STARTED Fri 4/1/22 (week 10, lecture 27)
@ 3 10 min mark

HANDOUT # 9

RANDOM EFFECTS MODELS AND NESTED MODELS

I. Mixed Effects Models: Random and Fixed Factor Levels

   a. Experiments with Random Treatment Levels

   b. Experiments with Fixed and Random Treatment Levels

   c. Derivation of Expected Mean Squares from Random and Mixed Effects Models

   d. Estimation of Variance Components

II. Nested Treatment Structure

   a. Experiments with Nested Treatment Structure

   b. Experiments with Nested and Crossed Treatment Structure

- **Supplemental Reading: Design & ANOVA Book - Ch. 17 & 18**

I. Experiments with Random Treatment Levels

Example of Experiment with Random Treatment Levels

A beverage company has encountered problems with the machinery that fills the containers with products. The amount of beverage placed in the containers varies too much around the specified amount listed on the container. A quality control engineer designs the following study to investigate the level of variability in the amount of beverage injected into the container due to the two major suspected sources: the filling machine and the operator of the machine.

Research Question: The company wanted to identify what proportion of the variability in the amount of beverage placed in the container was due to machinery differences and what proportion was due to operator variation.

Treatment Design: The quality control engineer randomly selected 3 filling machines from the 130 machines used by the company and randomly selected 4 workers from its work force to operate the filling machines. The 12 treatments consist of two factors: Machine M_1, M_2, M_3 with 3 random levels and Operator O_1, O_2, O_3, O_4 with 4 random levels.

Experimental Design: Each operator filled 50 containers using each of the three machines in a random order. That is, the treatments consist of the 12 random combinations: (M_i, O_j). A random sample of 10 containers were then randomly assigned to each of these 12 random combinations (M_i, O_j) (treatments). The 10 containers were then filled with the beverage. The response variable Y_{ijk} is the deviation from the nominal amount on container (ijk). The complete experiment was then repeated 5 times with a new randomization during each of the 5 repeats. There are a total of 600 responses.

It is important to note that in experiments involving treatments having random levels that there are two randomizations involved in the experiment.

1. First the levels of the treatment are randomly selected from a population of treatments.
2. Second, the EU's are randomly assigned to the randomly selected treatments.
3. In some instances, the order in which the responses from the EU's are obtained is also randomized.

Example of Experiment with Fixed and Random Treatment Levels: Mixed Factor Levels

Design and Analysis of Experiments by Gary W. Dehlert

Dental Fillings Experiment

Dental fillings made from gold can vary in hardness depending on how the metal is treated prior to its placement in the tooth. Two factors thought to influence the hardness: the gold alloy and the condensation method. In addition, some dentists performing the dental work are better at some types of filling than others. Five dentists were randomly selected and agreed to participate in the experiment. Each dentist prepares 24 fillings (in random order), one for each of the combinations of condensation method (three levels) and alloy (eight levels). The fillings were then measured for hardness using the Diamond Pyramid Hardness Number (big scores are better). The data are contained in the following table:

		Alloy							
Dentist	Method	1	2	3	4	5	6	7	8
1	1	792	824	813	792	792	907	792	835
1	2	772	772	782	698	665	1115	835	870
1	3	782	803	752	620	835	847	560	585
2	1	803	803	715	803	813	858	907	882
2	2	752	772	772	782	743	933	792	824
2	3	715	707	835	715	673	698	734	681
3	1	715	724	743	627	752	858	762	724
3	2	792	715	813	743	613	824	847	782
3	3	762	606	743	681	743	715	824	681
4	1	673	946	792	743	762	894	792	649
4	2	657	743	690	882	772	813	870	858
4	3	690	245	493	707	289	715	813	312
5	1	634	715	707	698	715	772	1048	870
5	2	649	724	803	665	752	824	933	835
5	3	724	627	421	483	405	536	405	312

The treatments consist of combining a level of the Factor having random levels with the levels of the Factors having fixed levels.

1. First the levels of the factor having random levels are randomly selected from a population of levels, the Dentist.

2. Second the treatments are constructed by combining the levels of the fixed level factors, Method and Gold Alloy, with the randomly selected levels of the Dentist factor.

3. Third, the EU's are randomly assigned to the treatments. In the above example, the 24 fillings are randomly assigned to the Dentist.

4. In some experiments, the order in which the responses from the EU's are obtained is also randomized (the preparation of 24 fillings by each Dentist was done in a random order).

Statistical Model for CRD With Treatments Constructed From Two Factors

We will consider three cases, both factors having fixed levels, both factors having random levels, and one with random levels and one with fixed levels.

<u>Case 1</u> **Both Factors Having Fixed Levels:** Suppose we have two factors F_1 at a fixed levels and F_2 at b fixed levels. There are r EU's randomly assigned to each of the $t = ab$ treatments.

MODEL: $y_{ijk} = \mu + \tau_i + \gamma_j + (\tau\gamma)_{ij} + e_{ijk}$ with

i. e_{ijk}'s iid $N(0, \sigma_e^2)$

ii. $\tau_a = 0, \quad \gamma_b = 0 \quad (\tau\gamma)_{aj} = (\tau\gamma)_{ib} = 0 \quad \text{for all} \quad (i,j)$

The above conditions yield the following results: $y_{ijk} \overset{\mathcal{D}}{\sim} N(\mu_{ij},\ \sigma_e^2)$

a. $\mu_{ij} = E[y_{ijk}] = \mu + \tau_i + \gamma_j + (\tau\gamma)_{ij}$ many of the terms, $\tau_i, \gamma_j, (\tau\gamma)_{ij}$, are 0

b. $Var[y_{ijk}] = \sigma_e^2$

c. y_{ijk}'s are independent normally distributed random variables with the same variance but are not identically distributed due to different means.

d. Goal: Test hypotheses about the interaction and main effects of the two factors using contrasts in the μ_{ij}

e. Goal: Use multiple comparisons, contrasts, other procedures to investigate differences in the μ_{ij}'s

<u>Case 2</u> **Both Factors Having Random Levels :** Suppose we have two factors F_1 at a randomly selected levels and F_2 at b randomly selected levels. There are r EU's randomly assigned to each of the $t = ab$ treatments.

MODEL: $y_{ijk} = \mu + a_i + b_j + (ab)_{ij} + e_{ijk}$ with

i. e_{ijk}'s iid $N(0, \sigma_e^2)$

ii. a_i's iid $N(0, \sigma_{F_1}^2)$, b_i's iid $N(0, \sigma_{F_2}^2)$, and $(ab)_{ij}$'s iid $N(0, \sigma_{F_1 * F_2}^2)$

iii. The random variables a_i, b_j, $(ab)_{ij}$, and e_{ijk} are all independent

The above conditions yield the following results: $y_{ijk} \overset{\mathcal{D}}{\sim} N(\mu,\ \sigma_y^2)$

a. $\mu_{ij} = E[y_{ijk}] = \mu$

b. $\sigma_y^2 = Var[y_{ijk}] = \sigma_{F_1}^2 + \sigma_{F_2}^2 + \sigma_{F_1 * F_2}^2 + \sigma_e^2$

c. y_{ijk}'s are identically distributed but not independent:

$$Cov[y_{ijk}, y_{i'j'k'}] = \begin{cases} \sigma_{F_1}^2 + \sigma_{F_2}^2 + \sigma_{F_1 * F_2}^2 + \sigma_e^2 & if \quad i = i', j = j', k = k' \\ \sigma_{F_1}^2 + \sigma_{F_2}^2 + \sigma_{F_1 * F_2}^2 & if \quad i = i', j = j', k \neq k' \\ \sigma_{F_1}^2 & if \quad i = i', j \neq j' \\ \sigma_{F_2}^2 & if \quad i \neq i', j = j' \\ 0 & if \quad i \neq i', j \neq j' \end{cases}$$

d. Goal: Test hypotheses about the interaction and main effects of the two factors by testing the hypotheses:

Test of Interaction: $H_o : \sigma_{F_1 * F_2} = 0$ versus $H_1 : \sigma_{F_1 * F_2} > 0$

Test of Main effect of F_1: $H_o : \sigma_{F_1} = 0$ versus $H_1 : \sigma_{F_1} > 0$

Test of Main effect of F_2: $H_o : \sigma_{F_2} = 0$ versus $H_1 : \sigma_{F_2} > 0$

e. Proportionally allocate the variability in the measurements, σ_y^2 into the four sources of variation based on the relative sizes of $\sigma_{F_1}^2, \sigma_{F_2}^2, \sigma_{F_1 * F_2}^2$, and σ_e^2

f. Multiple comparisons and contrasts are not appropriate for factors with random factor levels. These methods are appropriate for inferences concerning population means but not for population variances.

<u>Case 3</u> **Mixed Factors: One Factor has Fixed Levels and one Factor Has Random Levels :** Suppose we have one factor F_1 at a fixed levels and F_2 at b randomly selected levels. There are r EU's randomly assigned to each of the $t = ab$ treatments.

MODEL: $y_{ijk} = \mu + \tau_i + b_j + (\tau b)_{ij} + e_{ijk}$ with

i. e_{ijk}'s iid $N(0, \sigma_e^2)$

ii. τ_i's are population parameters with $\tau_a = 0$, b_i's iid $N(0, \sigma_{F_2}^2)$, and $(\tau b)_{ij}$'s iid $N(0, \sigma_{F_1 * F_2}^2)$

iii. The random variables b_j, $(\tau b)_{ij}$, and e_{ijk} are all independent

The above conditions yield the following results: $y_{ijk} \overset{\mathcal{D}}{\sim} N(\mu_{ij},\ \sigma_y^2)$

a. $\mu_{ij} = E[y_{ijk}] = E[\mu + \tau_i + b_j + (\tau b)_{ij} + e_{ijk}] = \mu + \tau_i$

b. $\sigma_y^2 = Var[y_{ijk}] = Var[\mu + \tau_i + b_j + (\tau b)_{ij} + e_{ijk}] = \sigma_{F_2}^2 + \sigma_{F_1 * F_2}^2 + \sigma_e^2$

c. y_{ijk}'s are neither identically distributed (if $\tau_i \neq 0$) nor independent :

$$Cov[y_{ijk}, y_{i'j'k'}] = \begin{cases} \sigma_{F_2}^2 + \sigma_{F_1*F_2}^2 + \sigma_e^2 & if \quad i = i', j = j', k = k' \\ \sigma_{F_2}^2 + \sigma_{F_1*F_2}^2 & if \quad i = i', j = j', k \neq k' \\ \sigma_{F_2}^2 & if \quad i \neq i', j = j' \\ 0 & if \quad j \neq j' \end{cases}$$

d. Goal: Test hypotheses about the interaction and main effects of the two factors by testing the hypotheses:

Test of Interaction: $H_o : \sigma_{F_1 * F_2} = 0$ versus $H_1 : \sigma_{F_1 * F_2} > 0$

Test of Main effect of F_1: $H_o : \tau_i = 0$ for all i versus $H_1 : \tau_i \neq 0$ for at least one i

Test of Main effect of F_2: $H_o : \sigma_{F_2} = 0$ versus $H_1 : \sigma_{F_2} > 0$

e. Proportionally allocate the variability in the measurements, σ_y^2 into the three sources of variation based on the relative sizes of : $\sigma_{F_2}^2, \sigma_{F_1*F_2}^2$, and σ_e^2

f. Use multiple comparisons and contrasts to investigate differences in the levels of the factor having fixed factor levels, F_1. Because F_2 has random levels, we would evaluate the levels of F_1 averaged over the levels of F_2 even when there is significant evidence of an interaction between factors F_1 and F_2. For example, F_1 is type of Machine and F_2 is Operators of Machine, we would still want to compare the average responses of the different types of machines even if there is a Machine by Operator interaction.

For each of the Three Situations: Fixed Effects, Random Effects, and Mixed Effects, we need to determine the appropriate hypotheses and test statistics. In each of the above three types of experiments, the computational form of the sums of squares for the AOV table are identical when there are an equal number of replications for each of the $t = ab$ treatments. However, the expected Mean Squares are different depending on the randomness of the terms in the corresponding model. In order to identify the proper test statistics, it is necessary to determine the expected values of the Mean Squares.

Model I: Both F_1 and F_2 Fixed

Source of Variation	MS	Expected MS	Null Hypothesis	F-Statistic
F_1	MS_{F_1}	$\sigma_e^2 + brQ_{F_1}$	$H_o : Q_{F_1} = 0$	$\frac{MS_{F_1}}{MSE}$
F_2	MS_{F_2}	$\sigma_e^2 + arQ_{F_2}$	$H_o : Q_{F_2} = 0$	$\frac{MS_{F_2}}{MSE}$
$F_1 * F_2$	$MS_{F_1*F_2}$	$\sigma_e^2 + rQ_{F_1*F_2}$	$H_o : Q_{F_1*F_2} = 0$	$\frac{MS_{F_1*F_2}}{MSE}$
Error	MSE	σ_e^2		

where we define the following parameters

$Q_{F_1} = \frac{1}{a-1}\sum_{i=1}^{a}(\bar{\mu}_{i.} - \bar{\mu}_{..})^2$

$Q_{F_2} = \frac{1}{b-1}\sum_{j=1}^{b}(\bar{\mu}_{.j} - \bar{\mu}_{..})^2$

$Q_{F_1*F_2} = \frac{1}{(a-1)(b-1)}\sum_{i=1}^{a}\sum_{j=1}^{b}(\mu_{ij} - \bar{\mu}_{i.} - \bar{\mu}_{.j} + \bar{\mu}_{..})^2$

Model II: Both F_1 and F_2 Random

Source of Variation	MS	Expected MS	Null Hypothesis	F-Statistic
F_1	MS_{F_1}	$\sigma_e^2 + r\sigma_{F_1*F_2}^2 + br\sigma_{F_1}^2$	$H_o : \sigma_{F_1}^2 = 0$	$\frac{MS_{F_1}}{MS_{F_1*F_2}}$
F_2	MS_{F_2}	$\sigma_e^2 + r\sigma_{F_1*F_2}^2 + ar\sigma_{F_2}^2$	$H_o : \sigma_{F_2}^2 = 0$	$\frac{MS_{F_2}}{MS_{F_1*F_2}}$
$F_1 * F_2$	$MS_{F_1*F_2}$	$\sigma_e^2 + r\sigma_{F_1*F_2}^2$	$H_o : \sigma_{F_1*F_2}^2 = 0$	$\frac{MS_{F_1*F_2}}{MSE}$
Error	MSE	σ_e^2		

Model III: F_1-Fixed and F_2-Random

Source of Variation	MS	Expected MS	Null Hypothesis	F-Statistic
F_1	MS_{F_1}	$\sigma_e^2 + r\sigma_{F_1*F_2}^2 + brQ_{F_1}$	$H_o : Q_{F_1} = 0$	$\frac{MS_{F_1}}{MS_{F_1*F_2}}$
F_2	MS_{F_2}	$\sigma_e^2 + r\sigma_{F_1*F_2}^2 + ar\sigma_{F_2}^2$	$H_o : \sigma_{F_2}^2 = 0$	$\frac{MS_{F_2}}{MS_{F_1*F_2}}$
$F_1 * F_2$	$MS_{F_1*F_2}$	$\sigma_e^2 + r\sigma_{F_1*F_2}^2$	$H_o : \sigma_{F_1*F_2}^2 = 0$	$\frac{MS_{F_1*F_2}}{MSE}$
Error	MSE	σ_e^2		

WARNING: If the conditions on the model parameters in the mixed model are changed from $\tau_a = 0$ to $\sum_{i=1}^{a} \tau_i = 0$ and $\sum_{i=1}^{a} (\tau b)_{ij} = 0$ for $j = 1, \ldots, b$, then the above results are changed to the following table:

Model III**: F_1-Fixed and F_2-Random

Source of Variation	MS	Expected MS	Null Hypothesis	F-Statistic
F_1	MS_{F_1}	$\sigma_e^2 + r\sigma_{F_1*F_2}^2 + brQ_{F_1}$	$H_o : Q_{F_1} = 0$	$\frac{MS_{F_1}}{MS_{F_1*F_2}}$
F_2	MS_{F_2}	$\sigma_e^2 + ar\sigma_{F_2}^2$	$H_o : \sigma_{F_2}^2 = 0$	$\frac{MS_{F_2}}{MSE}$
$F_1 * F_2$	$MS_{F_1*F_2}$	$\sigma_e^2 + r\sigma_{F_1*F_2}^2$	$H_o : \sigma_{F_1*F_2}^2 = 0$	$\frac{MS_{F_1*F_2}}{MSE}$
Error	MSE	σ_e^2		

The derivation of the Expected Mean Squares will be illustrated as follows.

Derivation of Expected Mean Squares for a CRD with an $a \times b$ Treatment Structure and r Reps per Treatment (Equally replicated)

Suppose we have two factors F_1 with a levels and F_2 with b levels. There are r observations per treatment. The methodology for computing the Expected Mean Squares will be illustrated by computing $E[MS_{F_1}]$ under four possible Models:

$$MS_{F_1} = \frac{rb}{a-1}\sum_{i=1}^{a}(\bar{y}_{i..} - \bar{y}_{...})^2 \quad \Rightarrow EMS_{F_1} = E[MS_{F_1}] = \frac{rb}{a-1}\sum_{i=1}^{a} E[(\bar{y}_{i..} - \bar{y}_{...})^2]$$

MODEL 1: F_1 - Fixed Factor Levels and F_2 - Fixed Factor Levels

CLAIM: $EMS_{F_1} = brQ_{F_1} + \sigma_e^2$

Proof of CLAIM:

$y_{ijk} = \mu_{ij} + e_{ijk}$ with e_{ijk}'s iid $N(0, \sigma_e^2)$ random variables

$\bar{y}_{i..} = \bar{\mu}_{i.} + \bar{e}_{i..} \qquad \bar{y}_{...} = \bar{\mu}_{..} + \bar{e}_{...} \quad \Rightarrow \quad \bar{y}_{i..} - \bar{y}_{...} = (\bar{\mu}_{i.} - \bar{\mu}_{..}) + (\bar{e}_{i..} - \bar{e}_{...})$

$$\begin{aligned}(a.)\ \ E[(\bar{y}_{i..} - \bar{y}_{...})^2] &= (\bar{\mu}_{i.} - \bar{\mu}_{..})^2 + 2(\bar{\mu}_{i.} - \bar{\mu}_{..})E[(\bar{e}_{i..} - \bar{e}_{...})] + E[(\bar{e}_{i..} - \bar{e}_{...})^2] \\ &= (\bar{\mu}_{i.} - \bar{\mu}_{..})^2 + 0 + Var[\bar{e}_{i..} - \bar{e}_{...}]\end{aligned}$$

$\bar{e}_{i..} - \bar{e}_{...} = \bar{e}_{i..} - \frac{1}{a}\sum_{m=1}^{a}\bar{e}_{m..} = (1 - \frac{1}{a})\bar{e}_{i..} - \frac{1}{a}\sum_{m \neq i}^{a}\bar{e}_{m..} \quad \Rightarrow$

$$\begin{aligned}(b.)\ \ Var[\bar{e}_{i..} - \bar{e}_{...}] &= \left(1 - \frac{1}{a}\right)^2 Var[\bar{e}_{i..}] + \frac{1}{a^2}\sum_{m \neq i} Var[\bar{e}_{m..}] \\ &= \left(1 - \frac{1}{a}\right)^2 \frac{\sigma_e^2}{br} + \frac{1}{a^2}\sum_{m \neq i}\frac{\sigma_e^2}{br} \\ &= \left(\frac{a-1}{a}\right)^2 \frac{\sigma_e^2}{br} + \left(\frac{a-1}{a^2}\right)\frac{\sigma_e^2}{br} = \left(\frac{a-1}{abr}\right)\sigma_e^2\end{aligned}$$

From (a.) and (b.), we have

$$E[(\bar{y}_{i..} - \bar{y}_{...})^2] = (\bar{\mu}_{i.} - \bar{\mu}_{..})^2 + \left(\frac{a-1}{abr}\right)\sigma_e^2 \quad \Rightarrow$$

$$EMS_{F_1} = \frac{rb}{a-1}\sum_{i=1}^{a} E[(\bar{y}_{i..} - \bar{y}_{...})^2] = \frac{br}{a-1}\sum_{i=1}^{a}\left[(\bar{\mu}_{i.} - \bar{\mu}_{..})^2 + \left(\frac{a-1}{abr}\right)\sigma_e^2\right] = brQ_{F_1} + \sigma_e^2$$

MODEL 2: F_1 **- Random Factor Levels and** F_2 **- Random Factor Levels**

CLAIM: $EMS_{F_1} = br\sigma^2_{F_1} + r\sigma^2_{F_1*F_2} + \sigma^2_e$

Proof of CLAIM:

$$y_{ijk} = \mu + a_i + b_j + (ab)_{ij} + e_{ijk}$$

with $a_i, b_j, (ab)_{ij}, e_{ijk}$'s independent normally distributed random variables with expectation equal to 0 and variances $\sigma^2_{F_1}, \sigma^2_{F_2}, \sigma^2_{F_1*F_2}$, and σ^2_e, respectively.

$$\bar{y}_{i..} = \mu + a_i + \bar{b}_{.} + (\bar{ab})_{i.} + \bar{e}_{i..} \text{ and } \bar{y}_{...} = \mu + \bar{a}_{.} + \bar{b}_{.} + (\bar{ab})_{..} + \bar{e}_{...} \quad \Rightarrow$$

$$\bar{y}_{i..} - \bar{y}_{...} = (a_i - \bar{a}_{.}) + ((\bar{ab})_{i.} - (\bar{ab})_{..}) + (\bar{e}_{i..} - \bar{e}_{...})$$

$$\begin{aligned}
E[(\bar{y}_{i..} - \bar{y}_{...})^2] &= E[(a_i - \bar{a}_{.})^2] + E[((\bar{ab})_{i.} - (\bar{ab})_{..})^2] + E[(\bar{e}_{i..} - \bar{e}_{...})^2] \\
&+ 2E[(a_i - \bar{a}_{.})]E[(\bar{e}_{i..} - \bar{e}_{...})] + 2E[((\bar{ab})_{i.} - (\bar{ab})_{..})]E[(\bar{e}_{i..} - \bar{e}_{...})] \\
&+ 2E[(a_i - \bar{a}_{.})]E[((\bar{ab})_{i.} - (\bar{ab})_{..})] \\
&= Var\left[\left(1 - \frac{1}{a}\right) a_i - \frac{1}{a}\sum_{m \neq i} a_m\right] \\
&+ Var\left[\left(1 - \frac{1}{a}\right) (\bar{ab})_{i.} - \frac{1}{a}\sum_{m \neq i} (\bar{ab})_{m.}\right] \\
&+ Var\left[(1 - \frac{1}{a})\bar{e}_{i..} + \frac{1}{a}\sum_{m \neq i} \bar{e}_{m..}\right] + 0 \\
&= \left[\left(\frac{a-1}{a}\right)^2 \sigma^2_{F_1} + \left(\frac{a-1}{a^2}\right)\sigma^2_{F_1}\right] \\
&+ \left[\left(\frac{a-1}{a}\right)^2 \frac{\sigma^2_{F_1*F_2}}{b} + \left(\frac{a-1}{a^2}\right)\frac{\sigma^2_{F_1*F_2}}{b}\right] \\
&+ \left[\left(\frac{a-1}{a}\right)^2 \frac{\sigma^2_e}{br} + \left(\frac{a-1}{a^2}\right)\frac{\sigma^2_e}{br}\right] \\
&= \left(\frac{a-1}{a}\right)\sigma^2_{F_1} + \left(\frac{a-1}{ab}\right)\sigma^2_{F_1*F_2} + \left(\frac{a-1}{abr}\right)\sigma^2_e \quad \Rightarrow
\end{aligned}$$

$$\begin{aligned}
EMS_{F_1} &= \tfrac{rb}{a-1}\textstyle\sum_{i=1}^{a} E[(\bar{y}_{i..} - \bar{y}_{...})^2] \\
&= \tfrac{br}{a-1}\textstyle\sum_{i=1}^{a}\left[\left(\tfrac{a-1}{a}\right)\sigma^2_{F_1} + \left(\tfrac{a-1}{ab}\right)\sigma^2_{F_1*F_2} + \left(\tfrac{a-1}{abr}\right)\sigma^2_e\right] \\
&= rb\sigma^2_{F_1} + r\sigma^2_{F_1*F_2} + \sigma^2_e
\end{aligned}$$

MODEL 3: F_1 - Fixed Factor Levels and F_2 - Random Factor Levels

CLAIM: $EMS_{F_1} = brQ_{F_1} + r\sigma^2_{F_1*F_2} + \sigma^2_e$

Proof of CLAIM:

$$y_{ijk} = \mu + \tau_i + b_j + (\tau b)_{ij} + e_{ijk}$$

with $\tau_a = 0$ and $b_j, (\tau b)_{ij}, e_{ijk}$'s independent normally distributed random variables with expectation equal to 0 and variances $\sigma^2_{F_2}, \sigma^2_{F_1*F_2}$, and σ^2_e, respectively.

$$\bar{y}_{i..} = \mu + \tau_i + \bar{b}_{.} + (\bar{\tau b})_{i.} + \bar{e}_{i..} \qquad \bar{y}_{...} = \mu + \bar{\tau}_{.} + \bar{b}_{.} + (\bar{\tau b})_{..} + \bar{e}_{...} \quad \Rightarrow$$

$$\bar{y}_{i..} - \bar{y}_{...} = (\tau_i - \bar{\tau}_{.}) + ((\bar{\tau b})_{i.} - (\bar{\tau b})_{..}) + (\bar{e}_{i..} - \bar{e}_{...}) \quad \Rightarrow$$

$$\begin{aligned}
E[(\bar{y}_{i..} - \bar{y}_{...})^2] &= (\tau_i - \bar{\tau}_{.})^2 + E[(\bar{\tau b})_{i.} - (\bar{\tau b})_{..})^2] + E[(\bar{e}_{i..} - \bar{e}_{...})^2] \\
&+ 2(\tau_i - \bar{\tau}_{.})E[(\bar{e}_{i..} - \bar{e}_{...})] + 2E[((\bar{\tau b})_{i.} - (\bar{\tau b})_{..})]E[(\bar{e}_{i..} - \bar{e}_{...})] \\
&+ 2(\tau_i - \bar{\tau}_{.})E[((\bar{\tau b})_{i.} - (\bar{\tau b})_{..})] \\
&= (\tau_i - \bar{\tau}_{.})^2 + Var\left[\left(1 - \frac{1}{a}\right)(\bar{\tau b})_{i.} - \frac{1}{a}\sum_{m \neq i}(\bar{\tau b})_{m.}\right] \\
&+ Var\left[(1 - \frac{1}{a})\bar{e}_{i..} + \frac{1}{a}\sum_{m \neq i}\bar{e}_{m..}\right] + 0 \\
&= (\tau_i - \bar{\tau}_{.})^2 + \left[\left(\frac{a-1}{a}\right)^2 \frac{\sigma^2_{F_1*F_2}}{b} + \left(\frac{a-1}{a^2}\right)\frac{\sigma^2_{F_1*F_2}}{b}\right] \\
&+ \left[\left(\frac{a-1}{a}\right)^2 \frac{\sigma^2_e}{br} + \left(\frac{a-1}{a^2}\right)\frac{\sigma^2_e}{br}\right] \\
&= (\tau_i - \bar{\tau}_{.})^2 + \left(\frac{a-1}{ab}\right)\sigma^2_{F_1*F_2} + \left(\frac{a-1}{abr}\right)\sigma^2_e \quad \Rightarrow
\end{aligned}$$

$$\begin{aligned}
EMS_{F_1} &= \tfrac{rb}{a-1}\textstyle\sum_{i=1}^a E[(\bar{y}_{i..} - \bar{y}_{...})^2] \\
&= \tfrac{rb}{a-1}\textstyle\sum_{i=1}^a \left[(\tau_i - \bar{\tau}_{.})^2 + \left(\frac{a-1}{ab}\right)\sigma^2_{F_1*F_2} + \left(\frac{a-1}{abr}\right)\sigma^2_e\right] \\
&= rbQ_{F_1} + r\sigma^2_{F_1*F_2} + \sigma^2_e
\end{aligned}$$

where $Q_{F_1} = \frac{rb}{a-1}\sum_{i=1}^a(\tau_i - \bar{\tau}_{.})^2$

MODEL 4: F_1 - Random Factor Levels and F_2 - Fixed Factor Levels

CLAIM: $EMS_{F_1} = br\sigma^2_{F_1} + r\sigma^2_{F_1*F_2} + \sigma^2_e$

Proof of CLAIM:

$$y_{ijk} = \mu + a_i + \gamma_j + (a\gamma)_{ij} + e_{ijk}$$

with $\gamma_b = 0$ and $a_i, (a\gamma)_{ij}, e_{ijk}$'s independent normally distributed random variables with expectation equal to 0 and variances $\sigma^2_{F_1}, \sigma^2_{F_1*F_2}$, and σ^2_e, respectively.

$$\bar{y}_{i..} = \mu + s_i + \bar{\gamma}_. + (\bar{a\gamma})_{i.} + \bar{e}_{i..} \qquad \bar{y}_{...} = \mu + \bar{a}_. + \bar{\gamma}_. + (\bar{a\gamma})_{..} + \bar{e}_{...} \quad \Rightarrow$$

$$\begin{aligned}
E[(\bar{y}_{i..} - \bar{y}_{...})^2] &= E[(a_i - \bar{a}_.)^2] + E[((\bar{a\gamma})_{i.} - (\bar{a\gamma})_{..})^2] + E[(\bar{e}_{i..} - \bar{e}_{...})^2] \\
&+ 2E[(a_i - \bar{a}_.)]E[(\bar{e}_{i..} - \bar{e}_{...})] + 2E[((\bar{a\gamma})_{i.} - (\bar{a\gamma})_{..})]E[(\bar{e}_{i..} - \bar{e}_{...})] \\
&+ 2E[(a_i - \bar{a}_.)]E[((\bar{a\gamma})_{i.} - (\bar{a\gamma})_{..})] \\
&= Var\left[\left(1 - \frac{1}{a}\right) a_i - \frac{1}{a}\sum_{m \neq i} \bar{a}_{m.}\right] \\
&+ Var\left[\left(1 - \frac{1}{a}\right) (\bar{a\gamma})_{i.} - \frac{1}{a}\sum_{m \neq i} (\bar{a\gamma})_{m.}\right] \\
&+ Var\left[(1 - \frac{1}{a})\bar{e}_{i..} + \frac{1}{a}\sum_{m \neq i} \bar{e}_{m..}\right] + 0 \\
&= \left[\left(\frac{a-1}{a}\right)^2 \sigma^2_{F_1} + \left(\frac{a-1}{a^2}\right)\frac{\sigma^2_{F_1*F_2}}{b}\right] \\
&+ \left[\left(\frac{a-1}{a}\right)^2 \frac{\sigma^2_{F_1*F_2}}{b} + \left(\frac{a-1}{a^2}\right)\sigma^2_{F_1*F_2}\right] \\
&+ \left[\left(\frac{a-1}{a}\right)^2 \frac{\sigma^2_e}{br} + \left(\frac{a-1}{a^2}\right)\frac{\sigma^2_e}{br}\right] \\
&= \left(\frac{a-1}{a}\right)\sigma^2_{F_1} + \left(\frac{a-1}{ab}\right)\sigma^2_{F_1*F_2} + \left(\frac{a-1}{abr}\right)\sigma^2_e \quad \Rightarrow
\end{aligned}$$

$$\begin{aligned}
EMS_{F_1} &= \tfrac{rb}{a-1}\textstyle\sum_{i=1}^a E[(\bar{y}_{i..} - \bar{y}_{...})^2] \\
&= \tfrac{br}{a-1}\textstyle\sum_{i=1}^a \left[\left(\tfrac{a-1}{a}\right)\sigma^2_{F_1} + \left(\tfrac{a-1}{ab}\right)\sigma^2_{F_1*F_2} + \left(\tfrac{a-1}{abr}\right)\sigma^2_e\right] \\
&= rb\sigma^2_{F_1} + r\sigma^2_{F_1*F_2} + \sigma^2_e
\end{aligned}$$

Statistical Model for CRD With Three Factors

We will consider four cases,

1. all factors having fixed levels
2. all factors having random levels
3. one factor with random levels and two with fixed levels
4. two factors with random levels and one with fixed levels

<u>Case 1</u> **All Factors Having Fixed Levels:** Suppose we have three factors F_1 with a fixed levels, F_2 with b fixed levels, and F_3 with c fixed levels. There are r EU's randomly assigned to each of the $t = abc$ treatments.

MODEL: $y_{ijkl} = \mu + \tau_i + \gamma_j + \delta_k + (\tau\gamma)_{ij} + (\tau\delta)_{ik} + (\gamma\delta)_{jk} + (\tau\gamma\delta)_{ijk} + e_{ijkl}$ with

i. e_{ijkl}'s iid $N(0, \sigma_e^2)$

ii. $\tau_a = 0, \quad \gamma_b = 0 \quad \delta_c = 0$

iii. $(\tau\gamma)_{aj} = (\tau\gamma)_{ib} = (\tau\delta)_{ak} = (\tau\delta)_{ic} = (\gamma\delta)_{jc} = (\gamma\delta)_{bk} = 0 \quad \text{for all} \quad (i,j,k)$

iv. $(\tau\gamma\delta)_{ajk} = (\tau\gamma\delta)_{ibk} = (\tau\gamma\delta)_{ijc} = 0 \quad \text{for all} \quad (i,j,k)$

The above conditions yield the following results

a. $\mu_{ijk} = E[y_{ijkl}] = \mu + \tau_i + \gamma_j + \delta_k + (\tau\gamma)_{ij} + (\tau\delta)_{ik} + (\gamma\delta)_{jk} + (\tau\gamma\delta)_{ijk}$ with a large number of these terms equal to 0

b. $Var[y_{ijkl}] = \sigma_e^2$

c. y_{ijkl}'s are independently distributed as $N(\mu_{ijk}, \sigma_e^2)$ r.v.'s, that is, independent normally distributed random variables with the same variance but not identically distributed due to potentially different means, μ_{ijk}.

That is, $y_{ijkl} \overset{\mathcal{D}}{\sim} N(\mu_{ijk},\ \sigma_e^2)$

__Case 2__ **All Factors Having Random Levels :** Suppose we have three factors F_1 with a randomly selected levels, F_2 with b randomly selected levels, and F_3 with c randomly selected levels. There are r EU's randomly assigned to each of the $t = abc$ treatments.

MODEL: $y_{ijkl} = \mu + a_i + b_j + c_k + (ab)_{ij} + (ac)_{ik} + (bc)_{jk} + (abc)_{ijk} + e_{ijkl}$ with

i. e_{ijkl}'s iid $N(0, \sigma_e^2)$

ii. a_i's iid $N(0, \sigma_{F_1}^2)$, b_j's iid $N(0, \sigma_{F_2}^2)$, and c_k's iid $N(0, \sigma_{F_3}^2)$

iii. $(ab)_{ij}$'s iid $N(0, \sigma_{F_1*F_2}^2)$, $(ac)_{ik}$'s iid $N(0, \sigma_{F_1*F_3}^2)$, $(bc)_{jk}$'s iid $N(0, \sigma_{F_2*F_3}^2)$, $(abc)_{ijk}$'s iid $N(0, \sigma_{F_1*F_2*F_3}^2)$,

iv. The random variables $a_i, b_j, c_k, (ab)_{ij}, (ac)_{ik}, (bc)_{jk}, (abc)_{ijk}$, and e_{ijkl} are all independent

The above conditions yield the following results

a. $\mu_{ijk} = E[y_{ijkl}] = \mu$

b. $\sigma_y^2 = Var[y_{ijkl}] = \sigma_{F_1}^2 + \sigma_{F_2}^2 + \sigma_{F_3}^2 + \sigma_{F_1*F_2}^2 + \sigma_{F_1*F_3}^2 + \sigma_{F_2*F_3}^2 + \sigma_{F_1*F_2*F_3}^2 + \sigma_e^2$

c. y_{ijkl}'s are identically distributed as $N(\mu, \sigma_y^2)$ r.v.'s but they are not independent because of the following correlation:

$$\begin{aligned}
\text{Cov}(y_{ijk\ell}, y_{i'j'k'\ell'}) &= E[(y_{ijk\ell} - \mu)(y_{i'j'k'ell'} - \mu)] \\
&= E[(a_i + b_j + c_k + (ab)_{ij} + (ac)_{ik} + (bc)_{jk} + (abc)_{ijk} + e_{ijkl}) * \\
&\quad (a_{i'} + b_{j'} + c_{k'} + (ab)_{i'j'} + (ac)_{i'k'} + (bc)_{j'k'} + (abc)_{i'j'k'} + e_{i'j'k'l'})] \\
&= E[(a_i)(a_{i'})] + E[(b_j)(b_{j'})] + E[(c_k)(c_{k'})] + E[(ab)_{ij}(ab)_{i'j'}] + E[(ac)_{ik}(ac)_{i'k'}] + \\
&\quad E[(bc)_{jk}(bc)_{j'k'}] + E[(abc)_{ijk}(abc)_{i'j'k'}] + E[(e_{ijkl})(e_{i'j'k'l'})] + 0
\end{aligned}$$

We thus have the following expression for the covariance for the various combinations of (i, i', j, j', k, k', l, l')

$$\text{Cov}(y_{ijk\ell}, y_{i'j'k'\ell'}) = \begin{cases} \sigma^2_{F_1} + \sigma^2_{F_2} + \sigma^2_{F_3} + \sigma^2_{F_1*F_2} + \sigma^2_{F_1*F_3} + \sigma^2_{F_2*F_3} + \sigma^2_{F_1*F_2*F_3} + \sigma^2_e & \text{if } i = i', j = j', k = k', l = l' \\ = Var(y_{ijk\ell}) & \\ \sigma^2_{F_1} + \sigma^2_{F_2} + \sigma^2_{F_3} + \sigma^2_{F_1*F_2} + \sigma^2_{F_1*F_3} + \sigma^2_{F_2*F_3} + \sigma^2_{F_1*F_2*F_3} & \text{if } i = i', j = j', k = k', l \neq l' \\ \sigma^2_{F_2} + \sigma^2_{F_3} + \sigma^2_{F_2*F_3} & \text{if } i \neq i', j = j', k = k' \\ \sigma^2_{F_1} + \sigma^2_{F_3} + \sigma^2_{F_1*F_3} & \text{if } i = i', j \neq j', k = k' \\ \sigma^2_{F_1} + \sigma^2_{F_2} + \sigma^2_{F_1*F_2} & \text{if } i = i', j = j', k \neq k' \\ \sigma^2_{F_3} & \text{if } i \neq i', j \neq j', k = k' \\ \sigma^2_{F_2} & \text{if } i \neq i', j = j', k \neq k' \\ \sigma^2_{F_1} & \text{if } i = i', j \neq j', k \neq k' \\ 0 & \text{if } i \neq i', j \neq j', k \neq k' \end{cases}$$

__Case 3__ **Mixed Factors: One Factor has Random Levels and two Factors Have Fixed Levels :** Suppose we have three factors F_1 with a fixed levels, F_2 with b fixed levels, and F_3 with c randomly selected levels. There are r EU's randomly assigned to each of the $t = abc$ treatments.

MODEL: $y_{ijkl} = \mu + \tau_i + \gamma_j + c_k + (\tau\gamma)_{ij} + (\tau c)_{ik} + (\gamma c)_{jk} + (\tau\gamma c)_{ijk} + e_{ijkl}$ with

i. e_{ijkl}'s iid $N(0, \sigma_e^2)$

ii. τ_i's, γ_j's and $(\tau\gamma)_{ij}$'s are population parameters with

$\tau_a = 0, \quad \gamma_b = 0 \quad (\tau\gamma)_{aj} = (\tau\gamma)_{ib} = 0 \quad \text{for all} \quad (i, j)$

iii. c_k's iid $N(0, \sigma^2_{F_3})$, $(\tau c)_{ij}$'s iid $N(0, \sigma^2_{F_1 * F_3})$ $(\gamma c)_{ij}$'s iid $N(0, \sigma^2_{F_2 * F_3})$ $(\tau\gamma c)_{ij}$'s iid $N(0, \sigma^2_{F_1 * F_2 * F_3})$

iii. The random variables $c_k, (\tau c)_{ik}, (\gamma c)_{jk}, (\tau\gamma c)_{ijk},$ and e_{ijkl} are all independent

The above conditions yield the following results

a. $\mu_{ijk} = E[y_{ijk}] = \mu + \tau_i + \gamma_j + (\tau\gamma)_{ij}$

b. $\sigma_y^2 = Var[y_{ijk}] = \sigma^2_{F_3} + \sigma^2_{F_1 * F_3} + \sigma^2_{F_2 * F_3} + \sigma^2_{F_1 * F_2 * F_3} + \sigma_e^2$

c. y_{ijkl}'s are distributed $N(\mu_{ijk}, \sigma_y^2)$ and are thus not identically distributed nor independent because of the following correlation:

$$\begin{aligned}
\text{Cov}(y_{ijkl}, y_{i'j'k'l'}) &= E[(y_{ijkl} - \mu - \tau_i - \gamma_j - (\tau\gamma)_{ij})(y_{i'j'k'l'} - \mu - \tau_{i'} - \gamma j' - (\tau\gamma)_{i'j'})] \\
&= E[(c_k + (ac)_{ik} + (bc)_{jk} + (abc)_{ijk} + e_{ijkl}) * \\
&\quad (c_{k'} + (ac)_{i'k'} + (bc)_{j'k'} + (abc)_{i'j'k'} + e_{i'j'k'l'})] \\
&= E[(c_k)(c_{k'})] + E[(ac)_{ik}(ac)_{i'k'}] + E[(bc)_{jk}(bc)_{j'k'}] + \\
&\quad E[(abc)_{ijk}(abc)_{i'j'k'}] + E[(e_{ijkl})(e_{i'j'k'l'})] + 0
\end{aligned}$$

We thus have the following expression for the covariance for the various combinations of (i, i', j, j', k, k', l, l')

$$\mathrm{Cov}(y_{ijkl}, y_{i'j'k'l'}) = \begin{cases} \sigma^2_{F_3} + \sigma^2_{F_1*F_3} + \sigma^2_{F_2*F_3} + \sigma^2_{F_1*F_2*F_3} + \sigma^2_e & \text{if } i = i', j = j', k = k', l = l' \\ \sigma^2_{F_3} + \sigma^2_{F_1*F_3} + \sigma^2_{F_2*F_3} + \sigma^2_{F_1*F_2*F_3} & \text{if } i = i', j = j', k = k', l \neq l' \\ \sigma^2_{F_3} + \sigma^2_{F_2*F_3} & \text{if } i \neq i', j = j', k = k' \\ \sigma^2_{F_3} + \sigma^2_{F_1*F_3} & \text{if } i = i', j \neq j', k = k' \\ \sigma^2_{F_3} & \text{if } i \neq i', j \neq j', k = k' \\ 0 & \text{if } k \neq k' \end{cases}$$

<u>Case 4</u> **Mixed Factors: One Factor has Fixed Levels and two Factors Have Random Levels :** Suppose we have one factor F_1 with a fixed levels, F_2 with b randomly selected levels, and F_3 with c randomly selected levels. There are r EU's randomly assigned to each of the $t = abc$ treatments.

MODEL: $y_{ijkl} = \mu + \tau_i + b_j + c_k + (\tau b)_{ij} + (\tau c)_{ik} + (bc)_{jk} + (\tau bc)_{ijk} + e_{ijkl}$ with

i. e_{ijkl}'s iid $N(0, \sigma_e^2)$

ii. τ_i's are population parameters with $\tau_a = 0$ $\quad b_j$'s iid $N(0, \sigma^2_{F_2})$ $\quad c_k$'s iid $N(0, \sigma^2_{F_3})$

$(\tau b)_{ij}$'s iid $N(0, \sigma^2_{F_1*F_2})$ $\quad (\tau c)_{ij}$'s iid $N(0, \sigma^2_{F_1*F_3})$ $\quad (bc)_{ij}$'s iid $N(0, \sigma^2_{F_2*F_3})$ $\quad (\tau bc)_{ij}$'s iid $N(0, \sigma^2_{F_1*F_2*F_3})$

iii. The random variables $b_j, c_k, (\tau b)_{ij}, (\tau c)_{ik}, (bc)_{jk}, (\tau bc)_{ijk}$, and e_{ijkl} are all independent

The above conditions yield the following results

a. $\mu_{ijk} = E[y_{ijk}] = \mu + \tau_i$

b. $\sigma_y^2 = Var[y_{ijk}] = \sigma^2_{F_2} + \sigma^2_{F_3} + \sigma^2_{F_1*F_2} + \sigma^2_{F_1*F_3} + \sigma^2_{F_2*F_3} + \sigma^2_{F_1*F_2*F_3} + \sigma_e^2$

c. y_{ijkl}'s are distributed $N(\mu + \tau_i, \sigma_y^2)$ and thus, are not identically distributed nor independent because of the following correlation:

$$
\text{Cov}(y_{ijkl}, y_{i'j'k'l'}) = \begin{cases}
\sigma^2_{F_2} + \sigma^2_{F_3} + \sigma^2_{F_1*F_2} + \sigma^2_{F_1*F_3} + \sigma^2_{F_2*F_3} + \sigma^2_{F_1*F_2*F_3} + \sigma_e^2 & \text{if } i = i', j = j', k = k', l = l' \\
\sigma^2_{F_2} + \sigma^2_{F_3} + \sigma^2_{F_1*F_2} + \sigma^2_{F_1*F_3} + \sigma^2_{F_2*F_3} + \sigma^2_{F_1*F_2*F_3} & \text{if } i = i', j = j', k = k', l \neq l' \\
\sigma^2_{F_2} + \sigma^2_{F_3} + \sigma^2_{F_2*F_3} & \text{if } i \neq i', j = j', k = k' \\
\sigma^2_{F_3} + \sigma^2_{F_1*F_3} & \text{if } i = i', j \neq j', k = k' \\
\sigma^2_{F_2} + \sigma^2_{F_1*F_2} & \text{if } i = i', j = j', k \neq k' \\
\sigma^2_{F_3} & \text{if } i \neq i', j \neq j', k = k' \\
\sigma^2_{F_2} & \text{if } i \neq i', j = j', k \neq k' \\
0 & \text{if } j \neq j', k \neq k'
\end{cases}
$$

For each of the Three Situations: Fixed Effects, Random Effects, and Mixed Effects, we need to determine the appropriate hypotheses and test statistics. In each of the above four types of experiments, the computational form of the sums of squares for the AOV table are identical when there are an equal number of replications for each of the $t = ab$ treatments. However, the expected Mean Squares are different depending on the randomness of the terms in the corresponding model. In order to identify the proper test statistics, it is necessary to determine the expected values of the Mean Squares.

Model I: The Fixed Levels For F_1, F_2 and F_3

Source of Variation	MS	Expected MS	Null Hypothesis	F-Statistic
F_1	MS_{F_1}	$\sigma_e^2 + bcrQ_{F_1}$	$H_o : Q_{F_1} = 0$	$\frac{MS_{F_1}}{MSE}$
F_2	MS_{F_2}	$\sigma_e^2 + acrQ_{F_2}$	$H_o : Q_{F_2} = 0$	$\frac{MS_{F_2}}{MSE}$
F_3	MS_{F_3}	$\sigma_e^2 + abrQ_{F_3}$	$H_o : Q_{F_3} = 0$	$\frac{MS_{F_3}}{MSE}$
$F_1 * F_2$	$MS_{F_1*F_2}$	$\sigma_e^2 + crQ_{F_1*F_2}$	$H_o : Q_{F_1*F_2} = 0$	$\frac{MS_{F_1*F_2}}{MSE}$
$F_1 * F_3$	$MS_{F_1*F_3}$	$\sigma_e^2 + brQ_{F_1*F_3}$	$H_o : Q_{F_1*F_3} = 0$	$\frac{MS_{F_1*F_3}}{MSE}$
$F_2 * F_3$	$MS_{F_2*F_3}$	$\sigma_e^2 + arQ_{F_2*F_3}$	$H_o : Q_{F_2*F_3} = 0$	$\frac{MS_{F_2*F_3}}{MSE}$
$F_1 * F_2 * F_3$	$MS_{F_1*F_1*F_3}$	$\sigma_e^2 + rQ_{F_1*F_2*F_3}$	$H_o : Q_{F_1*F_2*F_3} = 0$	$\frac{MS_{F_1*F_2*F_3}}{MSE}$
Error	MSE	σ_e^2		

where we define the following parameters

$Q_{F_1} = \frac{1}{a-1}\sum_{i=1}^{a}(\bar{\mu}_{i..} - \bar{\mu}_{...})^2, \qquad Q_{F_2} = \frac{1}{b-1}\sum_{j=1}^{b}(\bar{\mu}_{.j.} - \bar{\mu}_{...})^2, \qquad Q_{F_3} = \frac{1}{c-1}\sum_{k=1}^{c}(\bar{\mu}_{..k} - \bar{\mu}_{...})^2$

$Q_{F_1*F_2} = \frac{1}{(a-1)(b-1)}\sum_{i=1}^{a}\sum_{j=1}^{b}(\bar{\mu}_{ij.} - \bar{\mu}_{i..} - \bar{\mu}_{.j.} + \bar{\mu}_{...})^2$

$Q_{F_1*F_3} = \frac{1}{(a-1)(c-1)}\sum_{i=1}^{a}\sum_{k=1}^{c}(\bar{\mu}_{ik.} - \bar{\mu}_{i..} - \bar{\mu}_{..k} + \bar{\mu}_{...})^2$

$Q_{F_2*F_3} = \frac{1}{(b-1)(c-1)}\sum_{j=1}^{b}\sum_{k=1}^{c}(\bar{\mu}_{.jk} - \bar{\mu}_{.j.} - \bar{\mu}_{..k} + \bar{\mu}_{...})^2$

$Q_{F_1*F_2*F_3} = \frac{1}{(a-1)(b-1)(c-1)}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{i=1}^{c}(\mu_{ijk} - \bar{\mu}_{ij.} - \bar{\mu}_{i.k} - \bar{\mu}_{.jk} + \bar{\mu}_{i..} + \bar{\mu}_{.j.} + \bar{\mu}_{..k} - \bar{\mu}_{...})^2$

Model II: All Factors, F_1, F_2 and F_3, are Random

Source of Variation	MS	Expected MS	Null Hypothesis	F-Statistic
F_1	MS_{F_1}	$\sigma_e^2 + r\sigma^2_{F_1*F_2*F_3} + rc\sigma^2_{F_1*F_2} + rb\sigma^2_{F_1*F_3} + rbc\sigma^2_{F_1}$	$H_o : \sigma^2_{F_1} = 0$	$*$
F_2	MS_{F_2}	$\sigma_e^2 + r\sigma^2_{F_1*F_2*F_3} + rc\sigma^2_{F_1*F_2} + ra\sigma^2_{F_2*F_3} + rac\sigma^2_{F_2}$	$H_o : \sigma^2_{F_2} = 0$	$**$
F_3	MS_{F_3}	$\sigma_e^2 + r\sigma^2_{F_1*F_2*F_3} + rb\sigma^2_{F_1*F_3} + ra\sigma^2_{F_2*F_3} + rab\sigma^2_{F_3}$	$H_o : \sigma^2_{F_3} = 0$	$***$
$F_1 * F_2$	$MS_{F_1*F_2}$	$\sigma_e^2 + r\sigma^2_{F_1*F_2*F_3} + rc\sigma^2_{F_1*F_2}$	$H_o : \sigma^2_{F_1*F_2} = 0$	$\frac{MS_{F_1*F_2}}{MS_{F_1*F_2*F_3}}$
$F_1 * F_3$	$MS_{F_1*F_3}$	$\sigma_e^2 + r\sigma^2_{F_1*F_2*F_3} + rb\sigma^2_{F_1*F_3}$	$H_o : \sigma^2_{F_1*F_3} = 0$	$\frac{MS_{F_1*F_3}}{MS_{F_1*F_2*F_3}}$
$F_2 * F_3$	$MS_{F_2*F_3}$	$\sigma_e^2 + r\sigma^2_{F_1*F_2*F_3} + ra\sigma^2_{F_2*F_3}$	$H_o : \sigma^2_{F_2*F_3} = 0$	$\frac{MS_{F_2*F_3}}{MS_{F_1*F_2*F_3}}$
$F_1 * F_2 * F_3$	$MS_{F_1*F_2*F_3}$	$\sigma_e^2 + r\sigma^2_{F_1*F_2*F_3}$	$H_o : \sigma^2_{F_1*F_2*F_3} = 0$	$\frac{MS_{F_1*F_2*F_3}}{MSE}$
Error	$MS_{F_1*F_2*F_3}$	σ_e^2		

Approximate F-tests are obtained by obtaining a linear combination of Mean Squares, M, such that the expected value of M equals the expected value of the MS for corresponding source of variation, e.g., $E[M] = E_{H:_o}[MS_{F_1}]$.

1. For F_1, $* = \frac{MS_{F_1}}{M}$ where $M = MS_{F_1*F_2} + MS_{F_1*F_3} - MS_{F_1*F_2*F_3}$ which follows from

   Under H_o, EMS_{F_1} contains $\sigma^2_{F_1*F_2}$ and $\sigma^2_{F_1*F_3}$

   Try $M = MS_{F_1*F_2} + MS_{F_1*F_3}$ but then $E[M]$ would have $2\sigma_e^2$ and $2\sigma^2_{F_1*F_2*F_3}$

   Try $M = MS_{F_1*F_2} + MS_{F_1*F_3} - MS_{F_1*F_2*F_3}$

   $E[M] = 2\sigma_e^2 + 2r\sigma^2_{F_1*F_2*F_3} + rc\sigma^2_{F_1*F_2} + rb\sigma^2_{F_1*F_3} - (\sigma_e^2 + r\sigma^2_{F_1*F_2*F_3}) = \sigma_e^2 + r\sigma^2_{F_1*F_2*F_3} + rc\sigma^2_{F_1*F_2} + rb\sigma^2_{F_1*F_3}$

   M has df$=\frac{(M)^2}{\frac{\left(MS_{F_1*F_2}\right)^2}{df_{F_1*F_2}} + \frac{\left(MS_{F_1*F_3}\right)^2}{df_{F_1*F_3}} + \frac{\left(-MS_{F_1*F_2*F_3}\right)^2}{df_{F_1*F_2*F_3}}}$

2. For F_2, $** = \frac{MS_{F_2}}{M}$ where $M = MS_{F_1*F_2} + MS_{F_2*F_3} - MS_{F_1*F_2*F_3}$

   Similar justification as in part 1.

3. For F_3, $*** = \frac{MS_{F_3}}{M}$ where $M = MS_{F_1*F_3} + MS_{F_2*F_3} - MS_{F_1*F_2*F_3}$

   Similar justification as in part 1.

Model III: F_1-Fixed, F_2-Random, and F_3-Random

Source of Variation	MS	Expected MS	Null Hypothesis	F-Statistic
F_1	MS_{F_1}	$\sigma_e^2 + r\sigma^2_{F_1*F_2*F_3} + rc\sigma^2_{F_1*F_2} + rb\sigma^2_{F_1*F_3} + rbcQ_{F_1}$	$H_o : Q_{F_1} = 0$	$*$
F_2	MS_{F_2}	$\sigma_e^2 + r\sigma^2_{F_1*F_2*F_3} + rc\sigma^2_{F_1*F_2} + ra\sigma^2_{F_2*F_3} + rac\sigma^2_{F_2}$	$H_o : \sigma^2_{F_2} = 0$	$**$
F_3	MS_{F_3}	$\sigma_e^2 + r\sigma^2_{F_1*F_2*F_3} + rb\sigma^2_{F_1*F_3} + ra\sigma^2_{F_2*F_3} + rab\sigma^2_{F_3}$	$H_o : \sigma^2_{F_3} = 0$	$* * *$
$F_1 * F_2$	$MS_{F_1*F_2}$	$\sigma_e^2 + r\sigma^2_{F_1*F_2*F_3} + rc\sigma^2_{F_1*F_2}$	$H_o : \sigma^2_{F_1*F_2} = 0$	$\frac{MS_{F_1*F_2}}{MS_{F_1*F_2*F_3}}$
$F_1 * F_3$	$MS_{F_1*F_3}$	$\sigma_e^2 + r\sigma^2_{F_1*F_2*F_3} + rb\sigma^2_{F_1*F_3}$	$H_o : \sigma^2_{F_1*F_3} = 0$	$\frac{MS_{F_1*F_3}}{MS_{F_1*F_2*F_3}}$
$F_2 * F_3$	$MS_{F_2*F_3}$	$\sigma_e^2 + r\sigma^2_{F_1*F_2*F_3} + ra\sigma^2_{F_2*F_3}$	$H_o : \sigma^2_{F_2*F_3} = 0$	$\frac{MS_{F_2*F_3}}{MS_{F_1*F_2*F_3}}$
$F_1 * F_2 * F_3$	$MS_{F_1*F_2*F_3}$	$\sigma_e^2 + r\sigma^2_{F_1*F_2*F_3}$	$H_o : \sigma^2_{F_1*F_2*F_3} = 0$	$\frac{MS_{F_1*F_2*F_3}}{MSE}$
Error	$MS_{F_1*F_2*F_3}$	σ_e^2		

Approximate F-tests are obtained by obtaining a linear combination of Mean Squares, M, such that the expected value of M equals the expected value of the MS for corresponding source of variation, e.g., $E[M] = E_{H:_o}[MS_{F_1}]$.

1. For F_1, $* = \frac{MS_{F_1}}{M}$ where $M = MS_{F_1*F_2} + MS_{F_1*F_3} - MS_{F_1*F_2*F_3}$

2. For F_2, $** = \frac{MS_{F_2}}{M}$ where $M = MS_{F_1*F_2} + MS_{F_2*F_3} - MS_{F_1*F_2*F_3}$

3. For F_3, $* * * = \frac{MS_{F_3}}{M}$ where $M = MS_{F_1*F_3} + MS_{F_2*F_3} - MS_{F_1*F_2*F_3}$

Model IV: F_1-Fixed, F_2-Fixed, and F_3-Random

Source of Variation	MS	Expected MS	Null Hypothesis	F-Statistic
F_1	MS_{F_1}	$\sigma_e^2 + r\sigma^2_{F_1*F_2*F_3} + rb\sigma^2_{F_1*F_3} + rbcQ_{F_1}$	$H_o : Q_{F_1} = 0$	$\frac{MS_{F_1}}{MS_{F_1*F_3}}$
F_2	MS_{F_2}	$\sigma_e^2 + r\sigma^2_{F_1*F_2*F_3} + ra\sigma^2_{F_2*F_3} + racQ_{F_2}$	$H_o : Q_{F_2} = 0$	$\frac{MS_{F_2}}{MS_{F_2*F_3}}$
F_3	MS_{F_3}	$\sigma_e^2 + r\sigma^2_{F_1*F_2*F_3} + rb\sigma^2_{F_1*F_3} + ra\sigma^2_{F_2*F_3} + rab\sigma^2_{F_3}$	$H_o : \sigma^2_{F_3} = 0$	$*$
$F_1 * F_2$	$MS_{F_1*F_2}$	$\sigma_e^2 + r\sigma^2_{F_1*F_2*F_3} + rcQ_{F_1*F_2}$	$H_o : Q_{F_1*F_2} = 0$	$\frac{MS_{F_1*F_2}}{MS_{F_1*F_2*F_3}}$
$F_1 * F_3$	$MS_{F_1*F_3}$	$\sigma_e^2 + r\sigma^2_{F_1*F_2*F_3} + rb\sigma^2_{F_1*F_3}$	$H_o : \sigma^2_{F_1*F_3} = 0$	$\frac{MS_{F_1*F_3}}{MS_{F_1*F_2*F_3}}$
$F_2 * F_3$	$MS_{F_2*F_3}$	$\sigma_e^2 + r\sigma^2_{F_1*F_2*F_3} + ra\sigma^2_{F_2*F_3}$	$H_o : \sigma^2_{F_2*F_3} = 0$	$\frac{MS_{F_2*F_3}}{MS_{F_1*F_2*F_3}}$
$F_1 * F_2 * F_3$	$MS_{F_1*F_2*F_3}$	$\sigma_e^2 + r\sigma^2_{F_1*F_2*F_3}$	$H_o : \sigma^2_{F_1*F_2*F_3} = 0$	$\frac{MS_{F_1*F_2*F_3}}{MSE}$
Error	$MS_{F_1*F_2*F_3}$	σ_e^2		

Approximate F-tests are obtained by obtaining a linear combination of Mean Squares, M, such that the expected value of M equals the expected value of the MS for corresponding source of variation, e.g., $E[M] = E_{H:_o}[MS_{F_1}]$.

For F_3, $* = \frac{MS_{F_3}}{M}$ where $M = MS_{F_1*F_3} + MS_{F_2*F_3} - MS_{F_1*F_2*F_3}$

- Next we will display an algorithm which will generate the EMS given in the AOV tables.

Expected Mean Square Rules when $n_{ijk} \equiv r$

The following rules for determining the expected mean squares apply to equally replicated designs under the set the last term equal to 0 restrictions for all fixed effects terms in the model. The rules will be illustrated for an experiment with

factors F_1 and F_2 having a and b fixed levels, respectively, and factor F_3 having c random levels.

1. Write out the linear model for the experiment:

   MODEL: $y_{ijkl} = \mu + \tau_i + \gamma_j + c_k + (\tau\gamma)_{ij} + (\tau c)_{ik} + (\gamma c)_{jk} + (\tau\gamma c)_{ijk} + e_{l(ijk)}$

   Note that the replication source of variation is nested within the ijk treatment combination.

2. Construct a two-way table with

   a. The first column containing an entry for each source of variation in the model, excluding μ, include the number of levels of each factor, and whether the factor is Fixed "F" or Random "R"

   b. A column for each random variance component, σ or fixed variance component Q

   c. Above each column write "R" if the component is a variance or "F" if the component is a fixed levels treatment difference

Factor F_1-Fixed, Factor F_2-Fixed, Factor F_3-Random

		F	F	R	F	R	R	R	R
SV	Levels	Q_{F_1}	Q_{F_2}	$\sigma^2_{F_3}$	$Q_{F_1*F_2}$	$\sigma^2_{F_1*F_3}$	$\sigma^2_{F_2*F_3}$	$\sigma^2_{F_1*F_2*F_3}$	$\sigma^2_{e(F_1,F_2,F_3)}$
F_1-F	a								
F_2-F	b								
F_3-R	c								
$F_1 * F_2$-F	ab								
$F_1 * F_3$-R	ac								
$F_2 * F_3$-R	bc								
$F_1 * F_2 * F_3$-R	abc								
Error(F_1, F_2, F_3)-R	$abcr$								

STOP Friday 4/1/22 (week 10, lecture 27)

START Monday 4/4/22 (Week 11, Lecture 28)

3. For each row, place the following values under each variance component:

   a. In the column for a **fixed** variance component place a 0 in all rows **except** for the row where the source of variation exactly matches the subscripts of the **fixed** variance component. For this row divide the number consisting of r times the number of levels of all factors, $rabc$ in this example, by the number of levels of the factors in the subscript of the variance component, then place the resulting number in the row of the SV that exactly matched the subscripts of the **fixed** variance component.

      E.g, For $F_1 * F_2$ where both F_1 and F_2 have fixed levels, divide $abcr$ by ab yielding the value cr. Then place cr in row for $F_1 * F_2$ under $Q_{F_1 * F_2}$

      E.g., For SV F_1, a 0 would be placed in the row for F_1 under $Q_{F_1 * F_2}$ but bcr would be placed in the row for F_1 under Q_{F_1}

   b. If the source of variation (SV) is not a part of the subscript of a variance component, then place a 0 in the row for the SV and column of that variance component.

      E.g., For SV F_1, a 0 would be placed under $\sigma_{F_2 * F_3}$ in the row for F_1

   c. If the source of variation is a part of or the complete subscript of a **random** variance component then divide the number consisting of r times the number of levels of all factors, $rabc$ in this example, by the number of levels of the factors in the subscript of the variance component, then place the resulting number in the row for that SV under the random variance component.

      E.g., For SV F_1 and random variance component $\sigma^2_{F_1 * F_3}$, divide $abcr$ by ac yielding the value br. Then place place br under $\sigma^2_{F_1 * F_3}$ in the row for F_1

   d. Place a 1 in each row of the column under $\sigma^2_e = \sigma^2_{e(F_1, F_2, F_3)}$ because all sources of variation are contained in the subscript and $abcr$ divided by $abcr$ is 1.

4. After completing all entries in the table, the expected mean square for each source of variation is obtained by multiplying the entries in the row corresponding to each source of variation by the corresponding variance component.

Factor F_1-Fixed, Factor F_2-Fixed, Factor F_3-Random

		F	F	R	F	R	R	R	R
SV	Levels	Q_{F_1}	Q_{F_2}	$\sigma^2_{F_3}$	$Q_{F_1*F_2}$	$\sigma^2_{F_1*F_3}$	$\sigma^2_{F_2*F_3}$	$\sigma^2_{F_1*F_2*F_3}$	σ^2_e
F_1	a	bcr	0	0	0	br	0	r	1
F_2	b	0	acr	0	0	0	ar	r	1
F_3	c	0	0	abr	0	br	ar	r	1
$F_1 * F_2$	ab	0	0	0	cr	0	0	r	1
$F_1 * F_3$	ac	0	0	0	0	br	0	r	1
$F_2 * F_3$	bc	0	0	0	0	0	ar	r	1
$F_1 * F_2 * F_3$	abc	0	0	0	0	0	0	r	1
Error	$abcr$	0	0	0	0	0	0	0	1

SV	Expected Mean Square
F_1	$bcrQ_{F_1} + br\sigma^2_{F_1*F_3} + r\sigma^2_{F_1*F_2*F_3} + \sigma^2_e$
F_2	$acrQ_{F_2} + ar\sigma^2_{F_2*F_3} + r\sigma^2_{F_1*F_2*F_3} + \sigma^2_e$
F_3	$abr\sigma^2_{F_3} + br\sigma^2_{F_1*F_3} + ar\sigma^2_{F_2*F_3} + r\sigma^2_{F_1*F_2*F_3} + \sigma^2_e$
$F_1 * F_2$	$crQ_{F_1*F_2} + r\sigma^2_{F_1*F_2*F_3} + \sigma^2_e$
$F_1 * F_3$	$br\sigma^2_{F_1*F_3} + r\sigma^2_{F_1*F_2*F_3} + \sigma^2_e$
$F_2 * F_3$	$ar\sigma^2_{F_2*F_3} + r\sigma^2_{F_1*F_2*F_3} + \sigma^2_e$
$F_1 * F_2 * F_3$	$r\sigma^2_{F_1*F_2*F_3} + \sigma^2_e$
Error	σ^2_e

Factor F_1-Fixed, Factor F_2-Random, Factor F_3-Random

SV	Levels	F Q_{F_1}	R $\sigma^2_{F_2}$	R $\sigma^2_{F_3}$	R $\sigma^2_{F_1F_2}$	R $\sigma^2_{F_1F_3}$	R $\sigma^2_{F_2F_3}$	R $\sigma^2_{F_1F_2F_3}$	R σ^2_e
F_1 (F)	a	bcr	0	0	cr	br	0	r	1
F_2 (R)	b	0	acr	0	cr	0	ar	r	1
F_3 (R)	c	0	0	abr	0	br	ar	r	1
$F_1 * F_2$ (R)	ab	0	0	0	cr	0	0	r	1
$F_1 * F_3$ (R)	ac	0	0	0	0	br	0	r	1
$F_2 * F_3$ (R)	bc	0	0	0	0	0	ar	r	1
$F_1 * F_2 * F_3$ (R)	abc	0	0	0	0	0	0	r	1
Error	abcr	0	0	0	0	0	0		1

SV	Expected Mean Square
F_1	
F_2	
F_3	
$F_1 * F_2$	
$F_1 * F_3$	
$F_2 * F_3$	
$F_1 * F_2 * F_3$	
Error	

Factor F_1-Fixed, Factor F_2-Random, Factor F_3-Random

SV	Levels	F Q_{F_1}	R $\sigma^2_{F_2}$	R $\sigma^2_{F_3}$	R $\sigma^2_{F_1*F_2}$	R $\sigma^2_{F_1*F_3}$	R $\sigma^2_{F_2*F_3}$	R $\sigma^2_{F_1*F_2*F_3}$	R σ^2_e
F_1	a	bcr	0	0	cr	br	0	r	1
F_2	b	0	acr	0	cr	0	ar	r	1
F_3	c	0	0	abr	0	br	ar	r	1
$F_1 * F_2$	ab	0	0	0	cr	0	0	r	1
$F_1 * F_3$	ac	0	0	0	0	br	0	r	1
$F_2 * F_3$	bc	0	0	0	0	0	ar	r	1
$F_1 * F_2 * F_3$	abc	0	0	0	0	0	0	r	1
Error	$abcr$	0	0	0	0	0	0	0	1

SV	Expected Mean Square
F_1	$bcrQ_{F_1} + cr\sigma^2_{F_1*F_2} + br\sigma^2_{F_1*F_3} + r\sigma^2_{F_1*F_2*F_3} + \sigma^2_e$
F_2	$acr\sigma^2_{F_2} + cr\sigma^2_{F_1*F_2} + ar\sigma^2_{F_2*F_3} + r\sigma^2_{F_1*F_2*F_3} + \sigma^2_e$
F_3	$abr\sigma^2_{F_3} + br\sigma^2_{F_1*F_3} + ar\sigma^2_{F_2*F_3} + r\sigma^2_{F_1*F_2*F_3} + \sigma^2_e$
$F_1 * F_2$	$cr\sigma^2_{F_1*F_2} + r\sigma^2_{F_1*F_2*F_3} + \sigma^2_e$
$F_1 * F_3$	$br\sigma^2_{F_1*F_3} + r\sigma^2_{F_1*F_2*F_3} + \sigma^2_e$
$F_2 * F_3$	$ar\sigma^2_{F_2*F_3} + r\sigma^2_{F_1*F_2*F_3} + \sigma^2_e$
$F_1 * F_2 * F_3$	$r\sigma^2_{F_1*F_2*F_3} + \sigma^2_e$
Error	σ^2_e

Multiple Comparison and Contrasts in the Mixed Model

I. Two Factors F_1 and F_2

Case 1: Both Factors have Fixed Levels

Model: $y_{ijk} = \mu + \tau_i + \gamma_j + (\tau\gamma)_{ij} + e_{ijk}$ for $i = 1, \ldots, a \quad j = 1, \ldots, b \quad k = 1, \ldots, r$

Case 1a: $F_1 * F_2$ not significant

Suppose there is not significant evidence of an interaction between F_1 and F_2, then comparisons in the marginal means would be of interest: $\bar{\mu}_{i.}, \quad i = 1, \ldots, a$ and $\bar{\mu}_{.j}, \quad j = 1, \ldots, b$. This will require modifications to our formulas for testing contrasts and multiple comparisons. Suppose we want to compare means across the levels of factor F_1 : $\bar{\mu}_{i.}$'s with $\hat{\bar{\mu}}_{i.} = \bar{y}_{i..}$

$$Var(\bar{y}_{i..}) = Var(\bar{e}_{i..}) = \frac{\overset{=MSE}{\sigma_e^2}}{br} \quad \Rightarrow \widehat{SE}(\bar{y}_{i..}) = \sqrt{\frac{MSE}{rb}}$$

$$Var(\bar{y}_{i..} - \bar{y}_{h..}) = Var(\bar{e}_{i..} - \bar{e}_{h..}) = 2Var(\bar{e}_{i..}) = \frac{2\sigma_e^2}{br} \quad \Rightarrow \widehat{SE}(\bar{y}_{i..} - \bar{y}_{h..}) = \sqrt{\frac{2MSE}{rb}}$$

Use the above in the formulas for Multiple Comparisons, Dunnett's Procedure, Hsu's Procedure, and testing contrasts where we are evaluating differences in levels of Factor F_1 averaged over the levels of factor F_2: $\bar{\mu}_{i.}, \quad i = 1, \ldots, a$

1. The F Test: State there is significant evidence that the contrast

   $C = \sum_{i=1}^{a} c_i \bar{\mu}_{i.}$ is not 0 if $\widehat{C} = \sum_{i=1}^{a} c_i \bar{y}_{i..}$ satisfies:

   $$|\widehat{C}| \geq \sqrt{MSE} \quad \sqrt{\sum_{i=1}^{a} \frac{c_i^2}{rb}} \sqrt{F_{\alpha,1,df_{MSE}}}$$

2. The Tukey-Kramer's HSD: State there is significant evidence that the marginal means $\bar{\mu}_{i.}$ and $\bar{\mu}_{h.}$ are different if

   $$|\bar{y}_{i..} - \bar{y}_{h..}| \geq q_{\alpha,a,df_{MSE}} \quad \sqrt{\frac{1}{2}\left(\widehat{SE}(\bar{y}_{i..} - \bar{y}_{h..})\right)^2} = q_{\alpha,a,df_{MSE}} \quad \sqrt{\frac{MSE}{rb}}$$

3. The Dunnett's HSD: State there is significant evidence that the control mean $\bar{\mu}_{c.}$ is less than $\bar{\mu}_{i.}$ if

   $$(\bar{y}_{i..} - \bar{y}_{c..}) \geq d_{\alpha,a-1,df_{MSE}} \quad \sqrt{\frac{2MSE}{rb}}$$

Similar modifications must be made for the other multiple comparison procedures on the marginal means.

Case 1b: $F_1 * F_2$ significant

If there is a significant interaction, then the comparisons are across the treatment means μ_{ij} globally or across the levels of $\mu_{ij}, \quad i = 1, \ldots, a$ for fixed levels of $j = 1, \ldots, b$. In this case, the formulas will be modified in a similar manner

$$Var(\bar{y}_{ij.}) = Var(\bar{e}_{ij.}) = \frac{\sigma_e^2}{r} \quad \Rightarrow \widehat{SE}(\bar{y}_{ij.}) = \sqrt{\frac{MSE}{r}}$$

$$Var(\bar{y}_{ij.} - \bar{y}_{hk.}) = Var(\bar{e}_{ij.} - \bar{e}_{hk.}) = 2Var(\bar{e}_{ij.}) = \frac{2\sigma_e^2}{r} \quad \Rightarrow \widehat{SE}(\bar{y}_{ij.} - \bar{y}_{hk.}) = \sqrt{\frac{2MSE}{r}}$$

with the number of treatments being $t = ab$.

For example, suppose we want to compare the levels of factor F_1 separately at each level of factor F_2: For each value of $j = 1, \ldots, b$, compare the a levels of Factor F_1:

1. The Bonferroni - F Test: For each value of $j = 1, 2, \ldots, b$; State there is significant evidence that the contrast

   $C = \sum_{i=1}^{a} c_i \mu_{ij}$ is not 0 if $\widehat{C} = \sum_{i=1}^{a} c_i \bar{y}_{ij.}$ satisfies :

   $$|\widehat{C}| \geq \sqrt{MSE} \quad \sqrt{\sum_{i=1}^{a} \frac{c_i^2}{r}} \quad \sqrt{F_{\frac{\alpha}{b},1,df_{MSE}}}$$

2. The Bonferroni - Tukey's HSD: For each value of $j = 1, 2, \ldots, b$; State there is significant evidence that the marginal means μ_{ij} and μ_{hj} are different if

   $$|\bar{y}_{ij.} - \bar{y}_{hj.}| \geq q_{\frac{\alpha}{b},a,df_{MSE}} \quad \sqrt{\frac{MSE}{r}}$$

   Alternatively, compare the adjusted Tukey p-values from the LSMEANS statement in the SAS output to $\alpha = .05$

   or Compare the unadjusted p-values from the LSMEANS statement in the SAS output to $\alpha_{PC} = .05/M$ where $M = ba(a-1)/2$

3. The Bonferroni - Dunnett's HSD: For each value of $j = 1, 2, \ldots, b$; State there is significant evidence that the control mean $\mu_{c.}$ is less than μ_{ij} if

   $$(\bar{y}_{ij.} - \bar{y}_{c..}) \geq d_{\frac{\alpha}{b},a-1,df_{MSE}} \quad \sqrt{\frac{2MSE}{r}}$$

Similar modifications must be made for the other multiple comparison procedures on the marginal means.

- assumptions $\tau_i = 0$ $b_j \sim N(0, \sigma_b^2)$

Case 2: F_1 - Fixed Levels and F_2 - Random Levels

Case 2a: $F_1 * F_2$ not significant

If there is not significant evidence of an interaction between F_1 and F_2, then comparisons of differences in the marginal treatment means $\bar{\mu}_{i.}$ across the levels of F_1. Comparisons across the levels of F_2 would be inappropriate.

$y_{ijk} = \mu + \tau_i + b_j + (\tau b)_{ij} + e_{ijk};$ for $i = 1, \ldots, a \quad j = 1, \ldots, b \quad k = 1, \ldots, r$

$$\begin{aligned} Var(\bar{y}_{i..} - \bar{y}_{h..}) &= Var((\bar{b}_. + (\bar{\tau b})_{i.} + \bar{e}_{i..}) - (\bar{b}_. + (\bar{\tau b})_{h.} + \bar{e}_{h..})) \\ &= 2Var((\bar{\tau b})_{i.}) + 2Var(\bar{e}_{i..}) \\ &= \frac{2\sigma^2_{F_1*F_2}}{b} + \frac{2\sigma^2_e}{rb} = \frac{2(r\sigma^2_{F_1*F_2} + \sigma^2_e)}{rb} = \frac{2(EMS_{F_1*F_2})}{rb} \quad \Rightarrow \end{aligned}$$

Using the Method of Moments AOV estimators,
$\widehat{SE}(\bar{y}_{i..} - \bar{y}_{h..}) = \sqrt{\frac{2(MS_{F_1*F_2})}{rb}}$, with df=(a-1)(b-1)

Use the above in the formulas for Multiple Comparisons, Dunnett's Procedure, Hsu's Procedure, and testing contrasts.

1. The F Test: State there is significant evidence that the contrast

   $C = \sum_{i=1}^a c_i\bar{\mu}_{i.}$ is not 0 if $\widehat{C} = \sum_{i=1}^a c_i\bar{y}_{i..}$ satisfies :

   $|\widehat{C}| \geq \sqrt{MS_{F_1*F_2}} \sqrt{\sum_{i=1}^a \frac{c_i^2}{rb}} \sqrt{F_{\alpha,1,\nu}}$ where $\nu = df_{MS_{F_1*F_2}} = (a-1)(b-1)$

2. The Tukey-Kramer's HSD: State there is significant evidence that the marginal means $\bar{\mu}_{i.}$ and $\bar{\mu}_{h.}$ are different if

   $|\hat{\bar{\mu}}_{i.} - \hat{\bar{\mu}}_{h.}| = |\bar{y}_{i..} - \bar{y}_{h..}| \geq q_{\alpha,a,\nu} \sqrt{MS_{F_1*F_2}} \sqrt{\frac{1}{rb}}$ where $\nu = df_{MS_{F_1*F_2}} = (a-1)(b-1)$

3. The Dunnett's HSD: State there is significant evidence that the control mean $\bar{\mu}_{c.}$ is less than $\bar{\mu}_{i.}$ if

   $(\hat{\bar{\mu}}_{i.} - \hat{\bar{\mu}}_{c.}) = (\bar{y}_{i..} - \bar{y}_{c..}) \geq d_{\alpha,a-1,\nu} \sqrt{\frac{2MS_{F_1*F_2}}{rb}}$ where $\nu = df_{MS_{F_1*F_2}} = (a-1)(b-1)$

Similar modifications must be made for the other multiple comparison procedures on the marginal means.

Case 2b: $F_1 * F_2$ significant

If there is a significant interaction, then the same results that were obtained in Case 2a would be appropriate because the levels of F_2 are random.

Case 3: F_1 - Random Levels and F_2 - Random Levels

Multiple comparisons and contrasts in either the treatment means or the marginal means are not of interest due to the random nature of the levels of both factors.

II. Three Factors F_1, F_2 and F_3

Case 1: All Factors have Random Levels

Multiple comparisons and contrasts are not appropriate due to the random nature of the levels of all three factors.

Case 2: All Factors have Fixed Levels

Straight forward modifications to the multiple comparison procedures and tests of contrasts with

A. MSE used as the estimator of σ^2 and $\nu = df_{MSE}$ used in looking up critical values in tables

B. In the formulas, replace t with the appropriate number of means being compared

$t = abc$ for comparison of μ_{ijk}

$t = ab$ for comparisons of $\bar{\mu}_{ij.}$

$t = a$ for comparisons of $\bar{\mu}_{i..}$;

C. In the formulas, replace r or n_i with the appropriate sample size which is determined by the number of terms averaged over to obtain the point estimator. For example,

$n_i = r$ for comparisons of $\mu_{ijk} \ \Rightarrow \hat{\mu}_{ijk} = \bar{y}_{ijk.}$

$n_i = rc$ for comparisons of $\mu_{ij.} \ \Rightarrow \hat{\mu}_{ij.} = \bar{y}_{ij..}$

$n_i = rbc$ for comparisons of $\mu_{i..} \ \Rightarrow \hat{\mu}_{i..} = \bar{y}_{i...}$

Situation 1: When Three-way Interaction $F_1 * F_2 * F_3$ Is Significant

For each combination $(j,\ k)$ of the levels of $(F_2,\ F_3)$, conduct comparisons across the levels of F_1 using $\hat{\mu}_{ijk}$ in Tukey HSD or using Contrasts in μ_{ijk}.

Situation 2: When Three-way Interaction $F_1 * F_2 * F_3$ Is Not Significant

i. **Both $F_1 * F_2$ and $F_1 * F_3$ are Significant**

a. For each level of j of F_2 conduct comparisons of the levels of F_1 using $\hat{\mu}_{ij.}$ in Tukey HSD or using Contrasts in $\mu_{ij.}$

b. For each level of k of F_3 conduct comparisons of the levels of F_1 using $\hat{\mu}_{i.k}$ in Tukey HSD or using Contrasts in $\mu_{i.k}$

ii. **If $F_1 * F_2$ is Significant but $F_1 * F_3$ is Not Significant**

For each level of j of F_2 conduct comparisons of the levels of F_1 using $\hat{\mu}_{ij.}$ in Tukey HSD or using Contrasts in $\mu_{ij.}$

iii. **If Both $F_1 * F_2$ and $F_1 * F_3$ are Not Significant**

Conduct comparisons of the levels of F_1 using $\hat{\mu}_{i..}$ in Tukey HSD or using Contrasts in $\mu_{i..}$

Case 3: F_1 - Fixed Levels, F_2 - Random Levels and F_3 - Random Levels

Because the levels of F_2 and F_3 are random, only comparisons of differences in the levels of F_1 are of interest, that is, differences in the marginal treatment means $\bar{\mu}_{i..}$ are of interest:

$$y_{ijkl} = \mu + \tau_i + b_j + c_k + (\tau b)_{ij} + (\tau c)_{ik} + (bc)_{jk} + (\tau bc)_{ijk} + e_{ijkl} \quad \Rightarrow$$

$$\begin{aligned}
Var(\bar{y}_{i...}) &= Var[(\bar{b}_. + \bar{c}_. + (\bar{\tau b})_{i.} + (\bar{\tau c})_{i.} + (\bar{bc})_{..} + (\bar{\tau bc})_{i..} + \bar{e}_{i...}) \\
&= Var(\bar{b}_.) + Var(\bar{c}_.) + Var((\bar{\tau b})_{i.}) + Var((\bar{\tau c})_{i.}) + Var((\bar{bc})_{..}) + Var((\bar{\tau bc})_{i..}) + Var(\bar{e}_{i...}) \\
&= \\
&= \frac{\sigma^2_{F_2}}{b} + \frac{\sigma^2_{F_3}}{c} + \frac{\sigma^2_{F_1*F_2}}{b} + \frac{\sigma^2_{F_1*F_3}}{c} + \frac{\sigma^2_{F_2*F_3}}{bc} + \frac{\sigma^2_{F_1*F_2*F_3}}{bc} + \frac{\sigma^2_e}{rbc} \\
&= \\
&= \frac{cr\sigma^2_{F_2} + br\sigma^2_{F_3} + cr\sigma^2_{F_1*F_2} + br\sigma^2_{F_1*F_3} + r\sigma^2_{F_2*F_3} + r\sigma^2_{F_1*F_2*F_3} + \sigma^2_e}{bcr} \\
&= \left(\frac{1}{abcr}\right)[EMS_{F_2} + EMS_{F_3} + (a-1)EMS_{F_1*F_2} + (a-1)EMS_{F_1*F_3} - EMS_{F_2*F_3} \\
&\quad +(1-a)EMS_{F_1*F_2*F_3}] \quad \Rightarrow
\end{aligned}$$

Using the Method of Moments AOV estimators,

$$\widehat{SE}(\bar{y}_{i...}) = \sqrt{\left(\frac{1}{abcr}\right)[MS_{F_2} + MS_{F_3} + (a-1)MS_{F_1*F_2} + (a-1)MS_{F_1*F_3} - MS_{F_2*F_3} + (1-a)MS_{F_1*F_2*F_3}]}$$

Similarly, we compute:

$$\begin{aligned}
Var(\bar{y}_{i...} - \bar{y}_{h...}) &= Var[(\bar{b}_{.} + \bar{c}_{.} + \overline{(\tau b)}_{i.} + \overline{(\tau c)}_{i.} + \overline{(bc)}_{..} + \overline{(\tau bc)}_{i..} + \bar{e}_{i...}) \\
&\qquad -(\bar{b}_{.} + \bar{c}_{.} + \overline{(\tau b)}_{h.} + \overline{(\tau c)}_{h.} + \overline{(bc)}_{..} + \overline{(\tau bc)}_{h..} + \bar{e}_{h...})] \\
&= 2Var(\overline{(\tau b)}_{i.}) + 2Var(\overline{(\tau c)}_{i.}) + 2Var(\overline{(\tau bc)}_{i..}) + 2Var(\bar{e}_{i...}) \\
&= \frac{2(cr\sigma^2_{F_1*F_2} + br\sigma^2_{F_1*F_3} + r\sigma^2_{F_1*F_2*F_3} + \sigma^2_e)}{bcr} \\
&= \left(\frac{2}{bcr}\right)(EMS_{F_1*F_2} + EMS_{F_1*F_3} - EMS_{F_1*F_2*F_3}) \quad \Rightarrow
\end{aligned}$$

Using the Method of Moments AOV estimators,

$$\widehat{SE}(\bar{y}_{i...} - \bar{y}_{h...}) = \sqrt{\left(\frac{2}{bcr}\right)(MS_{F_1*F_2} + MS_{F_1*F_3} - MS_{F_1*F_2*F_3})}$$

The *df* associated with the estimator of the difference in marginal means is given by the Satterthwaite approximation:

$$df = \frac{\left(MS_{F_1*F_2} + MS_{F_1*F_3} - MS_{F_1*F_2*F_3}\right)^2}{\frac{(MS_{F_1*F_2})^2}{df_{MS_{F_1*F_2}}} + \frac{(MS_{F_1*F_3})^2}{df_{MS_{F_1*F_3}}} + \frac{(MS_{F_1*F_2*F_3})^2}{df_{MS_{F_1*F_2*F_3}}}}$$

Use the above in the formulas for Multiple Comparisons, Dunnett's Procedure, Hsu's Procedure, and testing contrasts.

1. The F Test: State there is significant evidence that the contrast

   $C = \sum_{i=1}^{a} c_i \bar{\mu}_{i..}$ is not 0 if if $\widehat{C} = \sum_{i=1}^{a} c_i \bar{y}_{i...}$ satisfies :

   $|\widehat{C}| \geq \sqrt{M} \quad \sqrt{\sum_{i=1}^{a} \frac{c_i^2}{rbc}} \quad \sqrt{F_{\alpha,1,\nu}}$

   $M = MS_{F_2} + MS_{F_3} + (a-1)MS_{F_1*F_2} + (a-1)MS_{F_1*F_3} - MS_{F_2*F_3} + (1-a)EMS_{F_1*F_2*F_3}$

   ν is computed from the Satterthwaite Approximation.

2. The Tukey-Kramer's HSD: State there is significant evidence that the marginal means $\bar{\mu}_{i.}$ and $\bar{\mu}_{h.}$ are different if

   $|\bar{y}_{i..} - \bar{y}_{h..}| \geq q_{\alpha,a,\nu} \quad \sqrt{MS_{F_1*F_2} + MS_{F_1*F_3} - MS_{F_1*F_2*F_3}} \quad \sqrt{\frac{1}{rbc}}$

   ν is given above from the Satterthwaite Approximation.

3. The Dunnett's HSD: State there is significant evidence that the control mean $\bar{\mu}_{c.}$ is less than $\bar{\mu}_{i.}$ if

$$(\bar{y}_{i..} - \bar{y}_{c..}) \geq d_{\alpha,a-1,\nu} \sqrt{\left(\frac{2}{rbc}\right)(MS_{F_1*F_2} + MS_{F_1*F_3} - MS_{F_1*F_2*F_3})}$$

ν is given above from the Satterthwaite Approximation.

Similar modifications must be made for the other multiple comparison procedures on the marginal means.

Case 4: F_1 - Fixed Levels, F_2 - Fixed Levels and F_3 - Random Levels

Case 4a : $F_1 * F_2$ Not Significant

If there is not significant evidence of a two-way interaction $F_1 * F_2$, then comparisons of differences in the marginal treatment means $\bar{\mu}_{i..}$ across the levels of F_1 or the marginal treatment means $\bar{\mu}_{.j.}$ across the levels of F_2 are of interest:

$$y_{ijkl} = \mu + \tau_i + \gamma_j + c_k + (\tau\gamma)_{ij} + (\tau c)_{ik} + (\gamma c)_{jk} + (\tau\gamma c)_{ijk} + e_{ijkl}$$

1. Comparing the F_1 marginal treatment means $\bar{\mu}_{1..}, \ldots, \bar{\mu}_{a..}$

$$\begin{aligned}
Var(\bar{y}_{i...} - \bar{y}_{h...}) &= Var[(\bar{c}_{.} + \overline{(\tau c)}_{i.} + \overline{(\gamma c)}_{..} + \overline{(\tau\gamma c)}_{i..} + \bar{e}_{i...}) \\
&\qquad -(\bar{c}_{.} + \overline{(\tau c)}_{h.} + \overline{(\gamma c)}_{..} + \overline{(\tau\gamma c)}_{h..} + \bar{e}_{h...})] \\
&= 2Var(\overline{(\tau c)}_{i.}) + 2Var(\overline{(\tau\gamma c)}_{i..}) + 2Var(\bar{e}_{i...}) \\
&= \frac{2\sigma^2_{F_1*F_3}}{c} + \frac{2\sigma^2_{F_1*F_2*F_3}}{bc} + \frac{2\sigma^2_e}{bcr} \\
&= \frac{2(rb\sigma^2_{F_1*F_3} + r\sigma^2_{F_1*F_2*F_3} + \sigma^2_e)}{bcr} \\
&= \left(\frac{2}{bcr}\right)(EMS_{F_1*F_3}) \quad \Rightarrow
\end{aligned}$$

Using the Method of Moments AOV estimators,

$$\widehat{SE}(\bar{y}_{i...} - \bar{y}_{h...}) = \sqrt{\left(\frac{2}{bcr}\right)(MS_{F_1*F_3})}$$

with $df = df_{MS_{F_1*F_3}} = (a-1)(c-1)$

Use the above in the formulas for Multiple Comparisons, Dunnett's Procedure, Hsu's Procedure, and testing contrasts.

2. A similar argument yields:

Comparing the F_2 marginal treatment means $\bar{\mu}_{.1.}, \ldots, \bar{\mu}_{.b.}$

$$Var(\bar{y}_{.j..} - \bar{y}_{.k..}) = \left(\frac{2}{acr}\right)(EMS_{F_2*F_3}) \quad \Rightarrow$$

Using the Method of Moments AOV estimators,

$$\widehat{SE}(\bar{y}_{.j..} - \bar{y}_{.k..}) = \sqrt{\left(\frac{2}{acr}\right)(MS_{F_2*F_3})}$$

with $df = df_{MS_{F_2*F_3}} = (b-1)(c-1)$

Case 4b : $F_1 * F_2$ is Significant

If there is significant evidence of a two-way interaction $F_1 * F_2$, then comparisons of differences in the interaction means $\bar{\mu}_{ij.}$ may be of interest.

1. Comparing the treatment means $\bar{\mu}_{ij.}$ across the levels of F_1 separately at each level of F_2:

$$\bar{\mu}_{1j.}, \ldots, \bar{\mu}_{aj.} \text{ for } j = 1, \ldots, b$$

$$y_{ijkl} = \mu + \tau_i + \gamma_j + c_k + (\tau\gamma)_{ij} + (\tau c)_{ik} + (\gamma c)_{ij} + (\tau\gamma c)_{ijk} + e_{ijkl}$$

$$\begin{aligned}
Var(\bar{y}_{ij..} - \bar{y}_{hj..}) &= Var[(\bar{c}_. + \bar{(\tau c)}_{i.} + \bar{(\gamma c)}_{j.} + \bar{(\tau\gamma c)}_{ij.} + \bar{e}_{ij..}) \\
&\quad\; -(\bar{c}_. + \bar{(\tau c)}_{h.} + \bar{(\gamma c)}_{j.} + \bar{(\tau\gamma c)}_{hj.} + \bar{e}_{hj..})] \\
&= 2Var(\bar{(\tau c)}_{i.}) + 2Var(\bar{(\tau\gamma c)}_{ij.}) + 2Var(\bar{e}_{ij..}) \\
&= \frac{2\sigma^2_{F_1*F_3}}{c} + \frac{2\sigma^2_{F_1*F_2*F_3}}{c} + \frac{2\sigma^2_e}{rc} \\
&= \frac{2(r\sigma^2_{F_1*F_3} + r\sigma^2_{F_1*F_2*F_3} + \sigma^2_e)}{rc} \\
&= \frac{2\left[\frac{1}{b}(EMS_{F_1*F_3} - EMS_{F_1*F_2*F_3}) + EMS_{F_1*F_2*F_3}\right]}{rc} \\
&= \left(\frac{2}{brc}\right)(EMS_{F_1*F_3} + (b-1)EMS_{F_1*F_2*F_3})
\end{aligned}$$

Using the Method of Moments AOV estimators,

$$\widehat{SE}(\bar{y}_{ij..} - \bar{y}_{hj..}) = \sqrt{\left(\frac{2}{brc}\right)(MS_{F_1*F_3} + (b-1)MS_{F_1*F_2*F_3})}$$

The df associated with the estimator of the difference in marginal means is given by the Satterthwaite approximation:

$$df = \frac{\left(MS_{F_1*F_3} + (b-1)MS_{F_1*F_2*F_3}\right)^2}{\frac{(MS_{F_1*F_3})^2}{df_{MS_{F_1*F_3}}} + \frac{((b-1)MS_{F_1*F_2*F_3})^2}{df_{MS_{F_1*F_2*F_3}}}}$$

2. Comparing the treatment means $\bar{\mu}_{ij.}$ across the levels of F_2 separately at each level of F_1:

$$\bar{\mu}_{i1.}, \ldots, \bar{\mu}_{ib.} \text{ for } i = 1, \ldots, a$$

$$Var(\bar{y}_{ij..} - \bar{y}_{ik..}) = \left(\frac{2}{acr}\right)\left(EMS_{F_2*F_3} + (a-1)EMS_{F_1*F_2*F_3}\right) \quad \Rightarrow$$

Using the Method of Moments AOV estimators,

$$\widehat{SE}(\bar{y}_{ij..} - \bar{y}_{ik..}) = \sqrt{\left(\frac{2}{acr}\right)\left(MS_{F_2*F_3} + (a-1)MS_{F_1*F_2*F_3}\right)}$$

The df associated with the estimator of the difference in marginal means is given by the Satterthwaite approximation:

$$df = \frac{\left(MS_{F_2*F_3} + (a-1)MS_{F_1*F_2*F_3}\right)^2}{\frac{(MS_{F_2*F_3})^2}{df_{MS_{F_2*F_3}}} + \frac{((a-1)MS_{F_1*F_2*F_3})^2}{df_{MS_{F_1*F_2*F_3}}}}$$

3. Comparing all $ab(ab-1)/2$ pairs of treatment means $\bar{\mu}_{ij.}$ versus $\bar{\mu}_{i'j'.}$:

$$\begin{aligned}
Var(\bar{y}_{ij..} - \bar{y}_{hk..}) &= Var[(\bar{c}_{.} + (\overline{\tau c})_{i.} + (\overline{\gamma c})_{j.} + (\overline{\tau\gamma c})_{ij.} + \bar{e}_{ij..}) \\
&\qquad -(\bar{c}_{.} + (\overline{\tau c})_{h.} + (\overline{\gamma c})_{k.} + (\overline{\tau\gamma c})_{hk.} + \bar{e}_{hk..})] \\
&= 2Var((\overline{\tau c})_{i.}) + 2Var((\overline{\gamma c})_{j.}) + 2Var((\overline{\tau\gamma c})_{ij.}) + 2Var(\bar{e}_{ij..}) \\
&= \frac{2\sigma^2_{F_1*F_3}}{c} + \frac{2\sigma^2_{F_2*F_3}}{c} + \frac{2\sigma^2_{F_1*F_2*F_3}}{c} + \frac{2\sigma^2_e}{rc} \\
&= \frac{2(r\sigma^2_{F_1*F_3} + r\sigma^2_{F_2*F_3} + r\sigma^2_{F_1*F_2*F_3} + \sigma^2_e)}{rc} \\
&= \left(\frac{2}{abrc}\right)\left(a\ EMS_{F_1*F_3} + b\ EMS_{F_2*F_3} + (ab-b-a)\ EMS_{F_1*F_2*F_3}\right) \quad \Rightarrow
\end{aligned}$$

Using the Method of Moments AOV estimators,

$$\widehat{SE}(\bar{y}_{ij..} - \bar{y}_{hk..}) = \sqrt{\left(\frac{2}{abcr}\right)(a\ MS_{F_1*F_3} + b\ MS_{F_2*F_3} + (ab - b - a)\ MS_{F_1*F_2*F_3})}$$

The *df* associated with the estimator of the difference in marginal means is given by the Satterthwaite approximation:

$$df = \frac{(a\ MS_{F_1*F_3} + b\ MS_{F_2*F_3} + (ab - b - a)\ MS_{F_1*F_2*F_3})^2}{\frac{(a\ MS_{F_1*F_3})^2}{df_{MS_{F_1*F_3}}} + \frac{(b\ MS_{F_2*F_3})^2}{df_{MS_{F_2*F_3}}} + \frac{((ab-b-a)\ MS_{F_1*F_2*F_3})^2}{df_{MS_{F_1*F_2*F_3}}}}$$

Estimation of Variance Components

I. Using AOV-Method of Moments

The AOV-MOM estimation procedure simply equates observed Mean Squares with Expected Values of the Mean Squares and solves for all relevant variance components. These estimators have the limitation that was discussed previously in Handout # 6.

For example, suppose we have a mixed model with factors F_1 Fixed and F_2 random. We then have three variance components to estimate: $\sigma^2_{F_2}$, $\sigma^2_{F_1*F_2}$, and σ^2_e. From the AOV table we have the expected Mean Squares as follows:

F_1-Fixed, F_2-Random

Source of Variation	MS	Expected MS
F_1	MS_{F_1}	$\sigma^2_e + r\sigma^2_{F_1*F_2} + rbQ_{F_1}$
F_2	MS_{F_2}	$\sigma^2_e + r\sigma^2_{F_1*F_2} + ra\sigma^2_{F_2}$
$F_1 * F_2$	$MS_{F_1*F_2}$	$\sigma^2_e + r\sigma^2_{F_1*F_2}$
Error	MSE	σ^2_e

Equating Expected MS to MS from the above table, the estimators are obtained as follows:

$$MSE = \hat{\sigma}^2_e \quad \Rightarrow \quad \hat{\sigma}^2_e = MSE$$

$$MS_{F_1*F_2} = \hat{\sigma}^2_e + r\hat{\sigma}^2_{F_1*F_2} \quad \Rightarrow \quad \hat{\sigma}^2_{F_1*F_2} = \frac{MS_{F_1*F_2} - \hat{\sigma}^2_e}{r} = \frac{MS_{F_1*F_2} - MSE}{r}$$

$$MS_{F_2} = \hat{\sigma}^2_e + r\hat{\sigma}^2_{F_1*F_2} + ra\hat{\sigma}^2_{F_2} \quad \Rightarrow \quad \hat{\sigma}^2_{F_2} = \frac{MS_{F_2} - (\hat{\sigma}^2_e + r\hat{\sigma}^2_{F_1*F_2})}{ra} = \frac{MS_{F_2} - MS_{F_1*F_2}}{ra}$$

To proportionally allocate the variability in the responses σ^2_y, first compute $\sigma^2_y = Var(y_{ijk}) = \sigma^2_{F_1*F_2} + \sigma^2_{F_2} + \sigma^2_e$, then estimate the proportional allocation as follows:

Variance Component	Proportion of Total
F_2	$\frac{\hat{\sigma}^2_{F_2}}{\hat{\sigma}^2_y}$
$F_1 * F_2$	$\frac{\hat{\sigma}^2_{F_1*F_2}}{\hat{\sigma}^2_y}$
$Error$	$\frac{\hat{\sigma}^2_e}{\hat{\sigma}^2_y}$

STOP Monday 4/4/22 (Week 11, Lecture 28)

START Friday 4/8/22 (week 11, lecture 29)

II. Likelihood Based Estimators

The mixed model is written as

fixed random

$$\mathbf{y} = \underbrace{\mathbf{X}}_{\text{fixed}}\boldsymbol{\beta} + \underbrace{\mathbf{Z}}_{\text{random}}\mathbf{u} + \mathbf{e}$$

where $\mathbf{y}$ denotes the vector of observed responses, $\mathbf{X}$ is the known Design matrix related to the fixed factor effects, $\boldsymbol{\beta}$ is vector of fixed-effects parameters, $\mathbf{Z}$ is the known Design matrix related to the random effects and $\mathbf{u}$ is a vector of random-effects. $\mathbf{X}$ contains indicator variables constructed from the fixed effects and $\mathbf{Z}$ contains indicator variables constructed from random effects. Finally, $\mathbf{e}$ is the unobserved vector of independent and identically distributed Gaussian random errors.

Assume that $\mathbf{u}$ and $\mathbf{e}$ are independent Gaussian random variables that are uncorrelated and have expectations 0 and variance-covariance matrices $\mathbf{G}$ and $\mathbf{R}$, respectively:

$$E\begin{bmatrix} \mathbf{u} \\ \mathbf{e} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix} \qquad Var\begin{bmatrix} \mathbf{u} \\ \mathbf{e} \end{bmatrix} = \begin{bmatrix} \mathbf{G} & \mathbf{0} \\ \mathbf{0} & \mathbf{R} \end{bmatrix}$$

The above conditions yield that $\mathbf{y}$ has a Gaussian distribution with

$$E[\mathbf{y}] = \mathbf{X}\boldsymbol{\beta} \qquad Var[\mathbf{y}] = \mathbf{V} = \mathbf{Z}\mathbf{G}\mathbf{Z}' + \mathbf{R}$$

Z is the random effects design matrix, G is the random effects var/cov matrix, so if we have all fixed effects $ZGZ' = 0$.

where $\mathbf{G}$ contains the variance components and $\mathbf{R} = \sigma_e^2 \mathbf{I}_n$

In the mixed model it is necessary to estimate not only $\boldsymbol{\beta}$ but also the unknown parameters in $\mathbf{G}$ and $\mathbf{R}$. Because the responses y are no longer independent, that is, $\mathbf{V}$ is no longer a diagonal matrix with σ_e^2 along the diagonal, Least Squares estimation is no longer the best method of estimation. For the estimation of $\boldsymbol{\beta}$, Generalized Least Squares (GLS) is the more appropriate method of estimation, that is, find $\hat{\boldsymbol{\beta}}$ which minimizes

$$(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \quad \Rightarrow \quad \hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}(\mathbf{X}'\mathbf{V}^{-1})\mathbf{y}$$

The use of GLS forces the researcher to supply the values of $\mathbf{V}$, hence the values for $\mathbf{G}$ and $\mathbf{R}$. In most situations, the values for $\mathbf{G}$ and $\mathbf{R}$ are unknown. An alternative is to use estimated GLS, in which efficient estimators of $\mathbf{G}$ and $\mathbf{R}$ are used. This then provides estimates but not very efficient estimates of $\boldsymbol{\beta}$.

In most settings, the better approach is to use **likelihood-based** estimation procedures under the condition that $\mathbf{u}$ and $\mathbf{e}$ are restricted to have Gaussian distributions, a restriction that is not required in LSE and GLE. However, likelihood-based estimators yield more efficient estimators in the cases where $\mathbf{V}$ does not equal $\sigma_e^2 \mathbf{I}_n$

Matrix Representation of Mixed Model

Mixed model with factor A having 4 fixed levels, factor B having 3 random levels, and 2 reps :

$$Y_{ijk} = \mu + \tau_i + b_j + (\tau b)_{ij} + e_{ijk}; \quad i = 1,2,3,4; \quad j = 1,2,3; \quad k = 1,2$$

with $\tau_4 = 0$, b_j iid $N(0, \sigma_B^2)$, $(\tau b)_{ij}$ iid $N(0, \sigma_{A*B}^2)$, e_{ijk} iid $N(0, \sigma_e^2)$, all r.v.'s are independent
In matrix form, the above model can be represented as

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \boldsymbol{\epsilon},$$

where $\mathbf{Y}$ - responses, $\mathbf{X}$ - design matrix fixed effects, $\mathbf{Z}$ - design matrix random effects, $\boldsymbol{\beta}$ - fixed effects parameters, $\mathbf{u}$ - random effects, and $\boldsymbol{\epsilon}$ - residuals.

$$\mathbf{Y} = \begin{bmatrix} Y_{111} \\ Y_{112} \\ Y_{121} \\ Y_{122} \\ Y_{131} \\ Y_{132} \\ Y_{211} \\ Y_{212} \\ Y_{221} \\ Y_{222} \\ Y_{231} \\ Y_{232} \\ Y_{311} \\ Y_{312} \\ Y_{321} \\ Y_{322} \\ Y_{331} \\ Y_{332} \\ Y_{411} \\ Y_{412} \\ Y_{421} \\ Y_{422} \\ Y_{431} \\ Y_{432} \end{bmatrix}_{24x1} \qquad \mathbf{X} = \begin{bmatrix} 1_6 & 1_6 & 0_6 & 0_6 \\ 1_6 & 0_6 & 1_6 & 0_6 \\ 1_6 & 0_6 & 0_6 & 1_6 \\ 1_6 & 0_6 & 0_6 & 0_6 \end{bmatrix}_{24x4} \qquad \boldsymbol{\beta} = \begin{bmatrix} \mu \\ \tau_1 \\ \tau_2 \\ \tau_3 \end{bmatrix}_{4x1} \qquad \mathbf{e} = \begin{bmatrix} e_{111} \\ e_{112} \\ e_{121} \\ e_{122} \\ e_{131} \\ e_{132} \\ e_{211} \\ e_{212} \\ e_{221} \\ e_{222} \\ e_{231} \\ e_{232} \\ e_{311} \\ e_{312} \\ e_{321} \\ e_{322} \\ e_{331} \\ e_{332} \\ e_{411} \\ e_{412} \\ e_{421} \\ e_{422} \\ e_{431} \\ e_{432} \end{bmatrix}_{24x1}$$

$$\mathbf{Z} = \begin{bmatrix}
1_2 & 0_2 & 0_2 & 1_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 \\
0_2 & 1_2 & 0_2 & 0_2 & 1_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 \\
0_2 & 0_2 & 1_2 & 0_2 & 0_2 & 1_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 \\
1_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 1_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 \\
0_2 & 1_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 1_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 \\
0_2 & 0_2 & 1_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 1_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 \\
1_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 1_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 \\
0_2 & 1_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 1_2 & 0_2 & 0_2 & 0_2 & 0_2 \\
0_2 & 0_2 & 1_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 1_2 & 0_2 & 0_2 & 0_2 \\
1_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 1_2 & 0_2 & 0_2 \\
0_2 & 1_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 1_2 & 0_2 \\
0_2 & 0_2 & 1_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 1_2
\end{bmatrix}_{24x15} \qquad \mathbf{u} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ (\tau b)_{11} \\ (\tau b)_{12} \\ (\tau b)_{13} \\ (\tau b)_{21} \\ (\tau b)_{22} \\ (\tau b)_{23} \\ (\tau b)_{31} \\ (\tau b)_{32} \\ (\tau b)_{33} \\ (\tau b)_{41} \\ (\tau b)_{42} \\ (\tau b)_{43} \end{bmatrix}_{15x1}$$

Maximum Likelihood Estimation:

Select $\hat{\boldsymbol{\beta}}$, $\hat{\mathbf{G}}$, and $\hat{\mathbf{R}}$ to maximize the likelihood $L(\boldsymbol{\beta}, \mathbf{G}, \mathbf{R})$ or

minimize $l(\boldsymbol{\beta}, \mathbf{G}, \mathbf{R}) = -2log(L(\boldsymbol{\beta}, \mathbf{G}, \mathbf{R}))$:

$$l(\boldsymbol{\beta}, \mathbf{G}, \mathbf{R}) = nlog(2\pi) + log|\mathbf{V}| + (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^{'}\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$$

where $\mathbf{V} = \mathbf{Z}\mathbf{G}\mathbf{Z}^{'} + \mathbf{R}$.

Typically, obtaining the MLE involves repeated iteration between obtaining an estimate of $\boldsymbol{\beta}$ and then an estimate of $\mathbf{V}$ (for example, Newton-Raphson algorithm). There are serious numerical challenges in obtaining the estimates when $\mathbf{V}$ is of a complex form. The MLE's are biased estimators and the size of the bias can be serious in the case of several random factors. An alternative procedure is the restricted maximum likelihood estimators:

Restricted Maximum Likelihood Estimates (REML):

The -2log-likelihood function is partitioned into two components. The first component does not involve $\boldsymbol{\beta}$:

$$(n-r)log(2\pi) + log(|\mathbf{V}|) + log(|\mathbf{X}^{'}\mathbf{V}^{-1}\mathbf{X}|) + (n-r)log(\mathbf{P}^{'}\mathbf{V}^{-1}\mathbf{P})$$

where r is the rank of

$$\mathbf{P} = \mathbf{y} - \mathbf{X}(\mathbf{X}^{'}\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}^{'}\mathbf{V}^{-1}\mathbf{y}$$

Maximizing the above yields the REML estimator of the variance components, $\mathbf{V}$, $\hat{\mathbf{V}}$. This involves an iteration procedure as in the MLE.

The value of $\hat{\mathbf{V}}$ is then used to obtain an estimator of the fixed effects parameters, $\boldsymbol{\beta}$:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^{'}\hat{\mathbf{V}}^{-1}\mathbf{X})^{-1}\mathbf{X}^{'}\hat{\mathbf{V}}^{-1}\mathbf{y}$$

and the predicted values for the random effects $\mathbf{u}$

$$\hat{\mathbf{u}} = \hat{\mathbf{G}}\mathbf{Z}^{'}\hat{\mathbf{V}}^{-1}(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})$$

Statistical Properties of REML's

If $\mathbf{G}$ and $\mathbf{R}$ are known then $\hat{\boldsymbol{\beta}}$ is the Best Linear Unbiased Estimator (BLUE) of $\boldsymbol{\beta}$ and $\hat{\mathbf{u}}$ is the Best Linear Unbiased Predictor (BLUP) of $\mathbf{u}$, where best means minimum mean square error. The problem is that $\mathbf{G}$ and $\mathbf{R}$ are generally unknown. In this case, asymptotic properties of the estimators are used.

In most situations, REML are the preferred estimators of variance components. This is the default option in PROC MIXED when using SAS to estimate variance components and test hypotheses in mixed models.

RESIDUALS FROM MIXED MODEL

With the mixed model written as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e}$$

there are a number of ways of considering the residuals from the mixed model. The **marginal and conditional means** from this model are respectively,

$$E[\mathbf{y}] = \mathbf{X}\boldsymbol{\beta} \qquad \text{and} \qquad E[\mathbf{Y}|\mathbf{u}] = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u}$$

There are two possible sets of residuals depending on whether we use the marginal or conditional residuals.

The **Marginal Residual Vector** is given by

$$\mathbf{r}_m = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}$$

The **Conditional Residual Vector** is given by

$$\mathbf{r}_c = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}} - \mathbf{Z}\hat{\mathbf{u}} = \mathbf{r}_m - \mathbf{Z}\hat{\mathbf{u}}$$

From the above two definitions and considering two different approaches to standardizing the residuals we obtain the following six possible residuals from the fitted mixed model:

Residual Type	Marginal	Conditional
Raw	$r_{mi} = y_i - \mathbf{x}_i'\hat{\boldsymbol{\beta}}$	$r_{ci} = y_i - \mathbf{x}_i'\hat{\boldsymbol{\beta}} - \mathbf{z}_i'\hat{\mathbf{u}}$
Studentized	$rs_{mi} = \frac{r_{mi}}{\widehat{SE}(r_{mi})}$	$rs_{ci} = \frac{r_{ci}}{\widehat{SE}(r_{ci})}$
Pearson	$rp_{mi} = \frac{r_{mi}}{\widehat{SE}(y_i)}$	$rp_{ci} = \frac{r_{ci}}{\widehat{SE}(y_i|\mathbf{u})}$

Which of the residuals is more effective in assessing model conditions is open to debate. Both the Studentized residual adjusts the residual for its estimated variance and hence is a more meaningful measure of model fit than is the Raw residuals. The Pearson residual selects a slightly different value for scaling the residuals than the value selected in the Studentized residuals. The choice may depend on computational ease more than any other reason.

APPROXIMATE F-TESTS

In a number of the experiments involving random factor levels that we have discussed, the form of the F-test for testing the null hypothesis that a variance component is 0 is not evident after examining the Expected Mean Squares. For example, consider a three factor experiment with the levels F_1-Fixed, F_2-Random, and F_3-Random. We obtained the Expected Mean Squares but there was no exact F-test for testing the variance components corresponding to the main effects of the three factors:

Model III: F_1-Fixed, F_2-Random, and F_3-Random

Source of Variation	MS	Expected MS	Null Hypothesis	F-Statistic
F_1	MS_{F_1}	$\sigma_e^2 + r\sigma^2_{F_1*F_2*F_3} + rc\sigma^2_{F_1*F_2} + rb\sigma^2_{F_1*F_3} + rbcQ_{F_1}$	$H_o : Q_{F_1} = 0$	$*$
F_2	MS_{F_2}	$\sigma_e^2 + r\sigma^2_{F_1*F_2*F_3} + rc\sigma^2_{F_1*F_2} + ra\sigma^2_{F_2*F_3} + rac\sigma^2_{F_2}$	$H_o : \sigma^2_{F_2} = 0$	$**$
F_3	MS_{F_3}	$\sigma_e^2 + r\sigma^2_{F_1*F_2*F_3} + rb\sigma^2_{F_1*F_3} + ra\sigma^2_{F_2*F_3} + rab\sigma^2_{F_3}$	$H_o : \sigma^2_{F_3} = 0$	$***$
$F_1 * F_2$	$MS_{F_1*F_2}$	$\sigma_e^2 + r\sigma^2_{F_1*F_2*F_3} + rc\sigma^2_{F_1*F_2}$	$H_o : \sigma^2_{F_1*F_2} = 0$	$\frac{MS_{F_1*F_2}}{MS_{F_1*F_2*F_3}}$
$F_1 * F_3$	$MS_{F_1*F_3}$	$\sigma_e^2 + r\sigma^2_{F_1*F_2*F_3} + rb\sigma^2_{F_1*F_3}$	$H_o : \sigma^2_{F_1*F_3} = 0$	$\frac{MS_{F_1*F_3}}{MS_{F_1*F_2*F_3}}$
$F_2 * F_3$	$MS_{F_2*F_3}$	$\sigma_e^2 + r\sigma^2_{F_1*F_2*F_3} + ra\sigma^2_{F_2*F_3}$	$H_o : \sigma^2_{F_2*F_3} = 0$	$\frac{MS_{F_2*F_3}}{MS_{F_1*F_2*F_3}}$
$F_1 * F_2 * F_3$	$MS_{F_1*F_2*F_3}$	$\sigma_e^2 + r\sigma^2_{F_1*F_2*F_3}$	$H_o : \sigma^2_{F_1*F_2*F_3} = 0$	$\frac{MS_{F_1*F_2*F_3}}{MSE}$
Error	MSE	σ_e^2		

There are two solutions for obtaining the F-tests:

1. **Solution I: Use the Satterthwaite Approximation:**

Approximate F-tests are obtained by obtaining a linear combination of Mean Squares, M, such that the expected value of M equals the expected value of the MS, under H_o, for the corresponding source of variation, e.g., $E[M] = E_{H_o}[MS_{F_1}]$.

1. For F_1, $* = \frac{MS_{F_1}}{M}$ where $M = MS_{F_1*F_2} + MS_{F_1*F_3} - MS_{F_1*F_2*F_3}$

$$\begin{aligned} E[M] &= EMS_{F_1*F_2} + EMS_{F_1*F_3} - EMS_{F_1*F_2*F_3} \\ &= (\sigma_e^2 + r\sigma^2_{F_1*F_2*F_3} + rc\sigma^2_{F_1*F_2}) + (\sigma_e^2 + r\sigma^2_{F_1*F_2*F_3} + rb\sigma^2_{F_1*F_3}) - (\sigma_e^2 + r\sigma^2_{F_1*F_2*F_3}) \\ &= \sigma_e^2 + r\sigma^2_{F_1*F_2*F_3} + rc\sigma^2_{F_1*F_2} + rb\sigma^2_{F_1*F_3} \\ &= E_{H_o}[MS_{F_1}] \end{aligned}$$

The degrees of freedom for the F-test are (a-1) and ν, where ν is obtained using the Satterthwaite Aproximation:

$$df = \frac{(MS_{F_1*F_2} + MS_{F_1*F_3} - MS_{F_1*F_2*F_3})^2}{\frac{(MS_{F_1*F_2})^2}{df_{MS_{F_1*F_2}}} + \frac{(MS_{F_1*F_3})^2}{df_{MS_{F_1*F_3}}} + \frac{(MS_{F_1*F_2*F_3})^2}{df_{MS_{F_1*F_2*F_3}}}}$$

In a similar fashion, obtain the other two F-tests.

2. For F_2, $** = \frac{MS_{F_2}}{M}$ where $M = MS_{F_1*F_2} + MS_{F_2*F_3} - MS_{F_1*F_2*F_3}$

3. For F_3, $*** = \frac{MS_{F_3}}{M}$ where $M = MS_{F_1*F_3} + MS_{F_2*F_3} - MS_{F_1*F_2*F_3}$

2. **Solution II: Pool Mean Squares for Non-significant Interactions:**

In the situation just encountered with F_1 fixed levels and F_2, F_3 having random levels, suppose our test of $H_o : \sigma_{F_1*F_2*F_3} = 0$ is not rejected.

The test of $H_o : \sigma_{F_1*F_2} = 0$ versus $H_1 : \sigma_{F_1*F_2} \neq 0$ raises a question: "Should we consider $\sigma_{F_1*F_2*F_3} = 0$ and hence use the F-test $F = \frac{MS_{F_1*F_2}}{MSE}$ or take into account that a Type II error may have occurred and use the F-test $F = \frac{MS_{F_1*F_2}}{MS_{F_1*F_2*F_3}}$?

Bancroft(1964) *Biometrics*, pp. 427-442, suggests the following solution: Consider the general situation:

we have 3 variance components: $\sigma_1, \sigma_2, \sigma_3$ with associated Mean Squares: MS_1, MS_2, MS_3 having $EMS_1 = \sigma_1^2, EMS_2 = \sigma_1^2 + \sigma_2^2, EMS_3 = \sigma_1^2 + \sigma_2^2 + \sigma_3^2$.

We want to test $H_o : \sigma_3 = 0$ versus $H_1 : \sigma_3 \neq 0$. However, based on an earlier test $\sigma_2 = 0$ was not rejected. Thus, should we use $\frac{MS_3}{MS_2}$ or $\frac{MS_3}{MS_{pooled}}$ as the test statistic,

where $MS_{pooled} = \frac{df_1 MS_1 + df_2 MS_2}{df_1 + df_2}$?

(a) Test $H_o : \sigma_2 = 0$ versus $H_1 : \sigma_2 \neq 0$ using $F = \frac{MS_2}{MS_1}$ at level α_1

(b) If $H_o : \sigma_2 = 0$ is rejected then use $F = \frac{MS_3}{MS_2}$ at level α_2 to test $H_o : \sigma_3 = 0$ versus $H_1 : \sigma_3 \neq 0$

(c) If $H_o : \sigma_2 = 0$ is not rejected then use $F = \frac{MS_3}{MS_{pooled}}$ at level α_3 to test $H_o : \sigma_3 = 0$ versus $H_1 : \sigma_3 \neq 0$

Question: What are the appropriate values for α_1, α_2, and α_3, so that the Probability of a Type I error for testing $H_o : \sigma_3 = 0$ versus $H_1 : \sigma_3 \neq 0$ is α?

The exact value of α depends on **8** values:

$$\alpha_1, \quad \alpha_2, \quad \alpha_3, \quad df_{MS_1}, \quad df_{MS_2}, \quad df_{MS_3}, \quad \theta_1 = \frac{\sigma_2^2}{\sigma_1^2}, \quad \theta_2 = \frac{\sigma_2^2}{\sigma_3^2}$$

The problem is that only $df_{MS_1}, df_{MS_2}, df_{MS_3}$ are known values. The values of $\alpha_1, \alpha_2, \alpha_3$ are restricted because we want the value of $P[TypeIError] = \alpha$ for testing $H_o : \sigma_{F_1*F_2} = 0$ versus $H_1 : \sigma_{F_1*F_2} \neq 0$.

The paper by Bancroft suggests the following options:

Option 1: When $\theta_1 = \frac{\sigma_2^2}{\sigma_1^2}$ is small, say, $1 \leq \theta_1 \leq 2$,

then select $\alpha_1 = .50$ if $df_{MS_3} \geq df_{MS_2}$ and $df_{MS_1} \geq 5df_2$

otherwise, select $\alpha_1 = .25$.

In both cases, then, select $\alpha_2 = \alpha_3 = \alpha$.

Option 2: If θ_1 is large or totally unknown, select $\alpha_1 = .01$ and use $\alpha_2 = \alpha_3 = \alpha$.

Option 3: If θ_1 is large or totally unknown, always use $F = \frac{MS_3}{MS_2}$ at level α.

In our example with F_1 having fixed levels, F_2, F_3 having random levels,

$$EMSE = \sigma_e^2, \quad EMS_{F_1*F_2*F_3} = \sigma_e^2 + r\sigma^2_{F_1*F_2*F_3}, \quad EMS_{F_1*F_2} = \sigma_e^2 + r\sigma^2_{F_1*F_2*F_3} + rc\sigma^2_{F_1*F_2}$$

1. Test $H_o : \sigma_{F_1*F_2*F_3} = 0$ versus $H_1 : \sigma_{F_1*F_2*F_3} \neq 0$ using $F = \frac{MS_{F_1*F_2*F_3}}{MSE}$ at level α_1
2. If $H_o : \sigma_{F_1*F_2*F_3} = 0$ is rejected then use $F = \frac{MS_{F_1*F_2}}{MS_{F_1*F_2*F_3}}$ at level α_2 to test $H_o : \sigma_{F_1*F_2} = 0$ versus $H_1 : \sigma_{F_1*F_2} \neq 0$
3. If $H_o : \sigma_{F_1*F_2*F_3} = 0$ is not rejected then use $F = \frac{MS_{F_1*F_2}}{MS_{pooled}}$ at level α_3 to test $H_o : \sigma_{F_1*F_2} = 0$ versus $H_1 : \sigma_{F_1*F_2} \neq 0$

Question: What are the appropriate values for α_1, α_2, and α_3, so that the Probability of a Type I error for testing $H_o : \sigma_{F_1*F_2} = 0$ versus $H_1 : \sigma_{F_1*F_2} \neq 0$ is α?

Option 1: When $\theta_1 = \frac{\sigma^2_{F_1*F_2*F_3}}{\sigma_e^2}$ is small, say, $1 \leq \theta_1 \leq 2$,

then select $\alpha_1 = .50$ if $df_{MS_{F_1*F_2}} \geq df_{MS_{F_1*F_2*F_3}}$ and $df_{MSE} \geq 5df_{F_1*F_2*F_3}$

otherwise, select $\alpha_1 = .25$.

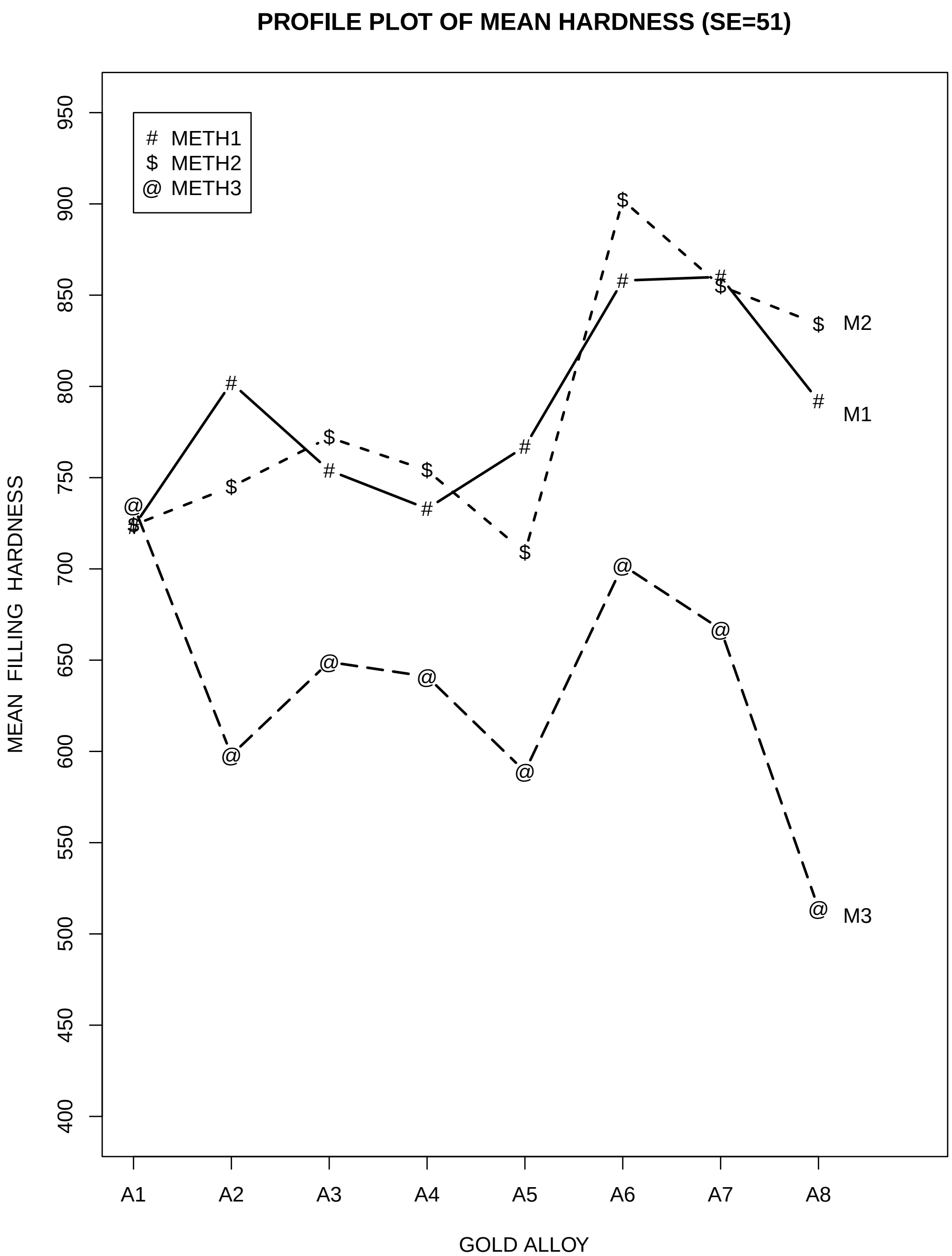
In both cases, then, select $\alpha_2 = \alpha_3 = \alpha$.

Option 2: If θ_1 is large or totally unknown, select $\alpha_1 = .01$ and use $\alpha_2 = \alpha_3 = \alpha$.

Option 3: If θ_1 is large or totally unknown, use $F = \frac{MS_{F_1*F_2}}{F_1*F_2*F_3}$ at level α.

Example of the Analysis of a Mixed Factors Experiment

We will now illustrate the analysis of the Dental Fillings Experiment in which we had three factors - one with random levels and two factors with fixed levels. The following SAS program and SAS output will be used to assess the significance of the three factors: Dentist, Condensation Method, and Gold Alloy on the hardness of the fillings. The following plot illustrates the relationship between type of alloy and condensation method:



```
* random3factors_mixed.sas;
ODS HTML; ODS GRAPHICS ON;
option ls=80 ps=50 nocenter nodate;
TITLE 'AOV - MIXED FACTOR LEVELS';
DATA RAW;
INPUT D $ M $ @@;
DO G = 1 TO 8;
INPUT Y @@; OUTPUT; END;
LABEL G='GOLD ALLOW' D = 'DENTIST'  M = 'CONDENSATION METHOD';
cards;
1           1    792 824 813 792 792 907  792 835
 1           2    772 772 782 698 665 1115 835 870
 1           3    782 803 752 620 835 847 560 585
 2           1    803 803 715 803 813 858 907 882
 2           2    752 772 772 782 743 933 792 824
 2           3    715 707 835 715 673 698 734 681
 3           1    715 724 743 627 752 858 762 724
 3           2    792 715 813 743 613 824 847 782
 3           3    762 606 743 681 743 715 824 681
 4           1    673 946 792 743 762 894 792 649
 4           2    657 743 690 882 772 813 870 858
 4           3    690 245 493 707 289 715 813 312
 5           1    634 715 707 698 715 772 1048 870
 5           2    649 724 803 665 752 824 933 835
 5           3    724 627 421 483 405 536 405 312
RUN;
PROC MIXED METHOD = TYPE3  CL;           & AOV-MOM
CLASS D M G;
MODEL Y = M G M*G;
RANDOM D D*M D*G/ CL ALPHA=.05;
LSMEANS M G M*G/CL ADJUST=TUKEY;
RUN;
PROC MIXED METHOD=REML CL;       & REML
CLASS D M G;
MODEL Y = M G M*G;
RANDOM D D*M D*G/ CL ALPHA=.05;
RUN;

ODS GRAPHICS OFF; ODS HTML CLOSE;
```

* see additional pdf output for reduced H.O. b/c there is a problem w/ the output below. (Random 3 factor Mixed)

AOV - MIXED FACTOR LEVELS

The Mixed Procedure

Model Information	
Data Set	WORK.RAW
Dependent Variable	Y
Covariance Structure	Variance Components
Estimation Method	Type 3
Residual Variance Method	Factor
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Containment

Class Level Information		
Class	**Levels**	**Values**
D	5	1 2 3 4 5
M	3	1 2 3
G	8	1 2 3 4 5 6 7 8

Dimensions	
Covariance Parameters	4
Columns in X	36
Columns in Z	60
Subjects	1
Max Obs per Subject	120

Number of Observations	
Number of Observations Read	120
Number of Observations Used	120
Number of Observations Not Used	0

Type 3 Analysis of Variance

Source	DF	Sum of Squares	Mean Square	Expected Mean Square	Error Term	Error DF	F Value	Pr > F
M	2	597615	298808	Var(Residual) + 8 Var(D*M) + Q(M,M*G)	MS(D*M)	8	9.07	0.0088
G	7	220338	31477	Var(Residual) + 3 Var(D*G) + Q(G,M*G)	MS(D*G)	28	4.22	0.0027
M*G	14	209773	14984	Var(Residual) + Q(M*G)	MS(Residual)	56	1.50	0.1403

AOV - MIXED FACTOR LEVELS

The Mixed Procedure

Type 3 Analysis of Variance								
Source	**DF**	**Sum of Squares**	**Mean Square**	**Expected Mean Square**	**Error Term**	**Error DF**	**F Value**	**Pr > F**
D	4	217576	54394	Var(Residual) + 3 Var(D*G) + 8 Var(D*M) + 24 Var(D)	MS(D*M) + MS(D*G) - MS(Residual)	6.6421	1.79	0.2403
D*M	8	263441	32930	Var(Residual) + 8 Var(D*M)	MS(Residual)	56	3.30	0.0037
D*G	28	208814	7457.652976	Var(Residual) + 3 Var(D*G)	MS(Residual)	56	0.75	0.7966
Residual	56	558258	9968.885119	Var(Residual)	.	.	.	.

Covariance Parameter Estimates				
Cov Parm	**Estimate**	**Alpha**	**Lower**	**Upper**
D	998.97	0.05	-2425.11	4423.05
D*M	2870.15	0.05	-1190.03	6930.34
D*G	-837.08	0.05	-2628.88	954.72
Residual	9968.89	0.05	7105.48	15002

Fit Statistics	
-2 Res Log Likelihood	1203.2
AIC (Smaller is Better)	1211.2
AICC (Smaller is Better)	1211.6
BIC (Smaller is Better)	1209.6

Solution for Random Effects											
Effect	**DENTIST**	**CONDENSATION METHOD**	**GOLD ALLOW**	**Estimate**	**Std Err Pred**	**DF**	**t Value**	**Pr > |t|**	**Alpha**	**Lower**	**Upper**
D	1			21.3112	25.4306	56	0.84	0.4056	0.05	-29.6325	72.2548
D	2			18.9604	25.4306	56	0.75	0.4590	0.05	-31.9833	69.9040
D	3			2.0092	25.4306	56	0.08	0.9373	0.05	-48.9345	52.9528
D	4			-16.1542	25.4306	56	-0.64	0.5279	0.05	-67.0978	34.7895
D	5			-26.1266	25.4306	56	-1.03	0.3087	0.05	-77.0702	24.8171
D*M	1	1		9.1664	38.4671	56	0.24	0.8125	0.05	-67.8926	86.2253
D*M	1	2		5.2965	38.4671	56	0.14	0.8910	0.05	-71.7625	82.3554
D*M	1	3		46.7667	38.4671	56	1.22	0.2292	0.05	-30.2922	123.83
D*M	2	1		13.8587	38.4671	56	0.36	0.7200	0.05	-63.2003	90.9176
D*M	2	2		-5.3512	38.4671	56	-0.14	0.8899	0.05	-82.4101	71.7078

AOV - MIXED FACTOR LEVELS

The Mixed Procedure

Solution for Random Effects											
Effect	DENTIST	CONDENSATION METHOD	GOLD ALLOW	Estimate	Std Err Pred	DF	t Value	Pr > |t|	Alpha	Lower	Upper
D*M	2	3		45.9680	38.4671	56	1.19	0.2371	0.05	-31.0909	123.03
D*M	3	1		-34.7407	38.4671	56	-0.90	0.3703	0.05	-111.80	42.3183
D*M	3	2		-15.7749	38.4671	56	-0.41	0.6833	0.05	-92.8338	61.2841
D*M	3	3		56.2881	38.4671	56	1.46	0.1490	0.05	-20.7708	133.35
D*M	4	1		6.7546	38.4671	56	0.18	0.8612	0.05	-70.3044	83.8135
D*M	4	2		9.1601	38.4671	56	0.24	0.8127	0.05	-67.8988	86.2191
D*M	4	3		-62.3276	38.4671	56	-1.62	0.1108	0.05	-139.39	14.7313
D*M	5	1		4.9611	38.4671	56	0.13	0.8978	0.05	-72.0979	82.0200
D*M	5	2		6.6694	38.4671	56	0.17	0.8630	0.05	-70.3896	83.7283
D*M	5	3		-86.6953	38.4671	56	-2.25	0.0281	0.05	-163.75	-9.6363
D*G	1		1	-4.3143	0	56	-Infty	<.0001	.	.	.
D*G	1		2	-14.4387	0	56	-Infty	<.0001	.	.	.
D*G	1		3	-5.2796	0	56	-Infty	<.0001	.	.	.
D*G	1		4	16.0468	0	56	Infty	<.0001	.	.	.
D*G	1		5	-11.4530	0	56	-Infty	<.0001	.	.	.
D*G	1		6	-31.6570	0	56	-Infty	<.0001	.	.	.
D*G	1		7	36.0262	0	56	Infty	<.0001	.	.	.
D*G	1		8	-2.7878	0	56	-Infty	<.0001	.	.	.
D*G	2		1	2.6666	0	56	Infty	<.0001	.	.	.
D*G	2		2	-2.8559	0	56	-Infty	<.0001	.	.	.
D*G	2		3	-4.0232	0	56	-Infty	<.0001	.	.	.
D*G	2		4	-6.8293	0	56	-Infty	<.0001	.	.	.
D*G	2		5	-5.9313	0	56	-Infty	<.0001	.	.	.
D*G	2		6	9.4461	0	56	Infty	<.0001	.	.	.
D*G	2		7	6.8645	0	56	Infty	<.0001	.	.	.
D*G	2		8	-15.2252	0	56	-Infty	<.0001	.	.	.
D*G	3		1	-8.3958	0	56	-Infty	<.0001	.	.	.
D*G	3		2	12.5713	0	56	Infty	<.0001	.	.	.
D*G	3		3	-12.6162	0	56	-Infty	<.0001	.	.	.
D*G	3		4	9.9448	0	56	Infty	<.0001	.	.	.
D*G	3		5	-3.5245	0	56	-Infty	<.0001	.	.	.
D*G	3		6	8.5979	0	56	Infty	<.0001	.	.	.

