ST 518 Homework 6

Eric Warren

October 22, 2023

Contents

1 Problem 1

A completely randomized experiment investigates the effects of increasing nitrogen (N) and copper (Cu) in the diet of chickens. Feed conversion ratio (FCR) is observed on $n = 4$ chickens for each of four treatment combinations(diets), with output below. Data are available online as "fcr.dat" so you can check your answers, but you should be able to complete these problems without software.

1.1 Part A

Write a factorial effects model for the 16 observed FCR measurements which assumes that, for a given diet, FCR is normally distributed, with variance σ^2 that is constant across diets.

We can say that this factorial effects model is $Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + E_{ijk}$ where for $i = 1, 2, 3, 4$ and $j = 1, 2, 3, 4$ and E_{ijk} are i.i.d $N(0, \sigma^2)$ errors.

1.2 Part B

Estimate the simple effect of increasing copper when $N = 25$.

We can estimate the simple effect by only looking at the output for when $N = 25$. This occurs when $Cu = 10$ and gives us a mean value of 133 and when $Cu = 100$ and gives us a mean value of 146. Therefore, the simple effect of increasing copper when $n = 25$ is 146 - 133 = 13. For future purposes we will call this answer $\hat{\theta_1} = \bar{y_3} - \bar{y_1} = 13$ where $\bar{y_3} = 146$ and $\bar{y_1} = 133$.

1.3 Part C

Estimate the simple effect of increasing copper when $N = 45$.

We can estimate the simple effect by only looking at the output for when $N = 45$. This occurs when $Cu = 10$ and gives us a mean value of 130 and when $Cu = 100$ and gives us a mean value of 127. Therefore, the simple effect of increasing copper when $n = 45$ is subtracting these two values of 127 - 130 = -3. For future purposes we will call this answer $\hat{\theta_2} = \bar{y_4} - \bar{y_2} = -3$ where $\bar{y_4} = 127$ and $\bar{y_2} = 130$.

1.4 Part D

Estimate the difference in the simple effects of increasing copper across levels of Nitrogen.

This is just finding the difference between **Part B** and **Part C**. For purposes of this problem, we will call this estimate $\hat{theta}_3 = \hat{\theta_1} - \hat{\theta_2} = \bar{y_3} - \bar{y_1} - (\bar{y_4} - \bar{y_2}) = \bar{y_3} - \bar{y_1} - \bar{y_4} + \bar{y_2} = -13 - (-3) = 16$. Therefore, this difference shown by $\hat{\theta_3} = 16$.

1.5 Part E

Using significance level $\alpha = .05$, test the hypothesis that the simple effects of copper are constant across levels of nitrogen.

Here we are going to do a F-test to see if this simple effects of copper are consistent across the different levels of nitrogen. Our null hypothesis is that the effects of copper are consistent across all levels of nitrogen. The alternative hypothesis is that they are not. We are going to use our result from **Part D** to help us get our appropriate F-value. In this case, we have shown that $\hat{\theta_3} = 16$. Now let us note the contrast which is $(-1, 1, 1, -1)'$ which will get us our c_i values later. To get our F-value, we must use the formula $F = \frac{SS(\hat{\theta_3})/(a-1)(b-1)}{MS(E)}$ where $a = 2$ because there are only 2 nitrogen levels, $b = 2$ because there are only 2 copper levels, $SS(\hat{\theta_3}) = \frac{(\hat{\theta_3})^2}{\sum_{i=1}^{4} \frac{c_i^2}{n_i}} = \frac{16^2}{\frac{-1^2}{4} + \frac{1^2}{4} + \frac{1^2}{4} + \frac{-1^2}{4}} = \frac{256}{4 * \frac{1}{4}} = \frac{256}{1} = 256$, $MS(E) = \frac{1}{4} * \sum_{i=1}^{4} s_i^2 = 36 + 16 + 58.6666667 + 57.3333333 = \frac{1}{4} * 168 = 42$. Now that we have found all the values for our F-test we can get our F-value to be $F = \frac{SS(\hat{\theta_3})/(a-1)(b-1)}{MS(E)} = \frac{256/(2-1)(2-1)}{42} = \frac{256}{42} = 6.095238$ on degrees of freedom $df = (a-1) * (b-1), n(ab-1) = (2-1)(2-1), 4(2*2-1) = 1*1, 4*3 = 1, 12$. We can get our F-critical value (or F^*) by using the `qf()` function in R. Our F^* value is `qf(1-alpha, df1, df2)` = `qf(1-.05, 1, 12)` = 4.7472253. Since the F-value we obtained is greater the F-critical value (6.095238 > 4.7472253) then we reject our null hypothesis and say that the simple effects of copper are **not** consistent across the different levels of nitrogen.

