}$$

```
anova(lm(ratio ~ as.factor(subject) + as.factor(protocol), data = resp))
```

```
## Analysis of Variance Table
##
## Response: ratio
##                     Df   Sum Sq   Mean Sq F value  Pr(>F)
## as.factor(subject)   8 0.037163 0.0046454  2.7672 0.03958 *
## as.factor(protocol)  2 0.000807 0.0004037  0.2405 0.78904
## Residuals           16 0.026859 0.0016787
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

After blocking on each subject, we see that the F-test result for subject is statistically significant but the treatment effect (protocol) is not statistically significant at the 5% level. The insignificance of the treatment effects suggests that there is no statistically meaningful difference between the different protocols and their effects on respiratory exchange ratio.

c)

Since the mean square of the block effect is greater than the mean square error by a factor of 2.7672366, we can say that block is useful in this experiment.

d)

```
anova(lm(ratio ~ as.factor(protocol), data = resp))
```

```
## Analysis of Variance Table
##
## Response: ratio
##                     Df   Sum Sq   Mean Sq F value Pr(>F)
## as.factor(protocol)  2 0.000807 0.0004037  0.1513 0.8604
## Residuals           24 0.064022 0.0026676
```

Comparing the results from the one-way ANOVA to the two-way ANOVA, we see that the F-statistic for protocol is still not statistically significant at the 5% level in either case and that the p-value of the F-statistic for the one-way ANOVA is relatively larger due to the residuals containing the unaccounted for sum of squares of blocking on each subject.

The one-way ANOVA has issues relative to the two-way ANOVA since we saw that because the F-statistic of blocking on each subject significant it should be included in the model. That is, there is more variation between blocks than within blocks and so accounting for these reduces the mean square error. Not blocking might lead us to incorrect conclude that protocol is not significant since the error sum of squares and therefore mean square error increases, which is the denominator in the F-statistic. The F-statistic would therefore be smaller and less likely to be significant for the given degrees of freedom.

Question 3

a)

```
anova(lm(strength ~ as.factor(trtmt), data = weld))
```

```
## Analysis of Variance Table
##
## Response: strength
##                  Df Sum Sq Mean Sq F value    Pr(>F)
## as.factor(trtmt) 14 1261.2  90.086  8.2395 0.0001092 ***
## Residuals        15  164.0  10.933
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Since the p-value is statistically significant at the 5% level, this suggests that we can reject the null hypothesis in favor of the alternative that at least two of the treatment combinations have different effects but not which combination of treatments is different. Therefore, it would make sense to do a two-way ANOVA to account the possible treatment combinations and see their effect on the response variable.

b)

```
anova(lm(strength ~ as.factor(gage) * as.factor(time), data = weld))
```

```
## Analysis of Variance Table
##
## Response: strength
##                                 Df Sum Sq Mean Sq F value    Pr(>F)
## as.factor(gage)                  2 278.60 139.300 12.7409 0.0005838 ***
## as.factor(time)                  4 385.53  96.383  8.8155 0.0007212 ***
## as.factor(gage):as.factor(time)  8 597.07  74.633  6.8262 0.0007535 ***
## Residuals                       15 164.00  10.933
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

From the results of the two-way ANOVA with interaction, we see that both the main effects terms for gage and time as well as their interaction are statistically significant at the 5% level. This suggests that we can reject the null hypothesis that there is no interaction between the factors and therefore it is appropriate to include the interaction term in the model.

c)

```
anova(lm(strength ~ as.factor(gage) + as.factor(time), data = weld))
```

```
## Analysis of Variance Table
##
## Response: strength
```

```
##                  Df Sum Sq Mean Sq F value  Pr(>F)
## as.factor(gage)  2 278.60 139.300  4.2097 0.02767 *
## as.factor(time)  4 385.53  96.383  2.9128 0.04369 *
## Residuals       23 761.07  33.090
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Given that the interaction term is statistically significant in part (b), we cannot trust the p-values from the model without the interaction term. This is due to the interaction term accounting for more of the variation and reducing the error sum of squares. A large mean square error results in a larger denominator in the F-statistic for both time and gage, which inflates the p-value by decreasing the F-statistic and makes it more difficult to detect whether the main effects are statistically significant.