AOV - MIXED FACTOR LEVELS

The Mixed Procedure

Solution for Random Effects											
Effect	DENTIST	CONDENSATION METHOD	GOLD ALLOW	Estimate	Std Err Pred	DF	t Value	Pr > |t|	Alpha	Lower	Upper
D*G	3		7	-4.3102	0	56	-Infty	<.0001	.	.	.
D*G	3		8	-3.9510	0	56	-Infty	<.0001	.	.	.
D*G	4		1	7.5792	0	56	Infty	<.0001	.	.	.
D*G	4		2	13.0567	0	56	Infty	<.0001	.	.	.
D*G	4		3	11.7772	0	56	Infty	<.0001	.	.	.
D*G	4		4	-33.5694	0	56	-Infty	<.0001	.	.	.
D*G	4		5	16.4914	0	56	Infty	<.0001	.	.	.
D*G	4		6	-6.1819	0	56	-Infty	<.0001	.	.	.
D*G	4		7	-20.9981	0	56	-Infty	<.0001	.	.	.
D*G	4		8	25.3811	0	56	Infty	<.0001	.	.	.
D*G	5		1	2.4644	0	56	Infty	<.0001	.	.	.
D*G	5		2	-8.3335	0	56	-Infty	<.0001	.	.	.
D*G	5		3	10.1419	0	56	Infty	<.0001	.	.	.
D*G	5		4	14.4071	0	56	Infty	<.0001	.	.	.
D*G	5		5	4.4174	0	56	Infty	<.0001	.	.	.
D*G	5		6	19.7949	0	56	Infty	<.0001	.	.	.
D*G	5		7	-17.5824	0	56	-Infty	<.0001	.	.	.
D*G	5		8	-3.4172	0	56	-Infty	<.0001	.	.	.

Type 3 Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
M	2	8	9.07	0.0088
G	7	28	4.22	0.0027
M*G	14	56	1.50	0.1403

Least Squares Means										
Effect	CONDENSATION METHOD	GOLD ALLOW	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper
M	1		786.15	31.6563	8	24.83	<.0001	0.05	713.15	859.15
M	2		786.95	31.6563	8	24.86	<.0001	0.05	713.95	859.95
M	3		636.85	31.6563	8	20.12	<.0001	0.05	563.85	709.85

AOV - MIXED FACTOR LEVELS

The Mixed Procedure

Least Squares Means										
Effect	**CONDENSATION METHOD**	**GOLD ALLOW**	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**	**Alpha**	**Lower**	**Upper**
G		1	727.47	29.8046	28	24.41	<.0001	0.05	666.41	788.52
G		2	715.07	29.8046	28	23.99	<.0001	0.05	654.01	776.12
G		3	724.93	29.8046	28	24.32	<.0001	0.05	663.88	785.99
G		4	709.27	29.8046	28	23.80	<.0001	0.05	648.21	770.32
G		5	688.27	29.8046	28	23.09	<.0001	0.05	627.21	749.32
G		6	820.60	29.8046	28	27.53	<.0001	0.05	759.55	881.65
G		7	794.27	29.8046	28	26.65	<.0001	0.05	733.21	855.32
G		8	713.33	29.8046	28	23.93	<.0001	0.05	652.28	774.39

Differences of Least Squares Means									
Effect	**CONDENSATION METHOD**	**GOLD ALLOW**	**CONDENSATION METHOD**	**GOLD ALLOW**	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**
M	1		2		-0.8000	40.5772	8	-0.02	0.9848
M	1		3		149.30	40.5772	8	3.68	0.0062
M	2		3		150.10	40.5772	8	3.70	0.0061
G		1		2	12.4000	31.5334	28	0.39	0.6971
G		1		3	2.5333	31.5334	28	0.08	0.9365
G		1		4	18.2000	31.5334	28	0.58	0.5684
G		1		5	39.2000	31.5334	28	1.24	0.2241
G		1		6	-93.1333	31.5334	28	-2.95	0.0063
G		1		7	-66.8000	31.5334	28	-2.12	0.0431
G		1		8	14.1333	31.5334	28	0.45	0.6575
G		2		3	-9.8667	31.5334	28	-0.31	0.7567
G		2		4	5.8000	31.5334	28	0.18	0.8554
G		2		5	26.8000	31.5334	28	0.85	0.4026
G		2		6	-105.53	31.5334	28	-3.35	0.0023
G		2		7	-79.2000	31.5334	28	-2.51	0.0181
G		2		8	1.7333	31.5334	28	0.05	0.9566
G		3		4	15.6667	31.5334	28	0.50	0.6232
G		3		5	36.6667	31.5334	28	1.16	0.2547
G		3		6	-95.6667	31.5334	28	-3.03	0.0052
G		3		7	-69.3333	31.5334	28	-2.20	0.0363

AOV - MIXED FACTOR LEVELS

The Mixed Procedure

Differences of Least Squares Means									
Effect	**CONDENSATION METHOD**	**GOLD ALLOW**	**CONDENSATION METHOD**	**GOLD ALLOW**	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**
G		3		8	11.6000	31.5334	28	0.37	0.7157
G		4		5	21.0000	31.5334	28	0.67	0.5109
G		4		6	-111.33	31.5334	28	-3.53	0.0015
G		4		7	-85.0000	31.5334	28	-2.70	0.0118
G		4		8	-4.0667	31.5334	28	-0.13	0.8983
G		5		6	-132.33	31.5334	28	-4.20	0.0002
G		5		7	-106.00	31.5334	28	-3.36	0.0023
G		5		8	-25.0667	31.5334	28	-0.79	0.4333
G		6		7	26.3333	31.5334	28	0.84	0.4107
G		6		8	107.27	31.5334	28	3.40	0.0020
G		7		8	80.9333	31.5334	28	2.57	0.0159

Differences of Least Squares Means									
Effect	**CONDENSATION METHOD**	**GOLD ALLOW**	**CONDENSATION METHOD**	**GOLD ALLOW**	**Adjustment**	**Adj P**	**Alpha**	**Lower**	**Upper**
M	1		2		Tukey-Kramer	0.9998	0.05	-94.3711	92.7711
M	1		3		Tukey-Kramer	0.0153	0.05	55.7289	242.87
M	2		3		Tukey-Kramer	0.0149	0.05	56.5289	243.67
G		1		2	Tukey-Kramer	0.9999	0.05	-52.1932	76.9932
G		1		3	Tukey-Kramer	1.0000	0.05	-62.0599	67.1265
G		1		4	Tukey-Kramer	0.9989	0.05	-46.3932	82.7932
G		1		5	Tukey-Kramer	0.9115	0.05	-25.3932	103.79
G		1		6	Tukey-Kramer	0.0993	0.05	-157.73	-28.5401
G		1		7	Tukey-Kramer	0.4282	0.05	-131.39	-2.2068
G		1		8	Tukey-Kramer	0.9998	0.05	-50.4599	78.7265
G		2		3	Tukey-Kramer	1.0000	0.05	-74.4599	54.7265
G		2		4	Tukey-Kramer	1.0000	0.05	-58.7932	70.3932
G		2		5	Tukey-Kramer	0.9882	0.05	-37.7932	91.3932
G		2		6	Tukey-Kramer	0.0420	0.05	-170.13	-40.9401
G		2		7	Tukey-Kramer	0.2316	0.05	-143.79	-14.6068
G		2		8	Tukey-Kramer	1.0000	0.05	-62.8599	66.3265
G		3		4	Tukey-Kramer	0.9996	0.05	-48.9265	80.2599
G		3		5	Tukey-Kramer	0.9359	0.05	-27.9265	101.26

AOV - MIXED FACTOR LEVELS

The Mixed Procedure

Differences of Least Squares Means									
Effect	**CONDENSATION METHOD**	**GOLD ALLOW**	**CONDENSATION METHOD**	**GOLD ALLOW**	**Adjustment**	**Adj P**	**Alpha**	**Lower**	**Upper**
G		3		6	Tukey-Kramer	0.0839	0.05	-160.26	-31.0735
G		3		7	Tukey-Kramer	0.3824	0.05	-133.93	-4.7401
G		3		8	Tukey-Kramer	0.9999	0.05	-52.9932	76.1932
G		4		5	Tukey-Kramer	0.9973	0.05	-43.5932	85.5932
G		4		6	Tukey-Kramer	0.0274	0.05	-175.93	-46.7401
G		4		7	Tukey-Kramer	0.1657	0.05	-149.59	-20.4068
G		4		8	Tukey-Kramer	1.0000	0.05	-68.6599	60.5265
G		5		6	Tukey-Kramer	0.0053	0.05	-196.93	-67.7401
G		5		7	Tukey-Kramer	0.0406	0.05	-170.59	-41.4068
G		5		8	Tukey-Kramer	0.9920	0.05	-89.6599	39.5265
G		6		7	Tukey-Kramer	0.9894	0.05	-38.2599	90.9265
G		6		8	Tukey-Kramer	0.0370	0.05	42.6735	171.86
G		7		8	Tukey-Kramer	0.2101	0.05	16.3401	145.53

Differences of Least Squares Means						
Effect	**CONDENSATION METHOD**	**GOLD ALLOW**	**CONDENSATION METHOD**	**GOLD ALLOW**	**Adj Lower**	**Adj Upper**
M	1		2		-116.74	115.14
M	1		3		33.3555	265.24
M	2		3		34.1555	266.04
G		1		2	-90.7211	115.52
G		1		3	-100.59	105.65
G		1		4	-84.9211	121.32
G		1		5	-63.9211	142.32
G		1		6	-196.25	9.9877
G		1		7	-169.92	36.3211
G		1		8	-88.9877	117.25
G		2		3	-112.99	93.2544
G		2		4	-97.3211	108.92
G		2		5	-76.3211	129.92
G		2		6	-208.65	-2.4123
G		2		7	-182.32	23.9211
G		2		8	-101.39	104.85

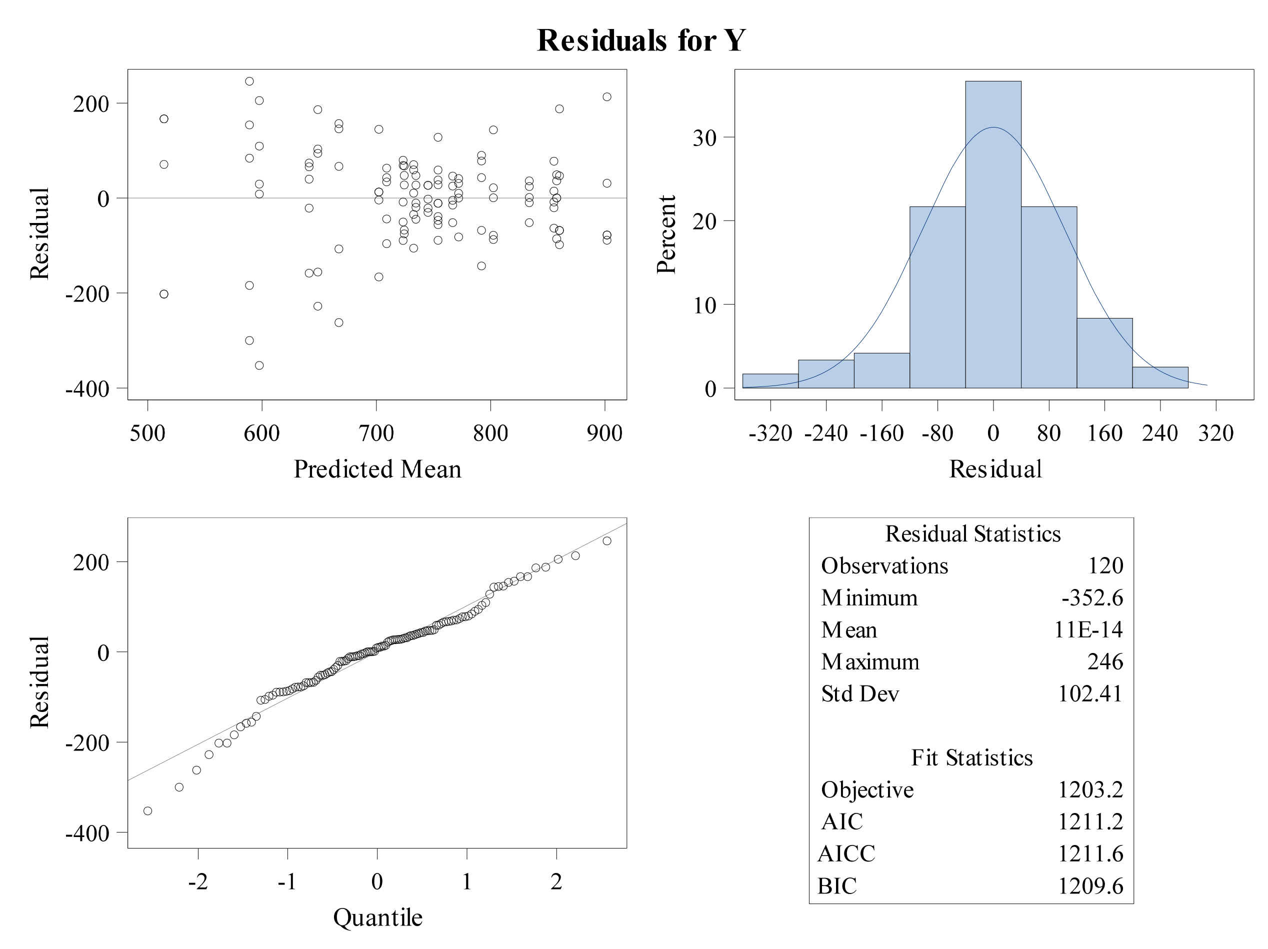
AOV - MIXED FACTOR LEVELS

The Mixed Procedure

Differences of Least Squares Means						
Effect	**CONDENSATION METHOD**	**GOLD ALLOW**	**CONDENSATION METHOD**	**GOLD ALLOW**	**Adj Lower**	**Adj Upper**
G		3		4	-87.4544	118.79
G		3		5	-66.4544	139.79
G		3		6	-198.79	7.4544
G		3		7	-172.45	33.7877
G		3		8	-91.5211	114.72
G		4		5	-82.1211	124.12
G		4		6	-214.45	-8.2123
G		4		7	-188.12	18.1211
G		4		8	-107.19	99.0544
G		5		6	-235.45	-29.2123
G		5		7	-209.12	-2.8789
G		5		8	-128.19	78.0544
G		6		7	-76.7877	129.45
G		6		8	4.1456	210.39
G		7		8	-22.1877	184.05

AOV - MIXED FACTOR LEVELS

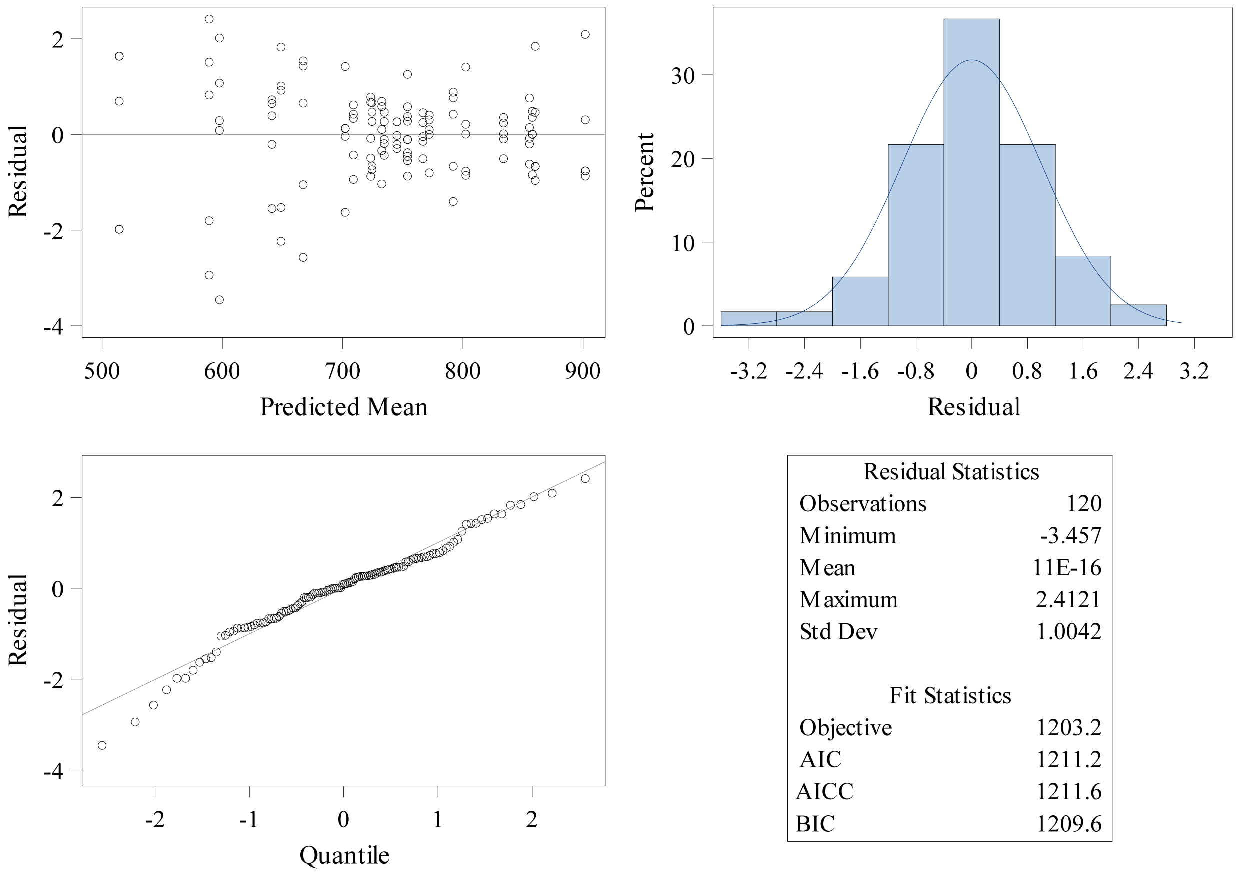
The Mixed Procedure



AOV - MIXED FACTOR LEVELS

The Mixed Procedure

Studentized Residuals for Y



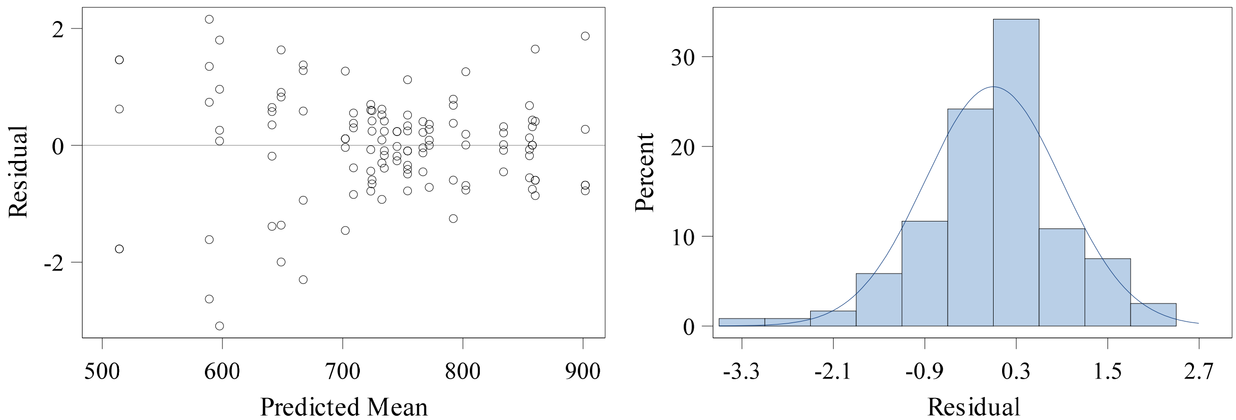
Residual Statistics	
Observations	120
Minimum	-3.457
Mean	11E-16
Maximum	2.4121
Std Dev	1.0042

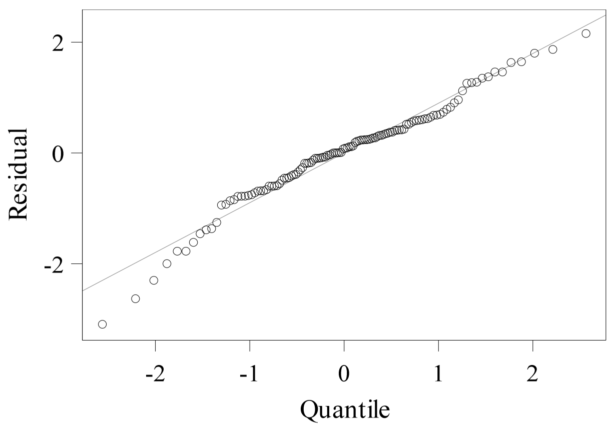
Fit Statistics	
Objective	1203.2
AIC	1211.2
AICC	1211.6
BIC	1209.6

AOV - MIXED FACTOR LEVELS

The Mixed Procedure

Pearson Residuals for Y



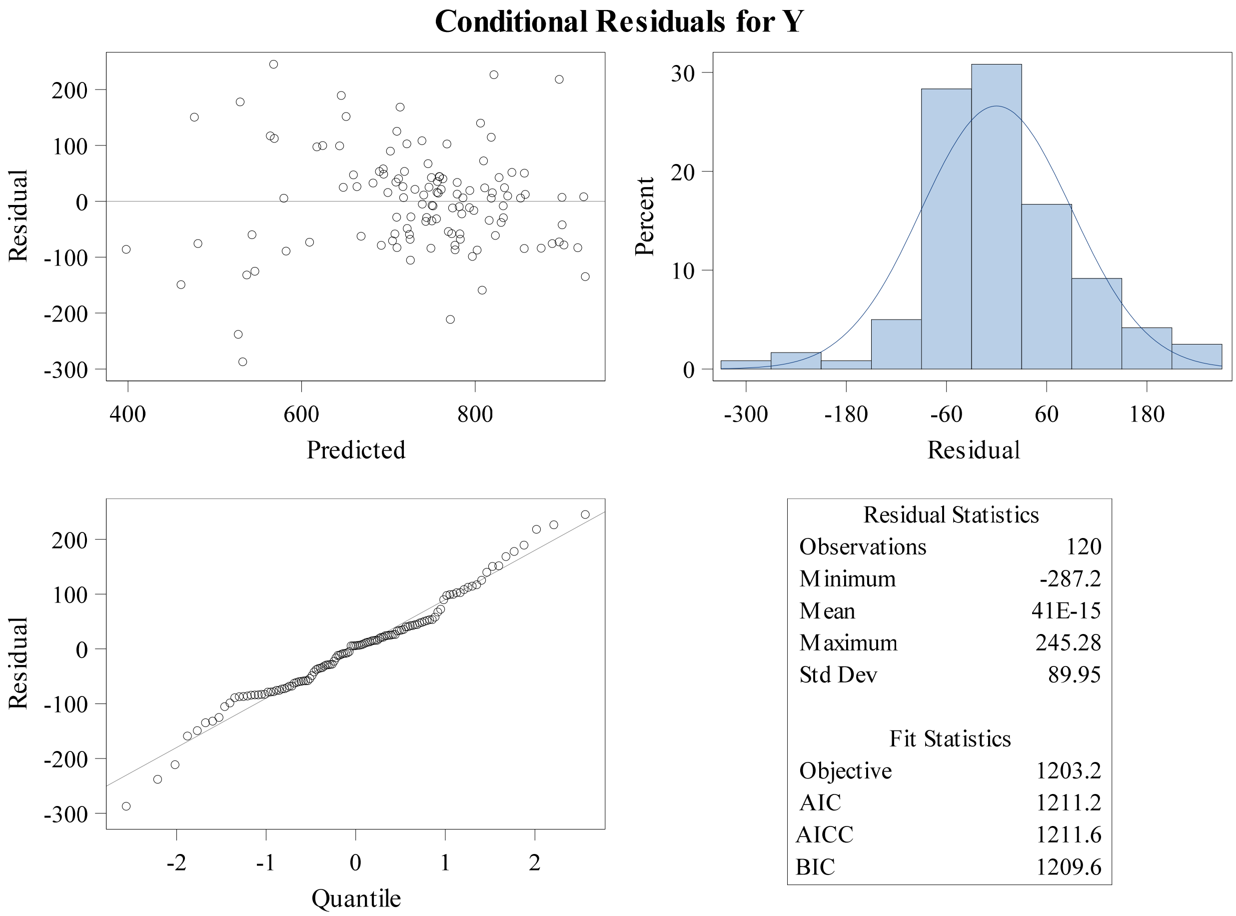


Residual Statistics	
Observations	120
Minimum	-3.092
Mean	1E-15
Maximum	2.1575
Std Dev	0.8982

Fit Statistics	
Objective	1203.2
AIC	1211.2
AICC	1211.6
BIC	1209.6

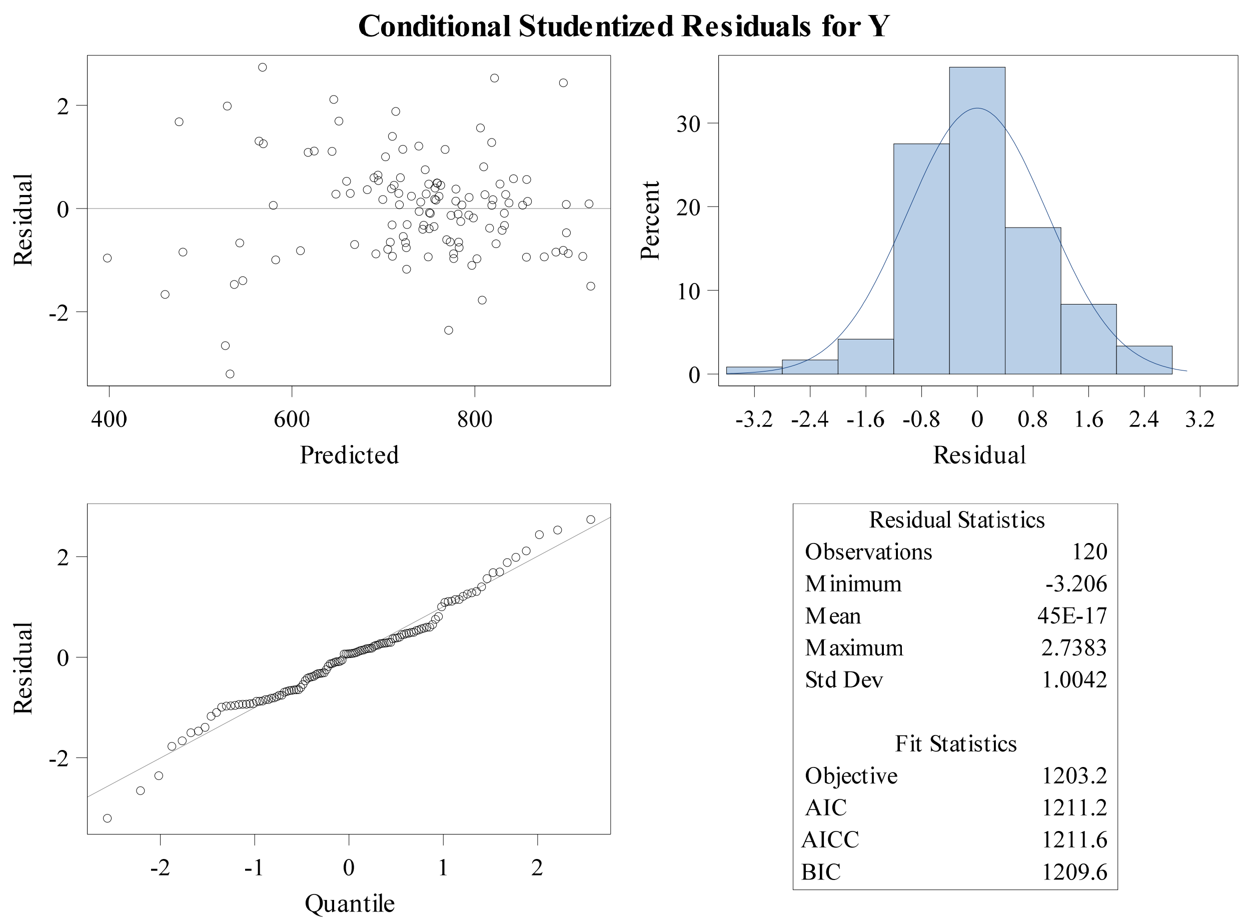
AOV - MIXED FACTOR LEVELS

The Mixed Procedure



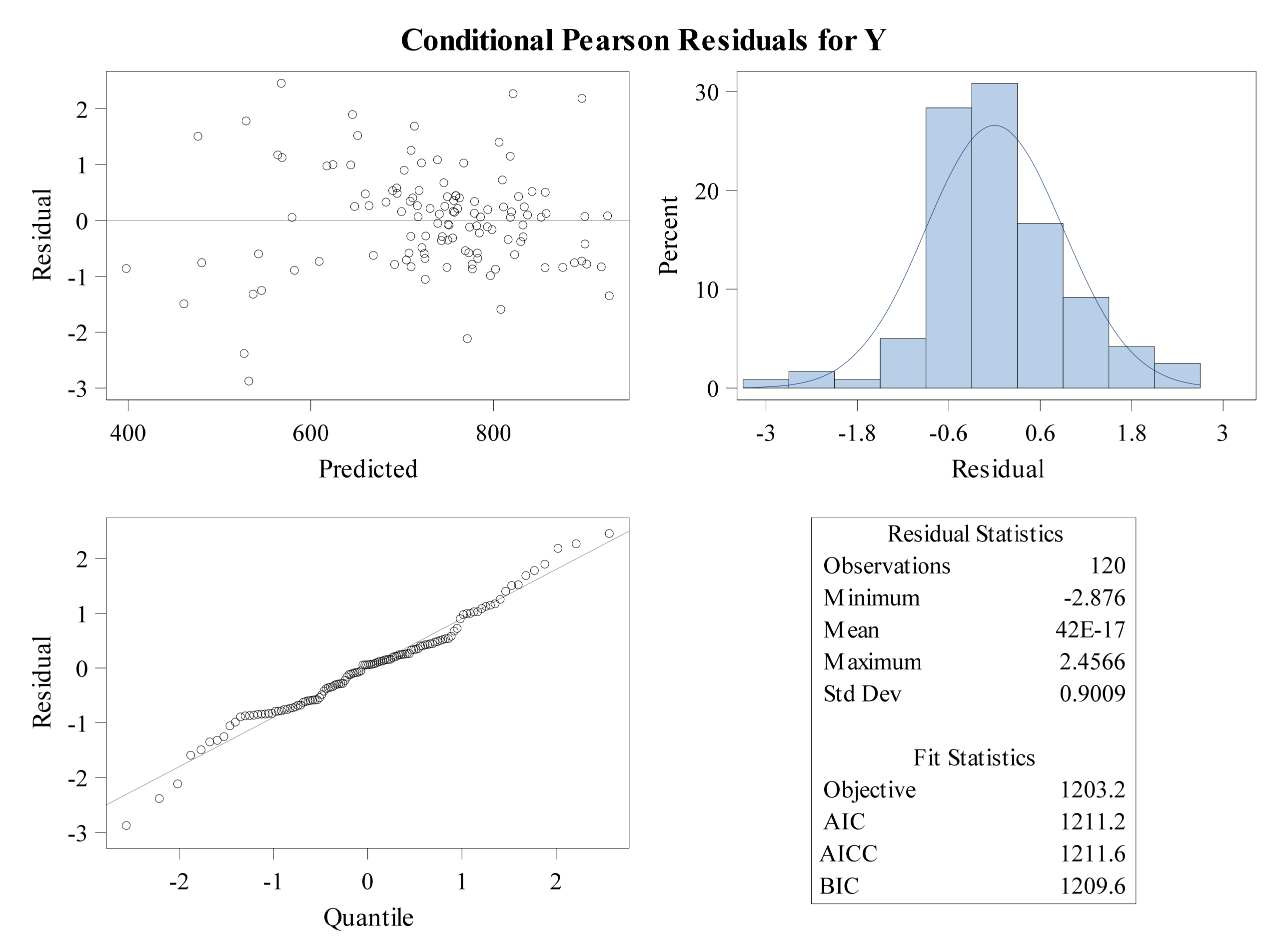
AOV - MIXED FACTOR LEVELS

The Mixed Procedure



AOV - MIXED FACTOR LEVELS

The Mixed Procedure



AOV - MIXED FACTOR LEVELS

The Mixed Procedure

Model Information	
Data Set	WORK.RAW
Dependent Variable	Y
Covariance Structure	Variance Components
Estimation Method	REML
Residual Variance Method	Profile
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Containment

Class Level Information		
Class	**Levels**	**Values**
D	5	1 2 3 4 5
M	3	1 2 3
G	8	1 2 3 4 5 6 7 8

Dimensions	
Covariance Parameters	4
Columns in X	36
Columns in Z	60
Subjects	1
Max Obs per Subject	120

Number of Observations	
Number of Observations Read	120
Number of Observations Used	120
Number of Observations Not Used	0

Iteration History			
Iteration	**Evaluations**	**-2 Res Log Like**	**Criterion**
0	1	1220.44921467	
1	2	1203.93606235	0.00000110
2	1	1203.93548863	0.00000000

AOV - MIXED FACTOR LEVELS

The Mixed Procedure

Convergence criteria met.

Covariance Parameter Estimates				
Cov Parm	**Estimate**	**Alpha**	**Lower**	**Upper**
D	894.69	0.05	133.33	4.1733E8
D*M	2973.69	0.05	1082.22	23204
D*G	0	.	.	.
Residual	9132.04	0.05	6895.68	12671

Fit Statistics	
-2 Res Log Likelihood	1203.9
AIC (Smaller is Better)	1209.9
AICC (Smaller is Better)	1210.2
BIC (Smaller is Better)	1208.8

Solution for Random Effects											
Effect	**DENTIST**	**CONDENSATION METHOD**	**GOLD ALLOW**	**Estimate**	**Std Err Pred**	**DF**	**t Value**	**Pr > |t|**	**Alpha**	**Lower**	**Upper**
D	1			19.0865	24.7414	56	0.77	0.4437	0.05	-30.4765	68.6496
D	2			16.9812	24.7414	56	0.69	0.4953	0.05	-32.5819	66.5442
D	3			1.7994	24.7414	56	0.07	0.9423	0.05	-47.7636	51.3625
D	4			-14.4679	24.7414	56	-0.58	0.5611	0.05	-64.0309	35.0952
D	5			-23.3993	24.7414	56	-0.95	0.3483	0.05	-72.9623	26.1638
D*M	1	1		9.4940	38.4816	56	0.25	0.8060	0.05	-67.5939	86.5819
D*M	1	2		5.4835	38.4816	56	0.14	0.8872	0.05	-71.6044	82.5714
D*M	1	3		48.4609	38.4816	56	1.26	0.2131	0.05	-28.6270	125.55
D*M	2	1		14.3575	38.4816	56	0.37	0.7105	0.05	-62.7304	91.4453
D*M	2	2		-5.5505	38.4816	56	-0.14	0.8858	0.05	-82.6384	71.5374
D*M	2	3		47.6337	38.4816	56	1.24	0.2209	0.05	-29.4542	124.72
D*M	3	1		-36.0037	38.4816	56	-0.94	0.3535	0.05	-113.09	41.0842
D*M	3	2		-16.3487	38.4816	56	-0.42	0.6726	0.05	-93.4366	60.7392
D*M	3	3		58.3333	38.4816	56	1.52	0.1352	0.05	-18.7546	135.42
D*M	4	1		7.0042	38.4816	56	0.18	0.8562	0.05	-70.0837	84.0921
D*M	4	2		9.4972	38.4816	56	0.25	0.8060	0.05	-67.5907	86.5851

AOV - MIXED FACTOR LEVELS

The Mixed Procedure

Solution for Random Effects											
Effect	**DENTIST**	**CONDENSATION METHOD**	**GOLD ALLOW**	**Estimate**	**Std Err Pred**	**DF**	**t Value**	**Pr > |t|**	**Alpha**	**Lower**	**Upper**
D*M	4	3		-64.5886	38.4816	56	-1.68	0.0988	0.05	-141.68	12.4993
D*M	5	1		5.1481	38.4816	56	0.13	0.8941	0.05	-71.9398	82.2360
D*M	5	2		6.9185	38.4816	56	0.18	0.8580	0.05	-70.1694	84.0064
D*M	5	3		-89.8393	38.4816	56	-2.33	0.0232	0.05	-166.93	-12.7514
D*G	1		1	0	.	.	.	.	.	.	.
D*G	1		2	0	.	.	.	.	.	.	.
D*G	1		3	0	.	.	.	.	.	.	.
D*G	1		4	0	.	.	.	.	.	.	.
D*G	1		5	0	.	.	.	.	.	.	.
D*G	1		6	0	.	.	.	.	.	.	.
D*G	1		7	0	.	.	.	.	.	.	.
D*G	1		8	0	.	.	.	.	.	.	.
D*G	2		1	0	.	.	.	.	.	.	.
D*G	2		2	0	.	.	.	.	.	.	.
D*G	2		3	0	.	.	.	.	.	.	.
D*G	2		4	0	.	.	.	.	.	.	.
D*G	2		5	0	.	.	.	.	.	.	.
D*G	2		6	0	.	.	.	.	.	.	.
D*G	2		7	0	.	.	.	.	.	.	.
D*G	2		8	0	.	.	.	.	.	.	.
D*G	3		1	0	.	.	.	.	.	.	.
D*G	3		2	0	.	.	.	.	.	.	.
D*G	3		3	0	.	.	.	.	.	.	.
D*G	3		4	0	.	.	.	.	.	.	.
D*G	3		5	0	.	.	.	.	.	.	.
D*G	3		6	0	.	.	.	.	.	.	.
D*G	3		7	0	.	.	.	.	.	.	.
D*G	3		8	0	.	.	.	.	.	.	.
D*G	4		1	0	.	.	.	.	.	.	.
D*G	4		2	0	.	.	.	.	.	.	.
D*G	4		3	0	.	.	.	.	.	.	.
D*G	4		4	0	.	.	.	.	.	.	.

AOV - MIXED FACTOR LEVELS

The Mixed Procedure

Solution for Random Effects											
Effect	DENTIST	CONDENSATION METHOD	GOLD ALLOW	Estimate	Std Err Pred	DF	t Value	Pr > |t|	Alpha	Lower	Upper
D*G	4		5	0	.	.	.	.	.	.	.
D*G	4		6	0	.	.	.	.	.	.	.
D*G	4		7	0	.	.	.	.	.	.	.
D*G	4		8	0	.	.	.	.	.	.	.
D*G	5		1	0	.	.	.	.	.	.	.
D*G	5		2	0	.	.	.	.	.	.	.
D*G	5		3	0	.	.	.	.	.	.	.
D*G	5		4	0	.	.	.	.	.	.	.
D*G	5		5	0	.	.	.	.	.	.	.
D*G	5		6	0	.	.	.	.	.	.	.
D*G	5		7	0	.	.	.	.	.	.	.
D*G	5		8	0	.	.	.	.	.	.	.

Type 3 Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
M	2	8	9.08	0.0088
G	7	28	3.45	0.0087
M*G	14	56	1.64	0.0964

Least Squares Means										
Effect	CONDENSATION METHOD	GOLD ALLOW	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper
M	1		786.15	31.6540	8	24.84	<.0001	0.05	713.16	859.14
M	2		786.95	31.6540	8	24.86	<.0001	0.05	713.96	859.94
M	3		636.85	31.6540	8	20.12	<.0001	0.05	563.86	709.84
G		1	727.47	31.4004	28	23.17	<.0001	0.05	663.15	791.79
G		2	715.07	31.4004	28	22.77	<.0001	0.05	650.75	779.39
G		3	724.93	31.4004	28	23.09	<.0001	0.05	660.61	789.25
G		4	709.27	31.4004	28	22.59	<.0001	0.05	644.95	773.59
G		5	688.27	31.4004	28	21.92	<.0001	0.05	623.95	752.59
G		6	820.60	31.4004	28	26.13	<.0001	0.05	756.28	884.92

AOV - MIXED FACTOR LEVELS

The Mixed Procedure

Least Squares Means										
Effect	**CONDENSATION METHOD**	**GOLD ALLOW**	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**	**Alpha**	**Lower**	**Upper**
G		7	794.27	31.4004	28	25.29	<.0001	0.05	729.95	858.59
G		8	713.33	31.4004	28	22.72	<.0001	0.05	649.01	777.65

Differences of Least Squares Means									
Effect	**CONDENSATION METHOD**	**GOLD ALLOW**	**CONDENSATION METHOD**	**GOLD ALLOW**	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**
M	1		2		-0.8000	40.5719	8	-0.02	0.9848
M	1		3		149.30	40.5719	8	3.68	0.0062
M	2		3		150.10	40.5719	8	3.70	0.0060
G		1		2	12.4000	34.8942	28	0.36	0.7250
G		1		3	2.5333	34.8942	28	0.07	0.9426
G		1		4	18.2000	34.8942	28	0.52	0.6061
G		1		5	39.2000	34.8942	28	1.12	0.2708
G		1		6	-93.1333	34.8942	28	-2.67	0.0125
G		1		7	-66.8000	34.8942	28	-1.91	0.0658
G		1		8	14.1333	34.8942	28	0.41	0.6885
G		2		3	-9.8667	34.8942	28	-0.28	0.7794
G		2		4	5.8000	34.8942	28	0.17	0.8692
G		2		5	26.8000	34.8942	28	0.77	0.4489
G		2		6	-105.53	34.8942	28	-3.02	0.0053
G		2		7	-79.2000	34.8942	28	-2.27	0.0311
G		2		8	1.7333	34.8942	28	0.05	0.9607
G		3		4	15.6667	34.8942	28	0.45	0.6569
G		3		5	36.6667	34.8942	28	1.05	0.3023
G		3		6	-95.6667	34.8942	28	-2.74	0.0105
G		3		7	-69.3333	34.8942	28	-1.99	0.0568
G		3		8	11.6000	34.8942	28	0.33	0.7420
G		4		5	21.0000	34.8942	28	0.60	0.5521
G		4		6	-111.33	34.8942	28	-3.19	0.0035
G		4		7	-85.0000	34.8942	28	-2.44	0.0215
G		4		8	-4.0667	34.8942	28	-0.12	0.9081
G		5		6	-132.33	34.8942	28	-3.79	0.0007

AOV - MIXED FACTOR LEVELS

The Mixed Procedure

Differences of Least Squares Means									
Effect	**CONDENSATION METHOD**	**GOLD ALLOW**	**CONDENSATION METHOD**	**GOLD ALLOW**	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**
G		5		7	-106.00	34.8942	28	-3.04	0.0051
G		5		8	-25.0667	34.8942	28	-0.72	0.4785
G		6		7	26.3333	34.8942	28	0.75	0.4568
G		6		8	107.27	34.8942	28	3.07	0.0047
G		7		8	80.9333	34.8942	28	2.32	0.0279

Differences of Least Squares Means									
Effect	**CONDENSATION METHOD**	**GOLD ALLOW**	**CONDENSATION METHOD**	**GOLD ALLOW**	**Adjustment**	**Adj P**	**Alpha**	**Lower**	**Upper**
M	1		2		Tukey-Kramer	0.9998	0.05	-94.3589	92.7589
M	1		3		Tukey-Kramer	0.0153	0.05	55.7411	242.86
M	2		3		Tukey-Kramer	0.0148	0.05	56.5411	243.66
G		1		2	Tukey-Kramer	1.0000	0.05	-59.0775	83.8775
G		1		3	Tukey-Kramer	1.0000	0.05	-68.9442	74.0108
G		1		4	Tukey-Kramer	0.9994	0.05	-53.2775	89.6775
G		1		5	Tukey-Kramer	0.9460	0.05	-32.2775	110.68
G		1		6	Tukey-Kramer	0.1742	0.05	-164.61	-21.6558
G		1		7	Tukey-Kramer	0.5530	0.05	-138.28	4.6775
G		1		8	Tukey-Kramer	0.9999	0.05	-57.3442	85.6108
G		2		3	Tukey-Kramer	1.0000	0.05	-81.3442	61.6108
G		2		4	Tukey-Kramer	1.0000	0.05	-65.6775	77.2775
G		2		5	Tukey-Kramer	0.9935	0.05	-44.6775	98.2775
G		2		6	Tukey-Kramer	0.0856	0.05	-177.01	-34.0558
G		2		7	Tukey-Kramer	0.3441	0.05	-150.68	-7.7225
G		2		8	Tukey-Kramer	1.0000	0.05	-69.7442	73.2108
G		3		4	Tukey-Kramer	0.9998	0.05	-55.8108	87.1442
G		3		5	Tukey-Kramer	0.9617	0.05	-34.8108	108.14
G		3		6	Tukey-Kramer	0.1517	0.05	-167.14	-24.1892
G		3		7	Tukey-Kramer	0.5076	0.05	-140.81	2.1442
G		3		8	Tukey-Kramer	1.0000	0.05	-59.8775	83.0775
G		4		5	Tukey-Kramer	0.9986	0.05	-50.4775	92.4775
G		4		6	Tukey-Kramer	0.0597	0.05	-182.81	-39.8558
G		4		7	Tukey-Kramer	0.2637	0.05	-156.48	-13.5225

AOV - MIXED FACTOR LEVELS

The Mixed Procedure

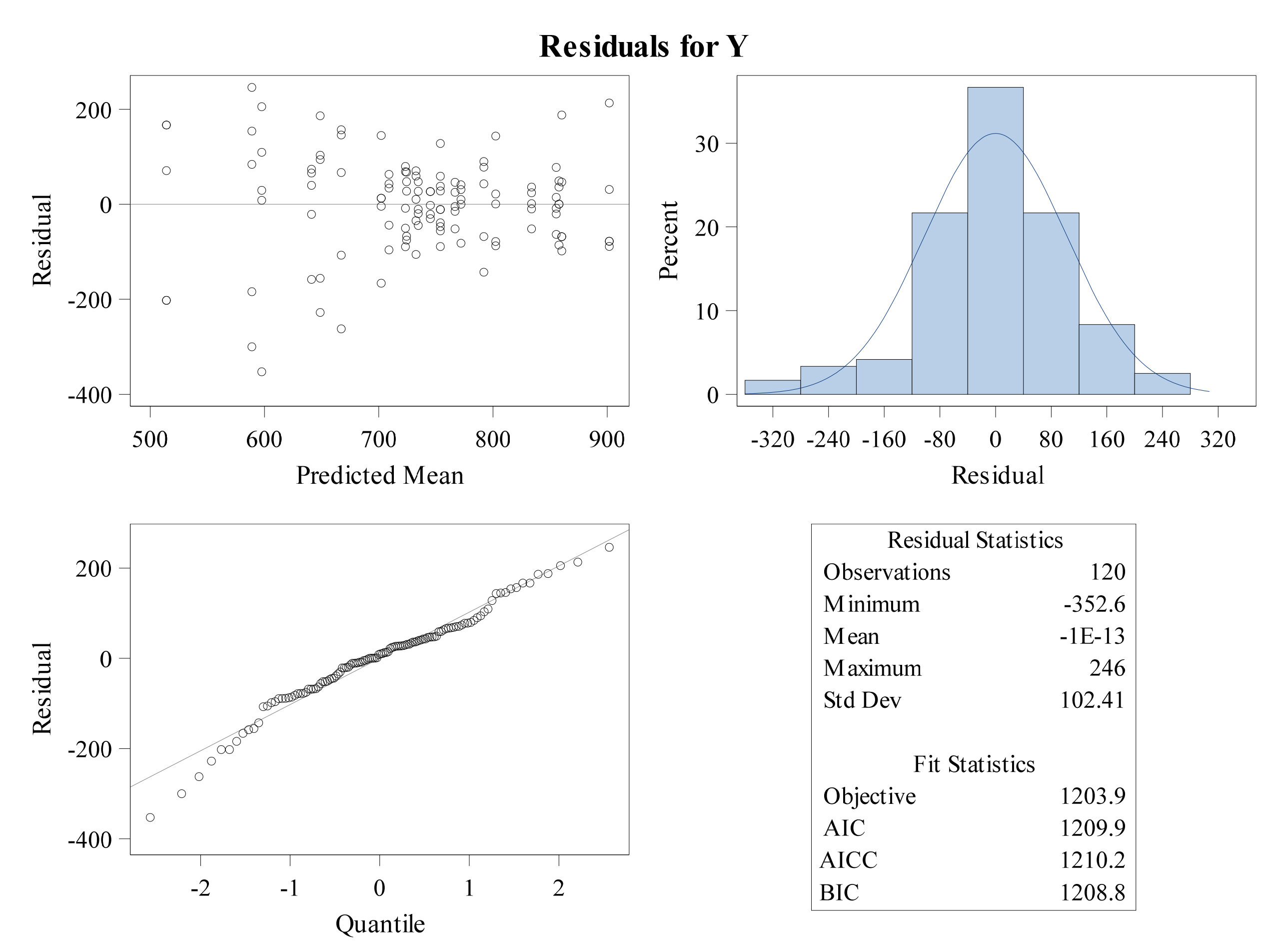
Differences of Least Squares Means									
Effect	**CONDENSATION METHOD**	**GOLD ALLOW**	**CONDENSATION METHOD**	**GOLD ALLOW**	**Adjustment**	**Adj P**	**Alpha**	**Lower**	**Upper**
G		4		8	Tukey-Kramer	1.0000	0.05	-75.5442	67.4108
G		5		6	Tukey-Kramer	0.0146	0.05	-203.81	-60.8558
G		5		7	Tukey-Kramer	0.0832	0.05	-177.48	-34.5225
G		5		8	Tukey-Kramer	0.9957	0.05	-96.5442	46.4108
G		6		7	Tukey-Kramer	0.9942	0.05	-45.1442	97.8108
G		6		8	Tukey-Kramer	0.0770	0.05	35.7892	178.74
G		7		8	Tukey-Kramer	0.3186	0.05	9.4558	152.41

Differences of Least Squares Means						
Effect	**CONDENSATION METHOD**	**GOLD ALLOW**	**CONDENSATION METHOD**	**GOLD ALLOW**	**Adj Lower**	**Adj Upper**
M	1		2		-116.73	115.13
M	1		3		33.3706	265.23
M	2		3		34.1706	266.03
G		1		2	-101.71	126.51
G		1		3	-111.58	116.65
G		1		4	-95.9117	132.31
G		1		5	-74.9117	153.31
G		1		6	-207.25	20.9784
G		1		7	-180.91	47.3117
G		1		8	-99.9784	128.25
G		2		3	-123.98	104.25
G		2		4	-108.31	119.91
G		2		5	-87.3117	140.91
G		2		6	-219.65	8.5784
G		2		7	-193.31	34.9117
G		2		8	-112.38	115.85
G		3		4	-98.4450	129.78
G		3		5	-77.4450	150.78
G		3		6	-209.78	18.4450
G		3		7	-183.45	44.7784
G		3		8	-102.51	125.71
G		4		5	-93.1117	135.11

