

# Advanced Data Analysis HW2

Ao Liu, al3472

1.

Suppose you have three different feeds that may affect the size of eggs that chickens lay. You randomly assign 10 chickens to each one of the three feeds and record the size of the eggs (maximum length, in centimeters) that the chickens lay the following week. The null hypothesis is that all the chicken feeds have the same effect on the length of the major axis. The alternative is that the feed has some causal effect. A partial output is

Source	df	SS	MS	F
feed		23.43		
error				
total		28.10		

(a) Complete the table above

Answer:

Source	df	SS	MS	F
feed	2	23.43	11.715	67.731
error	27	4.67	0.1730	
total	29	28.10		

(b)

Test the null hypothesis that all the chicken feeds have the same effect on the length of the major axis against the alternative that the feed has some causal effect. Use  $\alpha = 0.05$ .

Answer:

```
1 >qf(.95, df1=2, df2=27)
2 [1] 3.354131
```

Since  $F = 67.731 > f(0.95, 2, 29) = 3.354131$ , we reject the null hypothesis that all the chicken feeds have the same effect.

2.

Suppose you want to compare the types of popcorn popper and the brand of popcorn with respect to their yield (in terms of cups of popped corn). Factor A is the type of popper: oil-based versus air-based. Factor B is the brand of popcorn: gourmet versus national brand versus generic. For each combination of popper type and brand, you took three separate measurements. The ANOVA table is

Source	df	SS	MS	F
Propper(A)		4.5		
Corn(B)		15.75		
Interaction (A*B)				
error		1.67		
total		22.00		

(a)

Complete the table above.

**Answer:**

Source	df	SS	MS	F
Propper(A)	1	4.5	4.5	32.374
Corn(B)	2	15.75	7.875	56.655
Interaction (A*B)	2	0.08	0.04	0.2878
error	12	1.67	0.139	
total	17	22.00	1.294	

(b)

**Test  $H_0$  : No interaction against  $H_1$  : there is an interaction, use  $\alpha = 0.05$ .**

**Answer:**

```
1 >qf(.95, df1=2, df2=12)
2 [1] 3.885294
```

Since  $F = 32.374 > F(0.95, 1, 12) = 3.885294$ , we reject the null hypothesis that there is no interaction.

(c)

**It is decided to fit a model without an interaction and the partial results are**

Source	df	SS	MS	F
Propper(A)		4.5		
Corn(B)		15.75		
error		1.67		
total		22.00		

**Answer:**

(d)

**Complete the table above.**

**Answer:**

Source	df	SS	MS	F
Propper(A)	1	4.5	4.5	37.72
Corn(B)	2	15.75	7.875	66.01
error	14	1.67	0.1193	
total	17	22.00		

(e)

**Test  $H_0$ : No popper effect against  $H_1$ : there is a popper effect. Use  $\alpha = 0.05$**

**Answer:**

```
1 >qf(.95, df1=1, df2=14)
2 [1] 4.60011
3
```

Since  $F = 37.72 > F(0.95, 1, 14) = 4.6001$ , we reject the null hypothesis that there is no popper effect

(f)

Test  $H_0$ : No corn effect against  $H_1$  : there is a corn effect. Use  $\alpha = 0.05$ .

Answer:

```
1 >qf(.95, df1=2, df2=14)
2 [1] 3.738892
```

Since  $F = 66.01 > F(0.95, 2, 14) = 3.738892$ , we reject the null hypothesis that there is a corn effect

3.

In this exercise A and B are two fertilizers types, M, N, O and P are four wheat types and  $y_{ijk}$  values are wheat yields in bushels per plot (one third of an acre) corresponding to the different combinations of fertilizer type type and wheat type. Also, assume that this data was obtained by using a completely randomized experimental design. (see HW2data.csv)

(a)

Construct an interaction plot? Does it suggest that there is an interaction between fertilizer type and wheat type?

Answer:

(b)

Test  $H_0$  : No interaction against  $H_1$  : there is an interaction, use  $\alpha = 0.05$ .

Answer:

(c)

Fit a model without an interaction and test  $H_0$ : No fertilizer effect against  $H_1$  : there is a fertilizer effect. Use  $\alpha = 0.05$  if you reject  $H_0$ , use Tukeys method to do pairwise comparisons of the different fertilizer types.

Answer:

(d)

Test  $H_0$  : No wheat effect against  $H_1$  : there is a wheat effect. Use  $\alpha = 0.05$  if you reject  $H_0$ , use Tukeys method to do pairwise comparisons of the different wheat types.

Answer: