

# STAT 30040: Homework 5

Robert Winter

## Table of Contents

<b>1</b>	<b>Exercise 1: Two-Way ANOVA Interaction Effect</b>	<b>1</b>
1.1	Part (a) . . . . .	2
1.2	Part (b) . . . . .	3
1.3	Part (c) . . . . .	4
<b>2</b>	<b>Exercise 2: Two-Way Versus One-Way ANOVA</b>	<b>4</b>
2.1	Part (a) . . . . .	5
2.2	Part (b) . . . . .	6
2.3	Part (c) . . . . .	6
2.4	Part (d) . . . . .	7
2.5	Part (e) . . . . .	8
<b>3</b>	<b>Exercise 3: Two-Way Additive Model Without Replication</b>	<b>9</b>
3.1	Part (a) . . . . .	9
3.2	Part (b) . . . . .	11
3.3	Part (c) . . . . .	12
<b>4</b>	<b>Exercise 4: Sequential ANOVA</b>	<b>12</b>
4.1	Part (a) . . . . .	12
4.2	Part (b) . . . . .	13
4.3	Part (c) . . . . .	14

## 1 Exercise 1: Two-Way ANOVA Interaction Effect

Below is the partial output of an ANOVA table of a two-way ANOVA model with interaction.

## Analysis of Variance

Response: Survival

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Poison	1	23.290			
Method	1	8.251			
Poison:Method	1	0.012			
Residuals	44	196.960			

### 1.1 Part (a)

Complete the partial ANOVA table. (Note: The R command `pf(5.8, df1=3, df2=4)` computes  $\mathbb{P}(F \leq 5.8)$  where  $F \sim F_{3,4}$ .)

Each mean square is computed as the ratio of the corresponding sum of squares and degrees of freedom:

```
MS_poison = 23.290 / 1
MS_method = 8.251 / 1
MS_poisonxmethod = 0.012 / 1
MS_error = 196.960 / 44

c(MS_poison, MS_method, MS_poisonxmethod, MS_error)
```

```
[1] 23.290000  8.251000  0.012000  4.476364
```

Each  $F$  statistic is computed as the ratio of the corresponding mean square to  $MS_{error}$ :

```
F_poison = MS_poison / MS_error
F_method = MS_method / MS_error
F_poisonxmethod = MS_poisonxmethod / MS_error

c(F_poison, F_method, F_poisonxmethod)
```

```
[1] 5.202883834 1.843237206 0.002680747
```

Each  $p$ -value is equal to  $1 - \mathbb{P}(F \leq F_{1,44})$ , where  $F$  is the corresponding  $F$  statistic:

```
p_poison = 1 - pf(F_poison, df1 = 1, df2 = 44)
p_method = 1 - pf(F_method, df1 = 1, df2 = 44)
p_poisonxmethod = 1 - pf(F_poisonxmethod, df1 = 1, df2 = 44)

c(p_poison, p_method, p_poisonxmethod)
```

```
[1] 0.02744631 0.18149455 0.95894158
```

Thus, the completed ANOVA table is:

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Poison	1	23.290	23.29	5.2028838	0.0274463
Method	1	8.251	8.251	1.8432372	0.1814945
Poison:Method	1	0.012	0.012	0.0026807	0.9589416
Residuals	44	196.960	4.4763636		

## 1.2 Part (b)

Derive the ANOVA table for the additive (no interaction) two-way ANOVA model (for the same data). Do the  $p$ -values differ from the corresponding  $p$ -values in the model in part (a)? Explain why.

The degrees of freedom, sums of squares, and hence mean squares for the two main effects remain the same as in part (a). However, the residual source of variation “absorbs” the degrees of freedom and sum of squares that are no longer being separated out into the Poison-Method interaction. The new residual degrees of freedom is  $44 + 1 = 45$ , the new residual sum of squares is  $0.012 + 196.960 = 196.972$ , and the new residual mean square is the ratio of the latter to the former:

```
df_error2 = 44 + 1
SS_error2 = 0.012 + 196.960
MS_error2 = SS_error2 / df_error2

c(df_error2, SS_error2, MS_error2)
```

```
[1] 45.000000 196.972000 4.377156
```

Since  $MS_{error}$  has changed, the  $F$  statistics and  $p$ -values for Poison and Method change as well. Each new  $F$  statistic is computed as the ratio of the main effect’s mean square (which has not changed) to the “new”  $MS_{error}$ . Each new  $p$  value is now equal to  $1 - \mathbb{P}(F \leq F_{1,45})$ , where  $F$  is the main effect’s  $F$  statistic and the second degree of freedom of the  $F$  distribution is now 45, not 44.

```
F_poisson2 = MS_poisson / MS_error2
F_method2 = MS_method / MS_error2

c(F_poisson2, F_method2)
```

```
[1] 5.320807 1.885014
```

```
p_poisson2 = 1 - pf(F_poisson2, df1 = 1, df2 = 45)
p_method2 = 1 - pf(F_method2, df1 = 1, df2 = 45)

c(p_poisson2, p_method2)
```

```
[1] 0.02573133 0.17656877
```

Thus, the completed ANOVA table is:

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Poison	1	23.290	23.29	5.320807	0.0257313
Method	1	8.251	8.251	1.8850141	0.1765688
Residuals	45	196.972	4.3771556		

### 1.3 Part (c)

**Which model is better? Is including the interaction more appropriate?**

In Part (a), we found that the Poison-Method interaction effect on survival was highly non-significant, with a  $p$ -value of approximately 0.959. Since this interaction term does not contribute a significant, incremental explanation for the variation in survival, we prefer the model in Part (b) that includes main effects only. It is worth noting that the  $p$ -values for the main effects hardly change between the two models (approximately 0.027 vs 0.025 for Poison, and approximately 0.181 vs. 0.177 for Method), so our conclusions about the main effects of Poison and Method are robust to the inclusion/exclusion of the interaction term.

## 2 Exercise 2: Two-Way Versus One-Way ANOVA

A baseball league changed the strike zone each year for three years. The data below show the number of home runs (HR) that each of four players hit for the three years.

$Y_{ij}$	Player 1	Player 2	Player 3	Player 4
Zone 1	23	21	31	15
Zone 2	12	8	25	10
Zone 3	18	15	30	15

You may input the data to R and conduct the analysis below.

```
baseball = data.frame(HRs = c(23, 12, 18,
                             21, 8, 15,
                             31, 25, 30,
                             15, 10, 15),
                      player = c(1, 1, 1,
                                 2, 2, 2,
                                 3, 3, 3,
                                 4, 4, 4) %>% as.factor(),
                      zone = c(1, 2, 3,
                               1, 2, 3,
                               1, 2, 3,
                               1, 2, 3) %>% as.factor())
```

## 2.1 Part (a)

Perform a one-way analysis of variance for the Player effect on HR. Is there a significant Player effect by the analysis?

```
aova = aov(HRs ~ player,
           data = baseball)
summary(aova)
```

```

              Df Sum Sq Mean Sq F value Pr(>F)
player         3  436.3   145.42    6.369 0.0163 *
Residuals      8   182.7    22.83
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

We use a one-way ANOVA to test the null hypothesis that the average number of home runs hit was equal across the four players against the alternative hypothesis that the average number of home runs was different for at least two of the four players. We recover a  $p$ -value of  $0.016 < 0.05$ , which indicates that the average number of home runs *does* vary significantly between at least two players. In other words, there *is* a significant Player effect.

## 2.2 Part (b)

Perform a one-way analysis of variance for the Zone effect on HR. Is there a significant Zone effect by the analysis?

```
aovb = aov(HRs ~ zone,
           data = baseball)
summary(aovb)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
zone	2	158.2	79.08	1.545	0.265
Residuals	9	460.8	51.19		

We use a one-way ANOVA to test the null hypothesis that the average number of home runs hit was equal across the three strike zones against the alternative hypothesis that the average number of home runs was different for at least two of the three strike zones. We recover a  $p$ -value of  $0.265 > 0.05$ , meaning that there is not statistical evidence that the average number of home runs varies significantly across strike zones. In other words, there is not evidence of a significant Zone effect.

## 2.3 Part (c)

Perform a two-way analysis of variance for the additive effects of Player and Zone on HR. Is there a significant Player effect or Zone effect by the analysis?

```
aovc = aov(HRs ~ player + zone,
           data = baseball)
summary(aovc)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
player	3	436.3	145.42	35.61	0.000322 ***
zone	2	158.2	79.08	19.37	0.002413 **
Residuals	6	24.5	4.08		

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

We use a two-way ANOVA to test two null hypotheses.

(A) We test the null hypothesis that the average number of home runs hit was equal across the four players against the alternative hypothesis that the average number of home runs was different for at least two of the four players. We recover a  $p$ -value of  $0.0003 < 0.001$ , which indicates that the average number of home runs *does* vary significantly between at

least two players. In other words, there *is* a significant Player effect — as we found in part (a) above.

(B) We test the null hypothesis that the average number of home runs hit was equal across the three strike zones against the alternative hypothesis that the average number of home runs was different for at least two of the three strike zones. We recover a  $p$ -value of  $0.002 < 0.01$ , which indicates that the average number of home runs *does* vary significantly between at least two strike zones. In other words, there *is* a significant Zone effect — unlike our finding in part (b) above.

## 2.4 Part (d)

**Explain the reasons of getting different conclusions from the above analyses, and which conclusion you prefer.**

In each of Parts (a) and (c), we found significant Player effects, so there is no discrepancy here. However, there was a discrepancy in our results of Parts (b) and (c) with respect to Zone effects.

In Part (b), we performed a one-way ANOVA to assess whether variation in HR was explained by variation in Zone, and did not find statistical evidence that this was the case. In this analysis, all the variation in HR that was not explained by variation in Zone was attributed to the residuals, giving us a relatively large value of  $SS_{error}$ , hence  $MS_{error}$ . We computed an  $F$  statistic of  $F = \frac{MS_{zone}}{MS_{error}} \approx \frac{79.08}{51.19} \approx 1.545$ .

In Part (c), we performed a two-way ANOVA to assess whether variation in HR was explained by variation in (A) Player, and (B) Zone, and this time, we *did* find statistical evidence that variation in HR was explained by variation in Zone. By also taking into account how variation in HR was explained by variation in Player, this model had less unexplained/residual variation in HR than the model in part (b). As such, this model had lower values of  $SS_{error}$  and thus  $MS_{error}$ . Dividing the same  $MS_{zone}$  by a smaller  $MS_{error}$  gave us a much larger  $F$  statistic than before:  $F = \frac{MS_{zone}}{MS_{error}} \approx \frac{79.08}{4.08} \approx 19.37$ . This much larger  $F$  statistic is associated with a much smaller  $p$ -value, allowing us to find a significant Zone effect where we couldn't before. (It is worth noting that technically, we are comparing our “old” and “new”  $F$  statistics to two different distributions:  $F_{2,9}$  and  $F_{2,6}$ , respectively. But the shapes of the corresponding densities are similar enough that the much larger  $F$  statistic gives a much smaller  $p$ -value.)

We prefer the model in Part (c) to the model in Part (b). By taking into account both Player and Zone, the model in Part (c) is able to significantly reduce  $SS_{error}$  relative to the model in Part (b), meaning that it describes the data better overall. In particular, accounting for Player effects allows us to parse out Zone effects that were previously masked.

## 2.5 Part (e)

**Can we use a two-way analysis of variance to test the interaction effect of Player and Zone? Explain.**

We cannot use a two-way ANOVA to test the interaction effect of Player and Zone on HRs. Since there are four Players and three Zones in our data,

- the total number of degrees of freedom in the ANOVA is  $(3 \times 4) - 1 = 11$ ,
- the degrees of freedom associated with the Player source of variation is  $4 - 1 = 3$ ,
- the degrees of freedom associated with the Zone source of variation is  $3 - 1 = 2$ , and
- the degrees of freedom associated with the Player-Zone interaction source of variation is  $(4 - 1) \times (3 - 1) = 3 \times 2 = 6$ .

But since  $3 + 2 + 6 = 11$ , our sources of explained variation have “exhausted” the degrees of freedom available to the model; there are no degrees of freedom left over to attribute to the residuals/error. As such, we cannot compute  $MS_{error}$  or the  $F$  statistic necessary to test the Player-Zone interaction effect.

Another way to see the problem is through the number of observations. Since the strike zone was changed every year, there is only one home run tally per player per zone. So, if we were to incorporate a Player-Zone interaction into an ANOVA, every Player-Zone pair would correspond to just one data point. Without replicates for each Player-Zone pair, there is no residual variation left over after accounting for the Player-Zone interaction! Again, it is impossible to compute  $MS_{error}$  or the  $F$  statistic corresponding to the Player-Zone interaction.

Below, we try running ANOVAs including the Player-Zone interaction, both with and without main effects. Observe that in each case, R is not able to do any computations related to the residual variation, hence is unable to calculate  $F$  statistics for each explained source of variation.

```
aove1 = aov(HRs ~ player*zone,
            data = baseball)
summary(aove1)
```

	Df	Sum Sq	Mean Sq
player	3	436.3	145.42
zone	2	158.2	79.08
player:zone	6	24.5	4.08

```
aove2 = aov(HRs ~ player:zone,
            data = baseball)
summary(aove2)
```



```

      Df Sum Sq Mean Sq
player:zone 11  618.9   56.27

```

### 3 Exercise 3: Two-Way Additive Model Without Replication

For the two factor observations  $(y_{ij})$  in the following table,

Factor A \ B	B-1	B-2	B-3	B-5	B-5
A-1	4.4	4.3	4.0	3.6	3.3
A-2	4.1	4.1	4.0	4.0	3.9
A-3	3.2	3.9	3.8	4.1	4.4
A-4	3.9	4.0	3.8	4.1	4.4

consider the model (with  $i$  indexes row factor,  $j$  column factor):

$$\begin{aligned}
 Y_{ij} &= \mu + \alpha_i + \beta_j + \varepsilon_{ij}, \\
 \varepsilon_{ij} &\sim \mathcal{N}(0, \sigma^2), \\
 i &= 1, \dots, 4; j = 1, \dots, 5.
 \end{aligned}$$

```

q3data = matrix(c(4.4, 4.3, 4.0, 3.6, 3.3,
                  4.1, 4.1, 4.0, 4.0, 3.9,
                  3.2, 3.9, 3.8, 4.1, 4.4,
                  3.9, 4.0, 3.8, 4.1, 4.4),
                nrow = 4,
                ncol = 5,
                byrow = TRUE)
rownames(q3data) = c("A-1", "A-2", "A-3", "A-4")
colnames(q3data) = c("B-1", "B-2", "B-3", "B-4", "B-5")

```

#### 3.1 Part (a)

Under the constraints  $\sum_{i=1}^5 \alpha_i = 0$ ,  $\sum_{j=1}^5 \beta_j = 0$  (as in Rice §12.3.1), use simple averaging operations (differences, sums, and averages) to compute parameter estimates  $\hat{\mu}$ ,  $\hat{\alpha}_i$ , and  $\hat{\beta}_j$ .

$\hat{\mu}$  is the grand mean of all observations:  $\hat{\mu} = \bar{Y}_{\bullet\bullet}$ .

```

mu_hat = mean(q3data)
mu_hat

```

```
[1] 3.965
```

Each  $\hat{\alpha}_i$  is the difference between the mean for the  $i$ th group of Factor A and the grand mean:  $\hat{\alpha}_i = \bar{Y}_{i\bullet} - \bar{Y}_{\bullet\bullet}$ .

```
# Compute alpha estimates
a1_hat = mean(q3data["A-1", ]) - mu_hat
a2_hat = mean(q3data["A-2", ]) - mu_hat
a3_hat = mean(q3data["A-3", ]) - mu_hat
a4_hat = mean(q3data["A-4", ]) - mu_hat

# Vector of alphas
alpha = c(a1_hat, a2_hat, a3_hat, a4_hat)
alpha
```

```
[1] -0.045  0.055 -0.085  0.075
```

```
# alphas sum to 0, as desired
sum(alpha)
```

```
[1] 0
```

Each  $\hat{\beta}_j$  is the difference between the mean for the  $j$ th group of Factor B and the grand mean:  $\hat{\beta}_j = \bar{Y}_{\bullet j} - \bar{Y}_{\bullet\bullet}$ .

