Chapter 3 Homework

Jeremy Morris

March 8, 2006

24. *In a three factor completely randomized design, indicate which terms can be assumed negligible and pooled under the following circumstances.*

    We'll use the following table for reference.

	df	EMS
A_i	(I-1)	$\sigma_e^2 + \sigma_{ABC}^2 + J\sigma_{AC}^2 + K\sigma_{AB} + KJ\sigma_A^2$
B_j	(J-1)	$\sigma_e^2 + \sigma_{ABC}^2 + I\sigma_{BC}^2 + K\sigma_{AB}^2 + IK\sigma_B^2$
AB_{ij}	(I-1)(J-1)	$\sigma_e^2 + \sigma_{ABC}^2 + K\sigma_{AB}^2$
C_k	(K-1)	$\sigma_e^2 + \sigma_{ABC}^2 + I\sigma_{BC}^2 + J\sigma_{AC}^2 + IJ\sigma_C^2$
AC_{ik}	(I-1)(K-1)	$\sigma_e^2 + \sigma_{ABC}^2 + J\sigma_{AC}^2$
BC_{jk}	(J-1)(K-1)	$\sigma_e^2 + \sigma_{ABC}^2 + I\sigma_{BC}^2$
ABC_{ijk}	(I-1)(J-1)(K-1)	$\sigma_e^2 + \sigma_{ABC}^2$
$\varepsilon_{l(ijk)}$	IJK(L-1)	σ_e^2

    (a) *AB, AC and ABC insignificant at the 0.25 level.*

    If we remove each term in the EMS column corrosponding to the interactions that are insignificant we get :

	df	EMS
A_i	(I-1)	$\sigma_e^2 + KJ\sigma_A^2$
B_j	(J-1)	$\sigma_e^2 + I\sigma_{BC}^2 + IK\sigma_B^2$
AB_{ij}	(I-1)(J-1)	σ_e^2
C_k	(K-1)	$\sigma_e^2 + I\sigma_{BC}^2 + IJ\sigma_C^2$
AC_{ik}	(I-1)(K-1)	σ_e^2
BC_{jk}	(J-1)(K-1)	$\sigma_e^2 + I\sigma_{BC}^2$
ABC_{ijk}	(I-1)(J-1)(K-1)	σ_e^2
$\varepsilon_{l(ijk)}$	IJK(L-1)	σ_e^2

    And we see that we can pool AB, AC and ABC with the error term.

    (b) *C, BC and ABC insignificant at the 0.25 level.*

    If we remove each term in the EMS column corrosponding to the interactions that are insignificant we get :

	df	EMS
A_i	(I-1)	$\sigma_e^2 + J\sigma_{AC}^2 + K\sigma_{AB} + KJ\sigma_A^2$
B_j	(J-1)	$\sigma_e^2 + K\sigma_{AB}^2 + \sigma_B^2$
AB_{ij}	(I-1)(J-1)	$\sigma_e^2 + K\sigma_{AB}^2$
C_k	(K-1)	$\sigma_e^2 + J\sigma_{AC}^2$
AC_{ik}	(I-1)(K-1)	$\sigma_e^2 + J\sigma_{AC}^2$
BC_{jk}	(J-1)(K-1)	σ_e^2
ABC_{ijk}	(I-1)(J-1)(K-1)	σ_e^2
$\varepsilon_{l(ijk)}$	IJK(L-1)	σ_e^2

This shows that we can pool BC and ABC with the error term. We cannot pool C since AC is not insignificant at the 0.25 level.

(c) *AC and BC both insignificant at the 0.25 level.*

	df	EMS
A_i	(I-1)	$\sigma_e^2 + \sigma_{ABC}^2 + K\sigma_{AB} + KJ\sigma_A^2$
B_j	(J-1)	$\sigma_e^2 + \sigma_{ABC}^2 + K\sigma_{AB}^2 + IK\sigma_B^2$
AB_{ij}	(I-1)(J-1)	$\sigma_e^2 + \sigma_{ABC}^2 + K\sigma_{AB}^2$
C_k	(K-1)	$\sigma_e^2 + \sigma_{ABC}^2 + IJ\sigma_C^2$
AC_{ik}	(I-1)(K-1)	$\sigma_e^2 + \sigma_{ABC}^2$
BC_{jk}	(J-1)(K-1)	$\sigma_e^2 + \sigma_{ABC}^2$
ABC_{ijk}	(I-1)(J-1)(K-1)	$\sigma_e^2 + \sigma_{ABC}^2$
$\varepsilon_{l(ijk)}$	IJK(L-1)	σ_e^2

Cannot pool any of the interactions.