1.6 Part F

Report the smallest level of significance at which the difference between simple copper effects across levels of nitrogen may be declared significant.

I am assuming for this problem that we are essentially finding the p-value we obtain from our F-value and use that as the smallest significance level we can use for the difference between copper effects still be declared significant. Remember for hypothesis testing, we declare significance when the p-value we obtain is less than our significance level (α). Note in **Part E** we showed that the F-value we obtained was 6.095238. Now we can use the `pf()` function to get the smallest level of significance at which we can still declare significance. To do this, we will put into the function `pf(F-value, df1, df2, lower.tail = F)` = `pf(6.095238, 1, 12, lower.tail = F)` = 0.0295555. Therefore, the smallest level of significance at which the difference between simple copper effects across levels of nitrogen may be declared significant is 0.0295555.

1.7 Part G

Report a contrast sum of squares associated with the contrast tested in part (d).

Remember in **Part E**, we said this contrast was $(-1, 1, 1, -1)'$. We can use this for the c_i values in our sum of squares calculation. We are going to call this sum of squares calculation $SS(\hat{\theta_3})$. Therefore, $SS(\hat{\theta_3}) = \frac{(\hat{\theta_3})^2}{\sum_{i=1}^{4} \frac{c_i^2}{n_i}} = \frac{16^2}{\frac{-1^2}{4}+\frac{1^2}{4}+\frac{1^2}{4}+\frac{-1^2}{4}} = \frac{256}{4*\frac{1}{4}} = \frac{256}{1} = 256$. This was also calculated in **Part E** for our F-test, but this is now reinforcing what we did to get this value.

1.8 Part H

Estimate the simple effect of increasing N when $Cu = 100$. Report a standard error and a 95% confidence interval for the effect. In light of this interval, can you declare the observed effect "significant" at level of significance $\alpha = .05$?

First let us find the simple effect of increasing N when $Cu = 100$. We are going to call this effect $\hat{\theta_4} = \bar{y_4} - \bar{y_3} = 127 - 146 = -19$.

Next we are going to report a standard error which is $\hat{SE}(\sum c_i \bar{y_{i.}}) = \sqrt{MS(E) * \sum \frac{c_i^2}{n_i}}$. Now, we know this contrast is just $(0, 0, -1, 1)'$ because we said that $\hat{\theta_4} = \bar{y_4} - \bar{y_3}$. We also said in **Part E** that the $MS(E) = \frac{1}{4} * \sum_{i=1}^{4} s_i^2 = 36 + 16 + 58.6666667 + 57.3333333 = \frac{1}{4} * 168 = 42$. So we can say that $\hat{SE}(\sum c_i \bar{y_{i.}}) = \sqrt{MS(E) * \sum \frac{c_i^2}{n_i}} = \sqrt{42 * (\frac{0^2}{4} + \frac{0^2}{4} + \frac{-1^2}{4} + \frac{1^2}{4})} = \sqrt{42 * (\frac{2}{4})} = \sqrt{21} = 4.5825757$. So, $\hat{SE}(\sum c_i \bar{y_{i.}}) = 4.5825757$.

Now we can get the confidence interval for these two effects, which is just saying $\hat{\theta_4} \pm t(\alpha/2, N-t) * \sqrt{MS(E) * \sum \frac{c_i^2}{n_i}}$ Now we know that $\hat{\theta_4} = 19$, $\alpha = 0.05$ since it is a 95% confidence interval, $N - t = n(ab-1) = 4(2*2-1)4*3 = 12$, and $\sqrt{MS(E) * \sum \frac{c_i^2}{n_i}} = \hat{SE}(\sum c_i \bar{y_{i.}}) = \sqrt{21}$. Therefore, $\hat{\theta_4} \pm t(\alpha/2, N-t) * \sqrt{MS(E) * \sum \frac{c_i^2}{n_i}} = -19 \pm t(0.05/2, 12) * \sqrt{21} = -19 \pm t(0.025, 12) * \sqrt{21}$. So our 95% confidence interval is $(-19 - t(0.025, 12) * \sqrt{21}, -19 + t(0.025, 12) * \sqrt{21})$. Now $t(0.025, 12)$ = `qt(0.025, 12, lower.tail = F)` = 2.1788128. Therefore our confidence interval when plugging the t-value of $t(0.025, 12)$ into it which was $(-19 - t(0.025, 12) * \sqrt{21}, -19 + t(0.025, 12) * \sqrt{21})$ is (-28.9845747, -9.0154253).

Lastly, since we can see that 0 is not contained in our 95% confidence interval we have statistically significant evidence to conclude (or declare) that the observed effect is "significant" at level of significance of $\alpha = .05$.

1.9 Part I

Estimate the main effect of increasing Cu. Give the F-ratio for a test of no effect, along with degrees of freedom.

We can estimate the main effects of increasing Cu by finding the mean of the FCR scores when $Cu = 100$ and then subtract the mean scores when $Cu = 10$. Therefore, we can say that this estimation which we will call $\hat{\theta_5} = \frac{1}{2} * (\bar{y_3} + \bar{y_4}) - \frac{1}{2} * (\bar{y_1} + \bar{y_2}) = \frac{1}{2} * (\bar{y_3} + \bar{y_4} - \bar{y_1} - \bar{y_2}) = \frac{1}{2} * (146 + 127 - 133 - 130) = 5$. Therefore the estimate on the main effect of increasing Cu is $\hat{\theta_5} = 5$.

Now we are going to get a F-ratio for an F-test for a test of no effect. We know that we can get the F-value from the equation $F = \frac{SS(\hat{\theta_5})/df_1}{MS(E)}$ where $SS(\hat{\theta_5}) = \frac{\hat{\theta_5}^2}{\frac{1}{n_{group_1}} + \frac{1}{n_{group_2}}} = \frac{5^2}{\frac{1}{4+4} + \frac{1}{4+4}} = \frac{25}{\frac{2}{8}} = 100$, $df_1 = 1$ since we are only 2 different options (and then subtract 1 to get 1 numerator degree of freedom), and from **Part E** we showed the $MS(E) = \frac{1}{4} * \sum_{i=1}^{4} s_i^2 = 36 + 16 + 58.6666667 + 57.3333333 = \frac{1}{4} * 168 = 42$. Therefore our F-value is $F = \frac{SS(\hat{\theta_5})/df_1}{MS(E)} = \frac{100/1}{42} = 2.3809524$ which has 1 numerator degree of freedom and $n(ab - 1) = 4(2 * 2 - 1) = 12$ denominator degrees of freedom. Therefore our F-ratio is $F = 2.3809524$ with 1, 12 degrees of freedom. Note we cannot make a conclusion since we are not given a significance level.

1.10 Part J

Given the analysis you've done so far, is it appropriate to say the the observed effect of copper in this experiment is not significant (using level $\alpha = .05$)? Explain.

From **Part I** we showed the F-ratio we are using to do this test is $F = 2.3809524$ with 1, 12 degrees of freedom. To compare this to our F-critical value, we can get the F-critical value for a significance level of 0.05 by using the `qf()` function and inputting `qt(alpha, df1, df2, lower.tail = F)` which in this case is `qf(0.05, 1, 12, lower.tail = F)` which gives us the corresponding F-critical value of 4.7472253. Since this F-critical value is larger than the F-value we calculated from our test in **Part I** we fail to reject the null hypothesis that there is no effect of increasing copper. Therefore it is appropriate to say the the observed effect of copper in this experiment is not significant.

1.11 Part K

Report the contrast sums of squares for the main effect of copper and the main effect of nitrogen.

From **Part J** we found the contrast sums of squares for the main effect of copper which was done by $SS(\hat{\theta_5}) = \frac{\hat{\theta_5}^2}{\frac{1}{n_{group_1}} + \frac{1}{n_{group_2}}} = \frac{5^2}{\frac{1}{4+4} + \frac{1}{4+4}} = \frac{25}{\frac{2}{8}} = 100$. So the contrast sums of squares for the main effect of copper is $SS(\hat{\theta_5}) = 100$

Now let us find the estimate of the main effect of increasing nitrogen by finding the mean of the FCR scores when $N = 45$ and then subtract the mean scores when $N = 25$. Therefore, we can say that this estimation which we will call $\hat{\theta_6} = \frac{1}{2} * (\bar{y_2} + \bar{y_4}) - \frac{1}{2} * (\bar{y_1} + \bar{y_3}) = \frac{1}{2} * (\bar{y_2} + \bar{y_4} - \bar{y_1} - \bar{y_3}) = \frac{1}{2} * (130 + 127 - 133 - 146) = 5$. Therefore the estimate on the main effect of increasing Cu is $\hat{\theta_5} = -11$. Now we can find the contrast sums of squares for the main effect of nitrogen which can be done by $SS(\hat{\theta_6}) = \frac{\hat{\theta_6}^2}{\frac{1}{n_{group_1}} + \frac{1}{n_{group_2}}} = \frac{-11^2}{\frac{1}{4+4} + \frac{1}{4+4}} = \frac{121}{\frac{2}{8}} = 484$. So the contrast sums of squares for the main effect of nitrogen is $SS(\hat{\theta_6}) = 484$.

1.12 Part L

Obtain an ANOVA table which partitions the variability between the four treatments into meaningful components.

We found the sum of squares for Copper, Nitrogen, and Nitrogen:Copper (interaction). Keep in mind the degrees of freedom are all 1 since there are only 2 different values of Copper and 2 different values of Nitrogen. So $df_{Nitrogen} = 2 - 1 = 1$, $df_{Copper} = 2 - 1 = 1$, $df_{Nitrogen:Cooper} = (2 - 1)(2 - 1) = 1$. Moreover since the degrees of freedom are all 1 for these parts then the sum of squares for these values equals their respective mean squared values too (because of dividing the sum of squares values by 1 degree of freedom). Next, we know that the $MS(E) = 42$ from **Part E**. Going backwards, we have shown in earlier parts that the error degrees of freedom is 12 (from $n(ab-1) = 4(2*2-1) = 12$). So the $SS(Error) = MS(E)*df_{Error} = 42*12 =$

504. Lastly we can get the corrected total values by saying that degrees of freedom is $a*b-1 = 4*4-1 = 15$ and the sum of squares value is adding up the previous 4 sum of squares which is $100+484+256+504 = 1344$.

Next we need to show F-values. The F-value is shown by doing $F = \frac{SS(\hat{\theta}_i)/df}{MS(E)}$. We are going to show the different F-values below. - For Copper: $F = \frac{100/1}{42} = 2.3809524$ on degrees of freedom 1, 12 - For Nitrogen: $F = \frac{484/1}{42} = 11.5238095$ on degrees of freedom 1, 12 - For Copper and Nitrogen Interaction: $F = \frac{256/1}{42} = 6.0952381$ on degrees of freedom 1, 12

Lastly, we need to get the P-values for the respective things we are modeling. Thus, we are going to use the `pf()` function to do this by inputting `pf(F-value, df1, df2, lower.tail = F)`. We will show this below. - For Copper: p-value is `pf(100/42, 1, 12, lower.tail = F)` = 0.1487723 - For Nitrogen: p-value is `pf(484/42, 1, 12, lower.tail = F)` = 0.0053218 - For Copper and Nitrogen Interaction: p-value is `pf(256/42, 1, 12, lower.tail = F)` = 0.0295555

Now we can make our ANOVA table.

Source	DF	Sum of Squares	Mean Square	F-value	P-value
Copper	1	100	100	2.3809524	0.1487723
Nitrogen	1	484	484	11.5238095	0.0053218
Copper and Nitrogen Interaction	1	256	256	6.0952381	0.0295555
Error	12	504	42		
Corrected Total	15	1344			

Similarly we could make an ANOVA table combining all treatments (and interaction in on thing called "Model"). Here we would add the sum of squares for copper, nitrogen, and its interaction which would be $100 + 484 + 256 = 840$ with degrees of freedom $1 + 1 + 1 = 3$. The mean square value would be the sum of squares divided by the degrees of freedom which is $840/3 = 280$. The F-value would be this mean square value divided by the MS(E) which is $280/42 = 6.6666667$ on 3, 12 degrees of freedom with the p-value being `pf(280/42, 3, 12, lower.tail = F)` = 0.0067142.

Now we can make this version of the ANOVA table.

Source	DF	Sum of Squares	Mean Square	F-value	P-value
Model	3	840	280	6.6666667	0.0067142
Error	12	504	42		
Corrected Total	15	1344			

And now we have made appropriate ANOVA tables showing the effects of copper and nitrogen (along with its interaction effect).

1.13 Part M

Briefly characterize the observed effects of copper and nitrogen on FCR, reporting appropriate p-values along the way.