```
# Compute beta estimates
b1_hat = mean(q3data[, "B-1"]) - mu_hat
b2_hat = mean(q3data[, "B-2"]) - mu_hat
b3_hat = mean(q3data[, "B-3"]) - mu_hat
b4_hat = mean(q3data[, "B-4"]) - mu_hat
b5_hat = mean(q3data[, "B-5"]) - mu_hat

# Vector of betas
beta = c(b1_hat, b2_hat, b3_hat, b4_hat, b5_hat)
beta
```

```
[1] -0.065  0.110 -0.065 -0.015  0.035
```

```
# betas sum to 0, as desired (up to floating point error)
sum(beta)
```

```
[1] 4.440892e-16
```

In summary,

- $\hat{\mu} = 3.965$ ,
- $\hat{\alpha}_1 = -0.045$ ,
- $\hat{\alpha}_2 = 0.055$ ,
- $\hat{\alpha}_3 = -0.085$ ,
- $\hat{\alpha}_4 = 0.075$ ,
- $\hat{\beta}_1 = -0.065$ ,
- $\hat{\beta}_2 = 0.11$ ,
- $\hat{\beta}_3 = -0.065$ ,
- $\hat{\beta}_4 = -0.015$ , and
- $\hat{\beta}_5 = 0.035$ ,

### 3.2 Part (b)

Create the 4-by-5 table of residuals  $y_{ij} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}_j$ .

We present the table of residuals  $\hat{\varepsilon}_{ij} = y_{ij} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}_j$  below.

```
# Function to compute residuals based on the above formula
residualify = function(i, j){
  q3data[i,j] - mu_hat - alpha[i] - beta[j]
}

# Make an empty 4x5 matrix of residuals
resids = matrix(data = NA, nrow = 4, ncol = 5)
rownames(resids) = c("A-1", "A-2", "A-3", "A-4")
colnames(resids) = c("B-1", "B-2", "B-3", "B-4", "B-5")

# Populate matrix of residuals
for(i in 1:4){
  for(j in 1:5){
    resids[i,j] = residualify(i,j)
  }
}

# Display
resids
```

	B-1	B-2	B-3	B-4	B-5
A-1	0.545	0.27	0.145	-0.305	-0.655
A-2	0.145	-0.03	0.045	-0.005	-0.155
A-3	-0.615	-0.09	-0.015	0.235	0.485
A-4	-0.075	-0.15	-0.175	0.075	0.325

### 3.3 Part (c)

How does the size of the residuals in Part (b) compare to the size of the estimated row and column effects  $\hat{\alpha}_i$  and  $\hat{\beta}_j$  in Part (a)?

```
min_effect = min(min(abs(alpha)),
                  min(abs(beta)))

max_effect = max(max(abs(alpha)),
                  max(abs(beta)))
```

The magnitudes of the row and column effects  $\hat{\alpha}_i$  and  $\hat{\beta}_j$  are between 0.015 and 0.11. Thus, the row and column effects  $\hat{\alpha}_i$  and  $\hat{\beta}_j$  are on the same orders of magnitude of the residuals  $\hat{\varepsilon}_{ij}$ . In some cases, the row and column effects are even smaller than the corresponding residuals! For example,  $|\hat{\beta}_3| = 0.065 < \hat{\alpha}_4 = 0.075 < |\varepsilon_{43}| = 0.175$ . This is not very encouraging: there is still just as much (if not more) unexplained variation in  $y_{ij}$  after accounting for Factors A and B as there was variation in  $y_{ij}$  explained by Factors A and B.

## 4 Exercise 4: Sequential ANOVA

Below are two sequential ANOVA tables for the same data, obtained from a two factor treatment (A and B) experiment. The data are not shown.

SOURCE	DF	SS	MS
A	3	3.3255	1.1085
B	3	112.95	37.65
A*B	9	0.48787	0.054207
ERROR	14	0.8223	0.058736

SOURCE	DF	SS	MS
B	3	116.25	38.749
A	3		
A*B			
ERROR			

### 4.1 Part (a)

Complete the second table on the right.

Regardless of the order of the sequential ANOVA, the degrees of freedom for each source of variation are the same. The sequential sum of squares and mean squares of the interaction term  $A \times B$  and the error are the same, as well. The total sum of squares is also unchanged.

We can therefore “back into” the sum of squares for Factor A by subtraction ( $SS_A = SS_{Total} - SS_B - SS_{A \times B} - SS_{error}$ ) and then compute  $MS_A$  as usual:

```
SS_A2 = (3.3255 + 112.95 + 0.48787 + 0.8223) -
        (116.25 + 0.48787 + 0.8223)
MS_A2 = SS_A2 / 3

c(SS_A2, MS_A2)
```

```
[1] 0.0255 0.0085
```

SOURCE	DF	SS	MS
B	3	116.25	38.749
A	3	0.0255	0.0085
A*B	9	0.48787	0.054207
ERROR	14	0.8223	0.058736

## 4.2 Part (b)

What do you conclude about the significance of Treatment A effects, Treatment B effects, and interactions?

First, we complete the left ANOVA table (first analyze Factor A, then Factor B), including  $F$  statistics and  $p$ -values:

```
F_A1 = 1.1085 / 0.058736
F_B1 = 37.65 / 0.058736
F_AB1 = 0.054207 / 0.058736

p_A1 = 1 - pf(F_A1, 3, 14)
p_B1 = 1 - pf(F_B1, 3, 14)
p_AB1 = 1 - pf(F_AB1, 9, 14)

c(p_A1, p_B1, p_AB1)
```

```
[1] 3.438033e-05 3.219647e-15 5.339499e-01
```

Source	DF	SS	MS	F	p-value
A	3	3.3255	1.1085	18.8725824	$3.438 \times 10^{-5}$
B	3	112.95	37.65	641.0038137	$3.22 \times 10^{-15}$

Source	DF	SS	MS	F	p-value
A*B	9	0.48787	0.054207	0.9228923	0.5339
ERROR	14	0.8223	0.058736		

Next, we complete the right ANOVA table (first analyze Factor B, then Factor A), including  $F$  statistics and  $p$ -values:

```

F_B2 = 38.749 / 0.058736
F_A2 = MS_A2 / 0.058736
F_AB2 = 0.054207 / 0.058736

p_B2 = 1 - pf(F_B2, 3, 14)
p_A2 = 1 - pf(F_A2, 3, 14)
p_AB2 = 1 - pf(F_AB2, 9, 14)

c(p_B2, p_A2, p_AB2)

```

```
[1] 2.664535e-15 9.313630e-01 5.339499e-01
```

SOURCE	DF	SS	MS	F	p-value
B	3	116.25	38.749	659.7146554	$2.665 \times 10^{-15}$
A	3	0.0255	0.0085	0.1447153	0.9314
A*B	9	0.48787	0.054207	0.9228923	0.5339
ERROR	14	0.8223	0.058736		

As shown in the first table above, when we analyze Factor A first, then Factor B second, the effects of Factors A and B both appear significant. However, as shown in the second table, when we analyze Factor B first, then Factor A second, only Factor B's effect appears significant. In other words, Factor A does not significantly explain variation in our variable of interest once we have already accounted for variation in Factor B. We can therefore conclude that Treatment B has a significant effect on our variable of interest, but Treatment A does not. Meanwhile, regardless of the order in which we analyze Factors A and B, the interaction effect of  $A * B$  has a  $p$ -value of approximately  $0.534 > 0.05$ , so we can also conclude that the interaction of Factors A and B does not have a significant effect on our variable of interest.

### 4.3 Part (c)

Assume there are no empty cells, that is, there is at least one observation in every combination of the levels of the two treatments. What is your guess on the

**likely pattern (e.g., numbers) of replicates? Could all treatment combinations have the same number of replicates?**

Since Factors A and B are each associated with three degrees of freedom, each of these factors have four levels, making for  $4 \times 4 = 16$  A-B treatment combinations in total. If all 16 of these treatment combinations had the same number of replicates, then regardless of the sequence in which we analyzed Factors A and B, the resulting ANOVA tables (including the sums of squares, mean squares, etc.) would be identical. Since the ANOVA table for analyzing Factor A, then Factor B is different than the ANOVA table for analyzing Factor B, then Factor A, it is not possible that each of the 16 treatment combinations had the same number of replicates. In other words, the data must be unbalanced. Given that the  $p$ -value associated with Factor A differs so significantly between the two ANOVA sequences, it might even be the case that the data are highly unbalanced (i.e., that there are very different numbers of replicates for each A-B treatment combination).