(d) *B, AB, BC and ABC insignificant at the 0.25 level*

	df	EMS
A_i	(I-1)	$\sigma_e^2 + J\sigma_{AC}^2 + KJ\sigma_A^2$
B_j	(J-1)	σ_e^2
AB_{ij}	(I-1)(J-1)	σ_e^2
C_k	(K-1)	$\sigma_e^2 + J\sigma_{AC}^2 + IJ\sigma_C^2$
AC_{ik}	(I-1)(K-1)	$\sigma_e^2 + J\sigma_{AC}^2$
BC_{jk}	(J-1)(K-1)	σ_e^2
ABC_{ijk}	(I-1)(J-1)(K-1)	σ_e^2
$\varepsilon_{l(ijk)}$	IJK(L-1)	σ_e^2

Pool B, AB, BC and ABC with the error term.

(e) *A and ABC insignificant at the 0.25 level.*

	df	EMS
A_i	(I-1)	$\sigma_e^2 + J\sigma_{AC}^2 + K\sigma_{AB}$
B_j	(J-1)	$\sigma_e^2 + I\sigma_{BC}^2 + K\sigma_{AB}^2 + IK\sigma_B^2$
AB_{ij}	(I-1)(J-1)	$\sigma_e^2 + K\sigma_{AB}^2$
C_k	(K-1)	$\sigma_e^2 + I\sigma_{BC}^2 + J\sigma_{AC}^2 + IJ\sigma_C^2$
AC_{ik}	(I-1)(K-1)	$\sigma_e^2 + J\sigma_{AC}^2$
BC_{jk}	(J-1)(K-1)	$\sigma_e^2 + I\sigma_{BC}^2$
ABC_{ijk}	(I-1)(J-1)(K-1)	σ_e^2
$\varepsilon_{l(ijk)}$	IJK(L-1)	σ_e^2

Pool ABC with the error term.

25. *Give formulae for each of the variance terms estimable in the design summarized in Table 1.*

Here I will copy table 1 and add a column for the F statistic.

	df	EMS	F_{calc}
A_i	1	$\sigma^2 + 6\sigma_{AB}^2 + 24\Phi(A)$	$MS(A)/MS(AB)$
B_j	3	$\sigma^2 + 12\sigma_B^2$	$MS(B)/MS(\varepsilon)$
AB_{ij}	3	$\sigma^2 + 6\sigma_{AB}^2$	$MS(AB)/MS(\varepsilon)$
C_k	2	$\sigma^2 + 4\sigma_{BC}^2 + 16\Phi(C)$	$MS(C)/MS(BC)$
AC_{ik}	2	$\sigma^2 + 2\sigma_{ABC}^2 + 8\Phi(AC)$	$MS(AC)/MS(ABC)$
BC_{jk}	6	$\sigma^2 + 4\sigma_{BC}^2$	$MS(BC)/MS(\varepsilon)$
ABC_{ijk}	6	$\sigma^2 + 2\sigma_{ABC}^2$	$MS(ABC)/MS(\varepsilon)$
$\varepsilon_{l(ijk)}$	24	σ^2	

Table 1: Copy of table 3.4.4

Then we get the following equations for the estimable variances.

$$\begin{aligned}
\sigma_B^2 &= [MS(B) - MS(\varepsilon)]/12 \\
\sigma_{AB}^2 &= [MS(AB) - MS(\varepsilon)]/6 \\
\sigma_{BC}^2 &= [MS(BC) - MS(\varepsilon)]/4 \\
\sigma_{ABC}^2 &= [MS(ABC) - MS(\varepsilon)]/2
\end{aligned}$$

26. *Use the design summarized in Table 1 to find $\sigma^2_{\bar{y}_{AC1}-\bar{y}_{AC2}}$ used to compare two or more means in the AC_{ik} interaction and to find $\sigma^2_{\bar{y}_{AC}}$ used for confidence intervals on a single AC_{ik} mean. Give the appropriate estimates and degrees of freedom.*

The theoretical standard error for the comparison of means is

$$\sigma^2_{\bar{y}_{AC1}-\bar{y}_{AC2}} = (\sigma^2 + 2\sigma^2_{ABC})/8 \tag{1}$$

With the estimate

$$s^2_{\bar{y}_{AC1}-\bar{y}_{AC2}} = MS(ABC)/8 \tag{2}$$

and 6 degrees of freedom. The expressions for the variance of a particular AC_{ij} follow.

$$\sigma^2_{\bar{y}_{AC}} = \frac{\sigma^2_B}{J} + \frac{(I-1)\sigma^2_{AB}}{IJ} + \frac{(K-1)\sigma^2_{BC}}{JK} + \frac{(I-1)(K-1)\sigma^2_{ABC}}{IJK} + \frac{\sigma^2}{JL} \tag{3}$$

Then the equation for the estimator is

$$s^2_{\bar{y}_{AC}} = \frac{[MS(B) - MS(\varepsilon)]/12}{4} + \frac{[MS(AB) - MS(\varepsilon)]/6}{8} \tag{4}$$

$$+ \frac{2[MS(BC) - MS(\varepsilon)]/4}{12} + \frac{2[MS(ABC) - MS(\varepsilon)]/2}{24} + \frac{MS(\varepsilon)}{12} \tag{5}$$

Which reduces to

$$s^2_{\bar{y}_{AC}} = \frac{MS(B) + MS(AB) + 2MS(BC) + 2MS(ABC)}{48} \tag{6}$$

Then we get the following approximation for the degrees of freedom.

$$df \approx \frac{\left\{\frac{[MS(B)+MS(AB)+2MS(BC)+2MS(ABC)]}{48}\right\}^2}{\frac{(MS(B)/48)^2}{3} + \frac{(MS(AB)/48)^2}{3} + \frac{(MS(BC)/24)^2}{6} + \frac{(MS(ABC)/24)^2}{6}} \tag{7}$$

27. *Use the design summarized in Table 1 to find the coefficients appropriate for the $A_{linear} \times C_{quadratic}$ interaction. What term is used to test $A_{linear} \times C_{quadratic}$?*

    To find the coefficients to test for the $A_{linear} \times C_{quadratic}$ relationship, we look up the coefficients in Appendix 5 and multiply them together. Then we get the following table.

Interaction	$\text{Linear}_{I=2}$	$\text{Quadratic}_{K=3}$	Coefficient
AC_{11}	-1	1	-1
AC_{12}	-1	-2	2
AC_{13}	-1	1	-1
AC_{21}	1	1	1
AC_{22}	1	-2	-2
AC_{23}	1	1	1

    Because the interaction term AC_{ik} is tested using ABC_{ijk}, $A_{linear} \times C_{quadratic}$ is tested using $A_{linear} \times B \times C_{quadratic}$. With df $1 \times (J-1) = 3$.

35. *It visually appears as if there is a quadratic trend in Figure 3.3. Repeat the plot using 95% confidence intervals. It is now easy to see why the quadratic trend was not significant. This points out the danger of plotting just the mean values and not the confidence intervals.*

    We need first to calculate the confidence intervals. Following the instructions given in the text, we get the following for the theoretical error term.

$$\sigma^2_{\bar{y}_{Fe}} = \frac{\sigma^2_{Fi}}{J} + \frac{(K-1)\sigma^2_{FiFe}}{IJ} + \frac{\sigma^2}{IK} \tag{8}$$

$$= \frac{\sigma^2_{Fi}}{2} + \frac{4\sigma^2_{FiFe}}{10} + \frac{\sigma^2}{10} \tag{9}$$

And we have the following calculations for the estimator $s^2_{\bar{y}_{Fe}}$.

$$s^2_{\bar{y}_{Fe}} = \frac{(MS(Fi) - MS(\varepsilon))/25}{2} + \frac{4(MS(FiFe) - MS(\varepsilon))/5}{10} + \frac{MS(\varepsilon)}{10} \quad (10)$$

$$= \frac{MS(Fi) + 4MS(FiFe)}{50} = \frac{124.82 + 4(69.87)}{50} \quad (11)$$

$$= 8.086 \quad (12)$$

Then we calculate the expression for the degrees of freedom.

$$df \approx \frac{8.086^2}{\frac{(MS(Fi)/50)^2}{1} + \frac{(4MS(FiFe)/50)^2}{4}} \quad (13)$$

$$= \frac{65.383}{6.232 + 31.243} = 1.744 \quad (14)$$

To find the appropriate $t_{\alpha/2}$ statistic, we use linear interpolation between 1 and 2 degrees of freedom to find $t_{0.05} = 3.79$. This gives the confidence interval $\bar{y}_{\cdot j \cdot} \pm 11.876$. The following table shows the values needed to plot the points along with the associated confidence intervals.

Fe	yield	left	right
0	22.3	10.424	34.176
10	30.4	18.524	42.276
20	40.6	28.724	52.476
30	40.1	28.224	51.976
40	41.7	29.824	53.576

The plot in Figure 1 shows that it is not probable that there is a quadratic effect. It appears that after the Fertilizer level reaches 20, there is no effect.

51. *Compute Δ for each of the terms in Table 3.9.7 using the approximation to the approximate df given in (3.10.8) at the end of Section 3.10 where appropriate. (Do not use the pooled model of the actual MS's since the Δ's are supposed to be calculated before any data is collected.)*

Calculations for $H_0 : \sigma^2_E = 0$.

$$F_{calc} = \frac{MS(E)}{MS(EI) + MS(EF) - MS(EIF)} \quad (15)$$

This will give numerator $df = 3$ and we use the approximation given in (3.10.8) to find the df for the denominator.

$$df_{harmonic} = \frac{3}{\sum_{i=1}^{3} 1/df_i} = \frac{3}{1/9 + 1/3 + 1/9} = 5.4 \quad (16)$$

We can use linear interpolation between 5 and 6 degrees of freedom from the table in appendix 13 to get $\delta = 5.1214$ (note: the book does not interpolate and uses the value for df=3/5 which is 5.265). $C = 8$, which gives the following answer.

$$\Delta_E = \frac{5.1214}{\sqrt{8}} = 1.81 \quad (17)$$

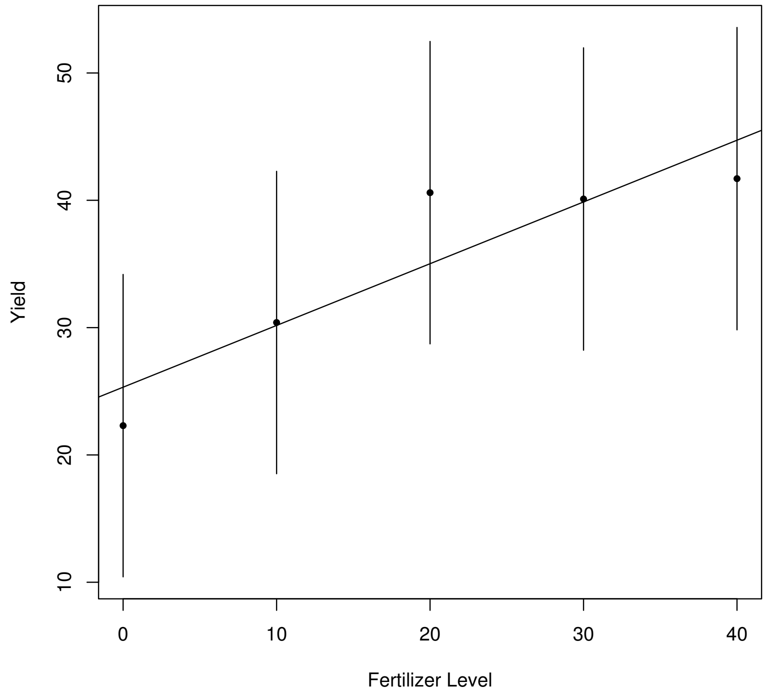

Figure 1: Effect of Fertilizer on Yield

Calculations for $H_0 : \sigma_I^2 = 0$.

$$F_{calc} = \frac{MS(E)}{MS(EI) + MS(IF) - MS(EIF)} \tag{18}$$

This will give numerator $df = 3$ and we use the approximation given in (3.10.8) to find the df for the denominator.

$$df_{harmonic} = \frac{3}{\sum_{i=1}^{3} 1/df_i} = \frac{3}{1/3 + 1/9 + 1/9} = 5.4 \tag{19}$$

Since $C = 8$, we get $\Delta_I = \Delta_E = 1.81$.

Calculations for $H_0 : \sigma_{EI}^2 = 0$.

$$F_{calc} = \frac{MS(EI)}{MS(EIF)} \tag{20}$$

This gives us $df = 9/9$. The table in appendix 13 only gives values for $(6/8, 10/8, 6/10, 10/10)$, if we use linear interpolation between these four values, we get $\delta = 2.654$, and $C = 2$, which gives $\Delta_{EI} = 1.87$.

Calculations for $H_0 : \sigma^2_F = 0$.

$$F_{calc} = \frac{MS(F)}{MS(IF) + MS(EF) - MS(EIF)} \tag{21}$$

$$df_{harmonic} = \frac{3}{\sum_{i=1}^{3} 1/df_i} = \frac{3}{1/3 + 1/3 + 1/9} = 3.857 \tag{22}$$

This gives us $df = 1/3.857$. Then we use interpolation between $df = 1/3$ and $df = 1/4$, to get $\delta = 21.0872$ and $C = 16$, so we get $\Delta_F = 5.272$.

Calculations for $H_0 : \sigma^2_{EF} = 0$.

$$F_{calc} = \frac{MS(EF)}{MS(EIF)} \tag{23}$$

$$df = 3/9 \tag{24}$$

Then we interpolate between $df = (3/8)$ and $df = (3/10)$ to get $\delta = 4.4005$. $C = 4$, so we get $\Delta_{EF} = 2.20$.

Calculations for $H_0 : \sigma^2_{IF} = 0$.

$$F_{calc} = \frac{MS(IF)}{MS(EIF)} \tag{25}$$

$$df = 3/9 \tag{26}$$

Then we interpolate between $df = (3/8)$ and $df = (3/10)$ to get $\delta = 4.4005$. $C = 4$, so we get $\Delta_{IF} = 2.20$.

53. *For the diesel engine example, which factor levels would you consider changing when redesigning the experiment. Why? Suggest a different design and explain your reasoning.*

I would drop the factor for Fuel and do two reps on each engine. It is quite unlikely that there will be significant variation between two tanks of gas from the same manufacturer, we can trust that the manufacturer of the gas has taken care of this for us.

The new model would be

$$y_{ijk} = \mu + E_i + I_j + EI_{ij} + \varepsilon_{k(ij)} \tag{27}$$

And the table below shows that if we make this modification, we now have direct tests for all factors and interactions.

	df	R/i	R/j	R/k	EMS
E_i	3	1	4	2	$\sigma^2 + 2\sigma^2_{EI} + 8\sigma^2_E$
I_j	3	4	1	2	$\sigma^2 + 2\sigma^2_{EI} + 8\sigma^2_I$
EI_{ij}	9	1	1	2	$\sigma^2 + 2\sigma^2_{EI}$
$\varepsilon_{k(ij)}$	16	1	1	1	σ^2

And we can look at the minimal detectible differences in the table below.

Factor	F_{calc}	df	δ	C	Δ
E_i	$MS(E)/MS(EI)$	3/9	4.4005	8	1.56
I_j	$MS(I)/MS(EI)$	3/9	4.4005	8	1.56
EI_{ij}	$MS(EI)/MS(\varepsilon)$	9/16	2.257	2	1.59