AOV - MIXED FACTOR LEVELS

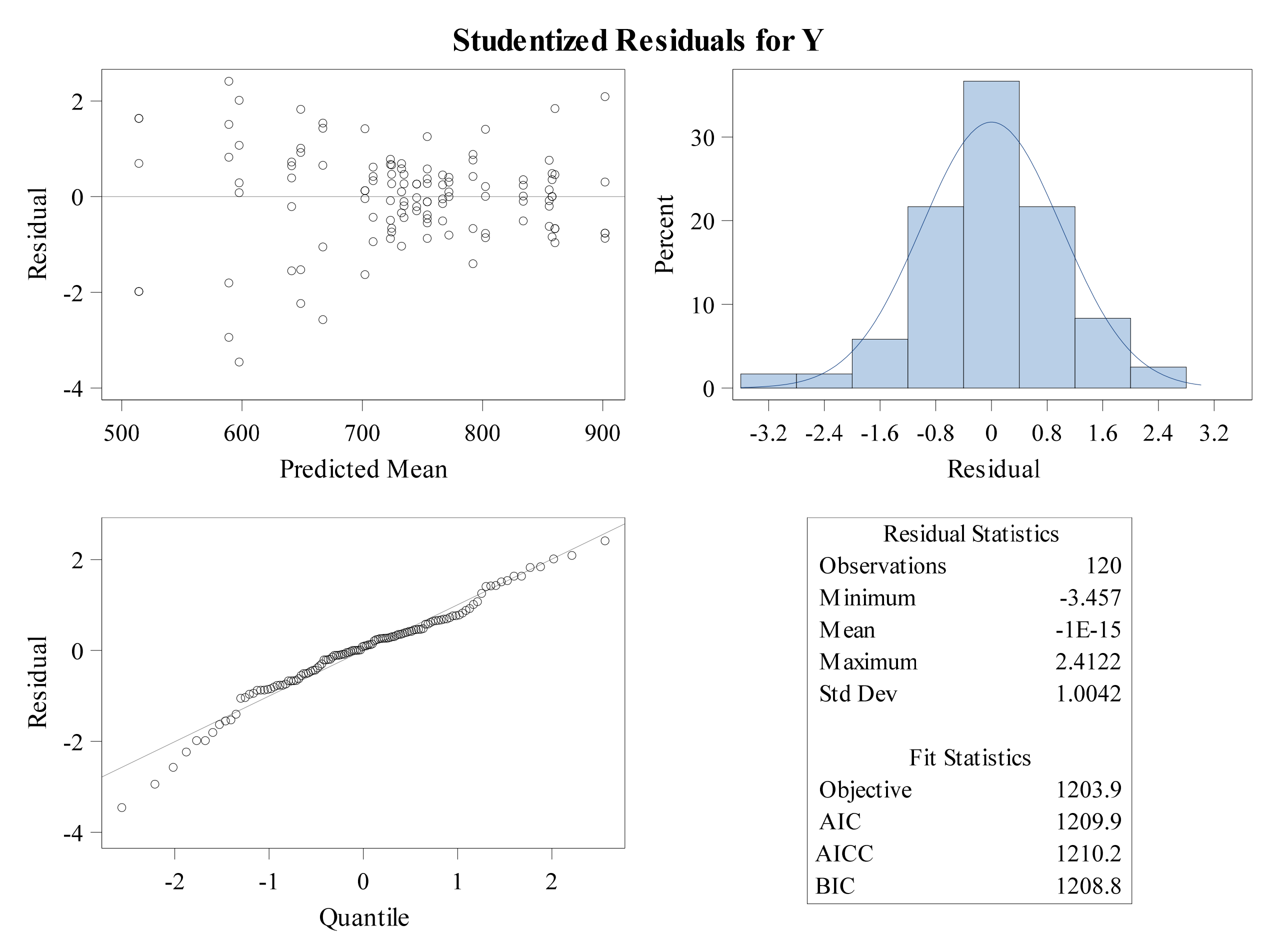
The Mixed Procedure

Differences of Least Squares Means						
Effect	**CONDENSATION METHOD**	**GOLD ALLOW**	**CONDENSATION METHOD**	**GOLD ALLOW**	**Adj Lower**	**Adj Upper**
G		4		6	-225.45	2.7784
G		4		7	-199.11	29.1117
G		4		8	-118.18	110.05
G		5		6	-246.45	-18.2216
G		5		7	-220.11	8.1117
G		5		8	-139.18	89.0450
G		6		7	-87.7784	140.45
G		6		8	-6.8450	221.38
G		7		8	-33.1784	195.05



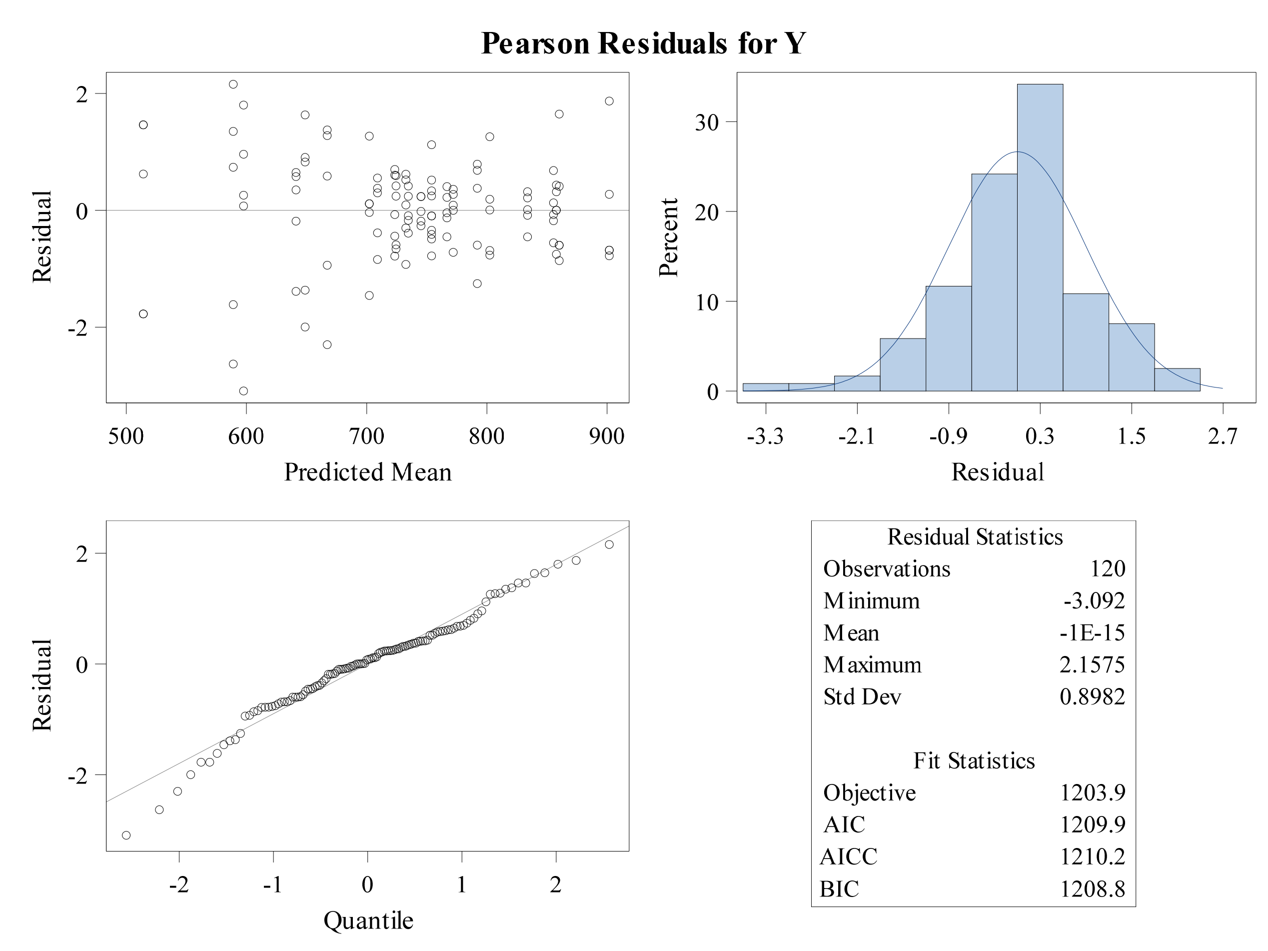
AOV - MIXED FACTOR LEVELS

The Mixed Procedure



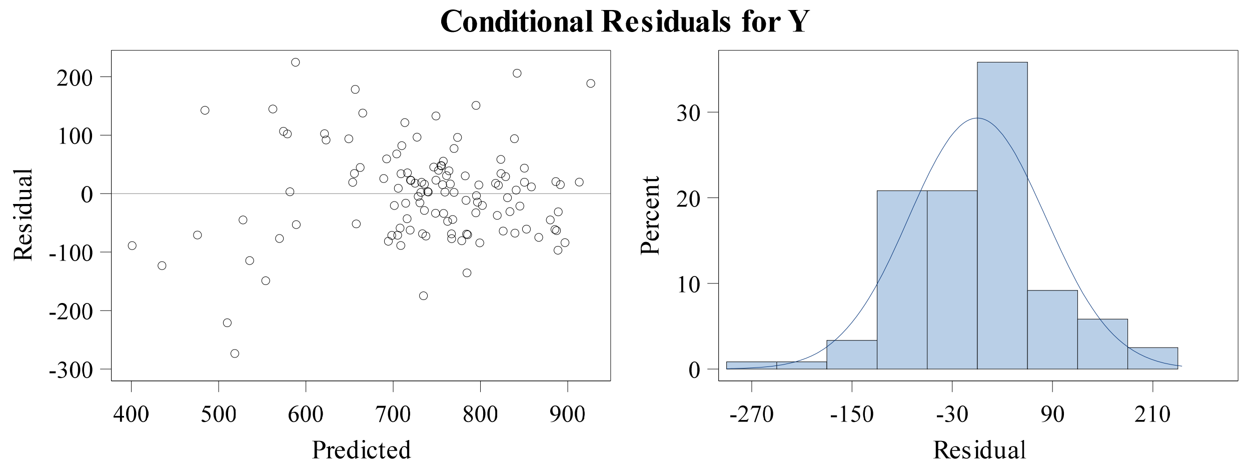
AOV - MIXED FACTOR LEVELS

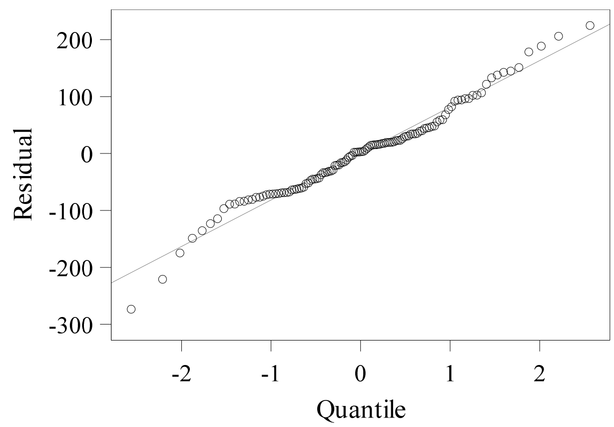
The Mixed Procedure



AOV - MIXED FACTOR LEVELS

The Mixed Procedure



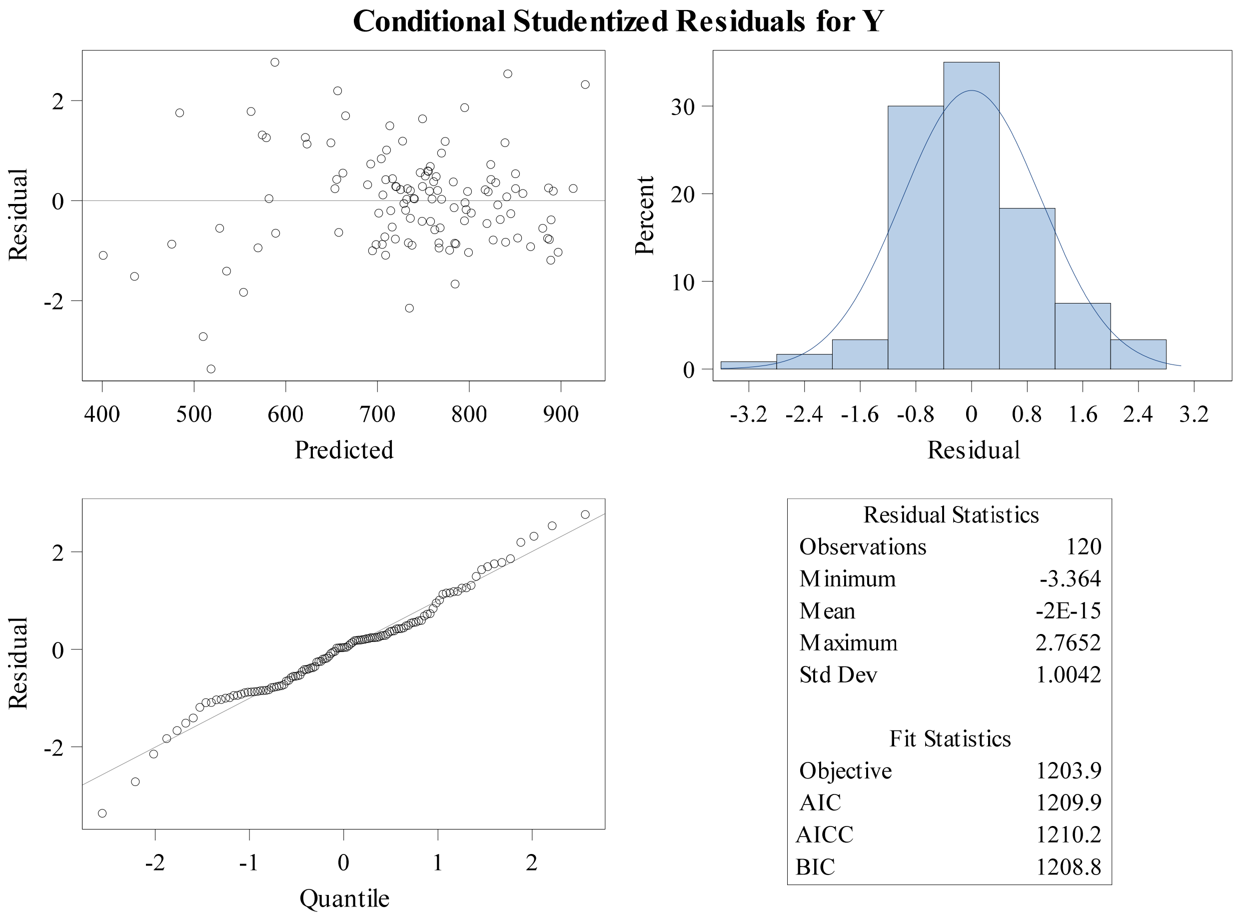


Residual Statistics	
Observations	120
Minimum	-273.5
Mean	-1E-13
Maximum	224.86
Std Dev	81.657

Fit Statistics	
Objective	1203.9
AIC	1209.9
AICC	1210.2
BIC	1208.8

AOV - MIXED FACTOR LEVELS

The Mixed Procedure

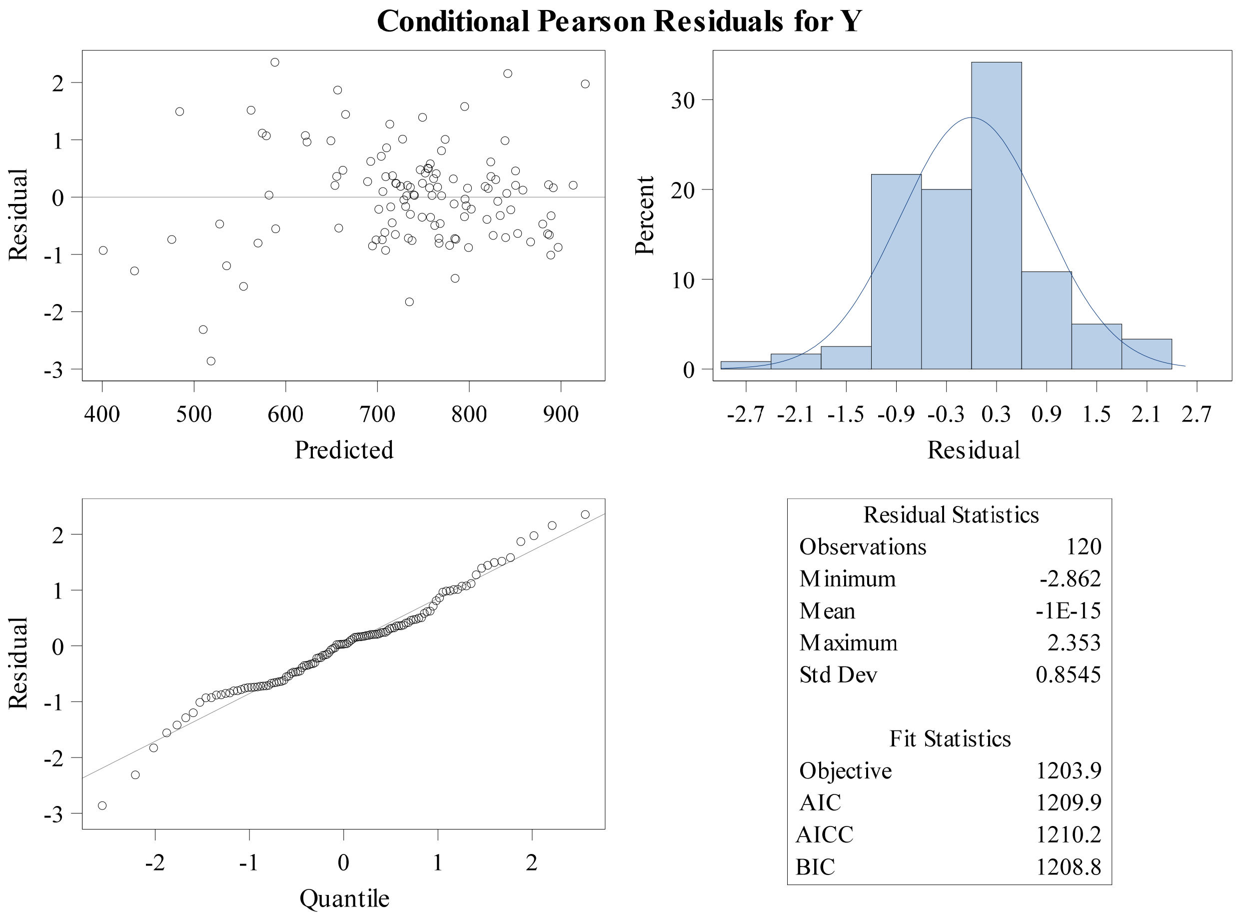


Residual Statistics	
Observations	120
Minimum	-3.364
Mean	-2E-15
Maximum	2.7652
Std Dev	1.0042

Fit Statistics	
Objective	1203.9
AIC	1209.9
AICC	1210.2
BIC	1208.8

AOV - MIXED FACTOR LEVELS

The Mixed Procedure



AOV - MIXED FACTOR LEVELS

The GLIMMIX Procedure

Model Information	
Data Set	WORK.RAW
Response Variable	Y
Response Distribution	Gaussian
Link Function	Identity
Variance Function	Default
Variance Matrix	Diagonal
Estimation Technique	Restricted Maximum Likelihood
Degrees of Freedom Method	Residual

Class Level Information		
Class	**Levels**	**Values**
D	5	1 2 3 4 5
M	3	1 2 3
G	8	1 2 3 4 5 6 7 8

Number of Observations Read	120
Number of Observations Used	120

Dimensions	
Covariance Parameters	1
Columns in X	36
Columns in Z	0
Subjects (Blocks in V)	1
Max Obs per Subject	120

Optimization Information	
Optimization Technique	None
Parameters	25
Lower Boundaries	1
Upper Boundaries	0
Fixed Effects	Not Profiled

AOV - MIXED FACTOR LEVELS

The GLIMMIX Procedure

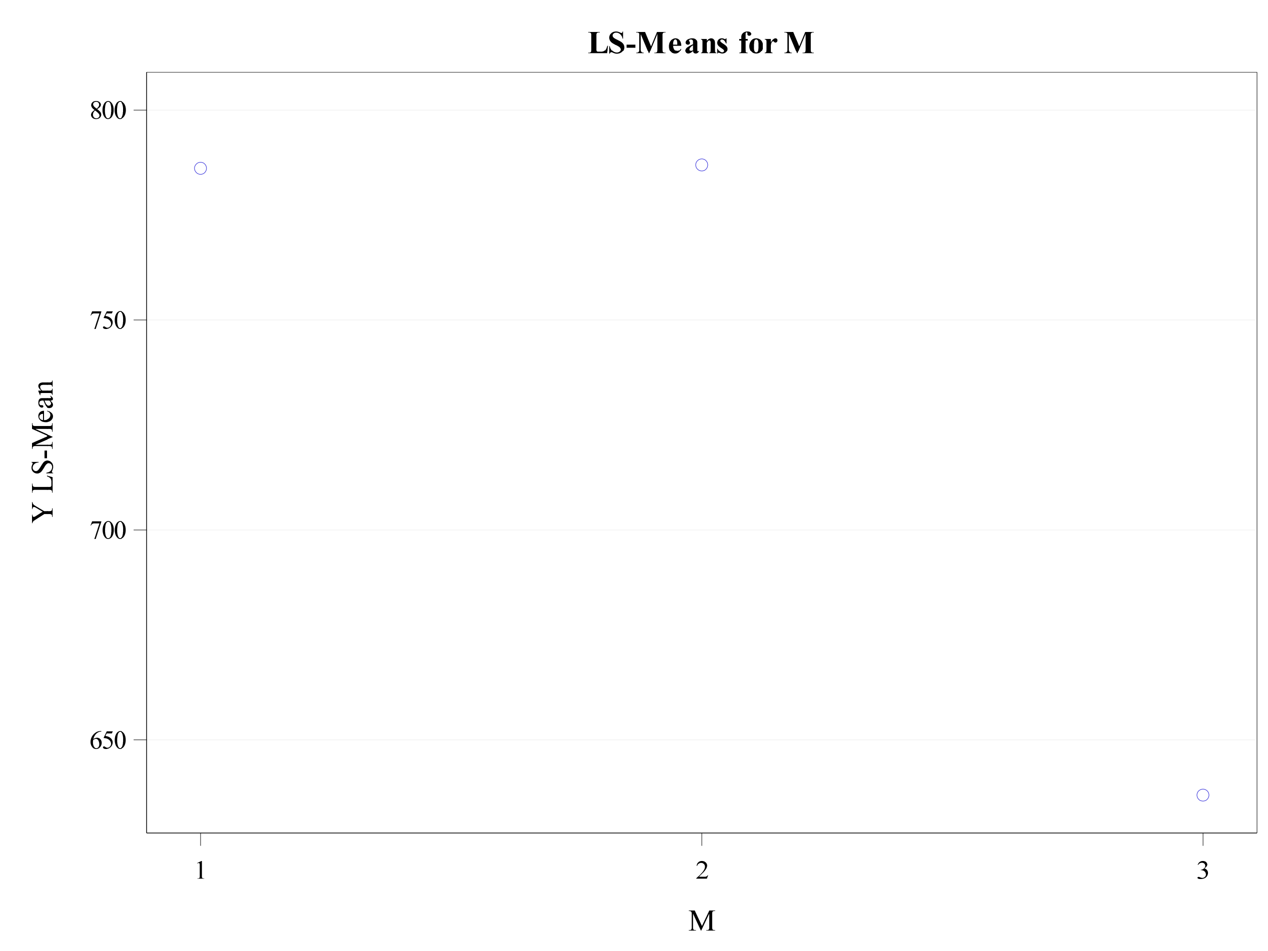
Fit Statistics	
-2 Res Log Likelihood	1220.45
AIC (smaller is better)	1270.45
AICC (smaller is better)	1289.02
BIC (smaller is better)	1334.56
CAIC (smaller is better)	1359.56
HQIC (smaller is better)	1296.36
Pearson Chi-Square	1248089
Pearson Chi-Square / DF	13000.93

Type III Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
M	2	96	22.98	<.0001
G	7	96	2.42	0.0251
M*G	14	96	1.15	0.3244

M Least Squares Means					
CONDENSATION METHOD	Estimate	Standard Error	DF	t Value	Pr > |t|
1	786.15	18.0284	96	43.61	<.0001
2	786.95	18.0284	96	43.65	<.0001
3	636.85	18.0284	96	35.32	<.0001

AOV - MIXED FACTOR LEVELS

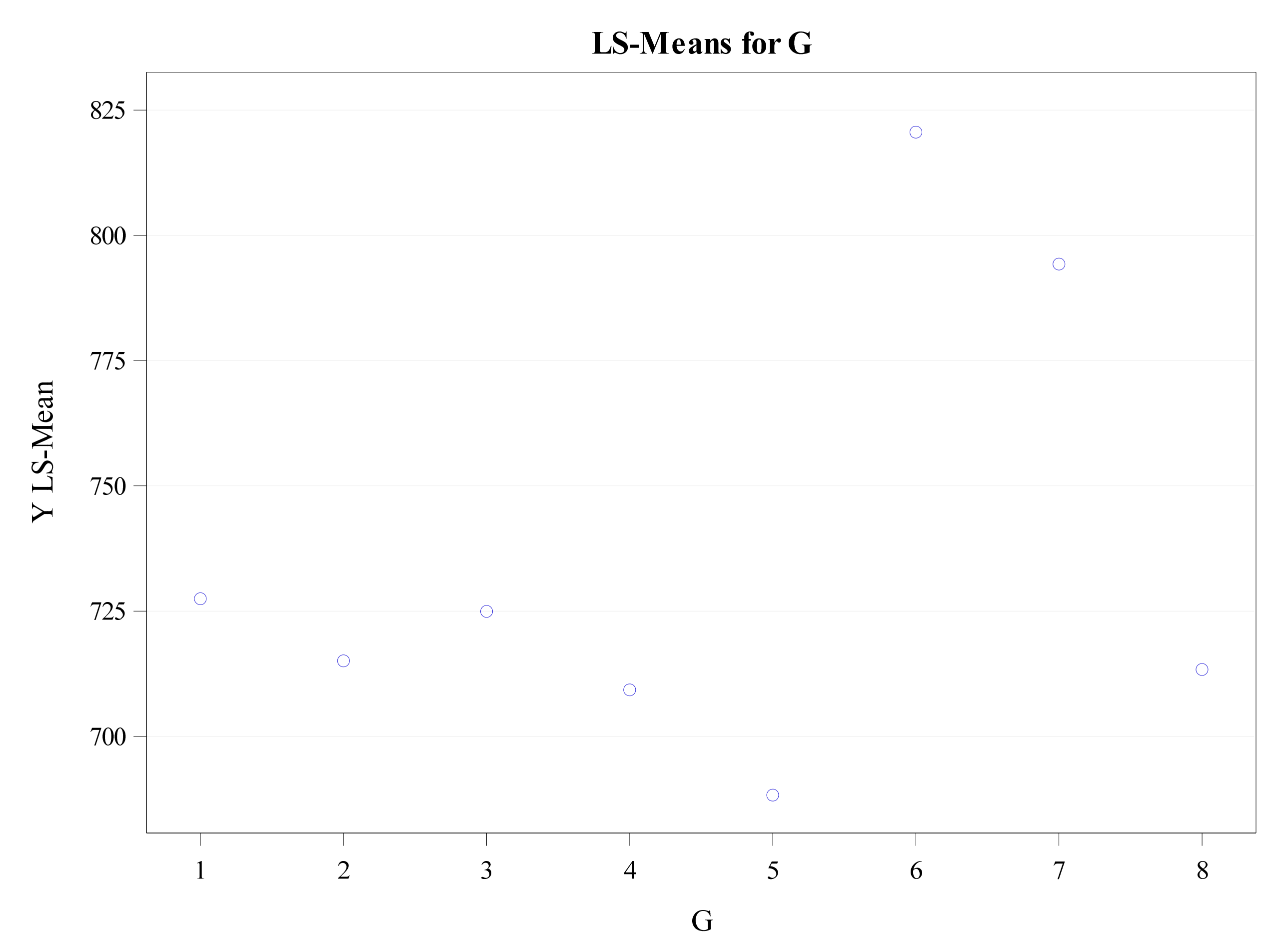
The GLIMMIX Procedure



G Least Squares Means					
GOLD ALLOW	Estimate	Standard Error	DF	t Value	Pr > |t|
1	727.47	29.4403	96	24.71	<.0001
2	715.07	29.4403	96	24.29	<.0001
3	724.93	29.4403	96	24.62	<.0001
4	709.27	29.4403	96	24.09	<.0001
5	688.27	29.4403	96	23.38	<.0001
6	820.60	29.4403	96	27.87	<.0001
7	794.27	29.4403	96	26.98	<.0001
8	713.33	29.4403	96	24.23	<.0001

AOV - MIXED FACTOR LEVELS

The GLIMMIX Procedure

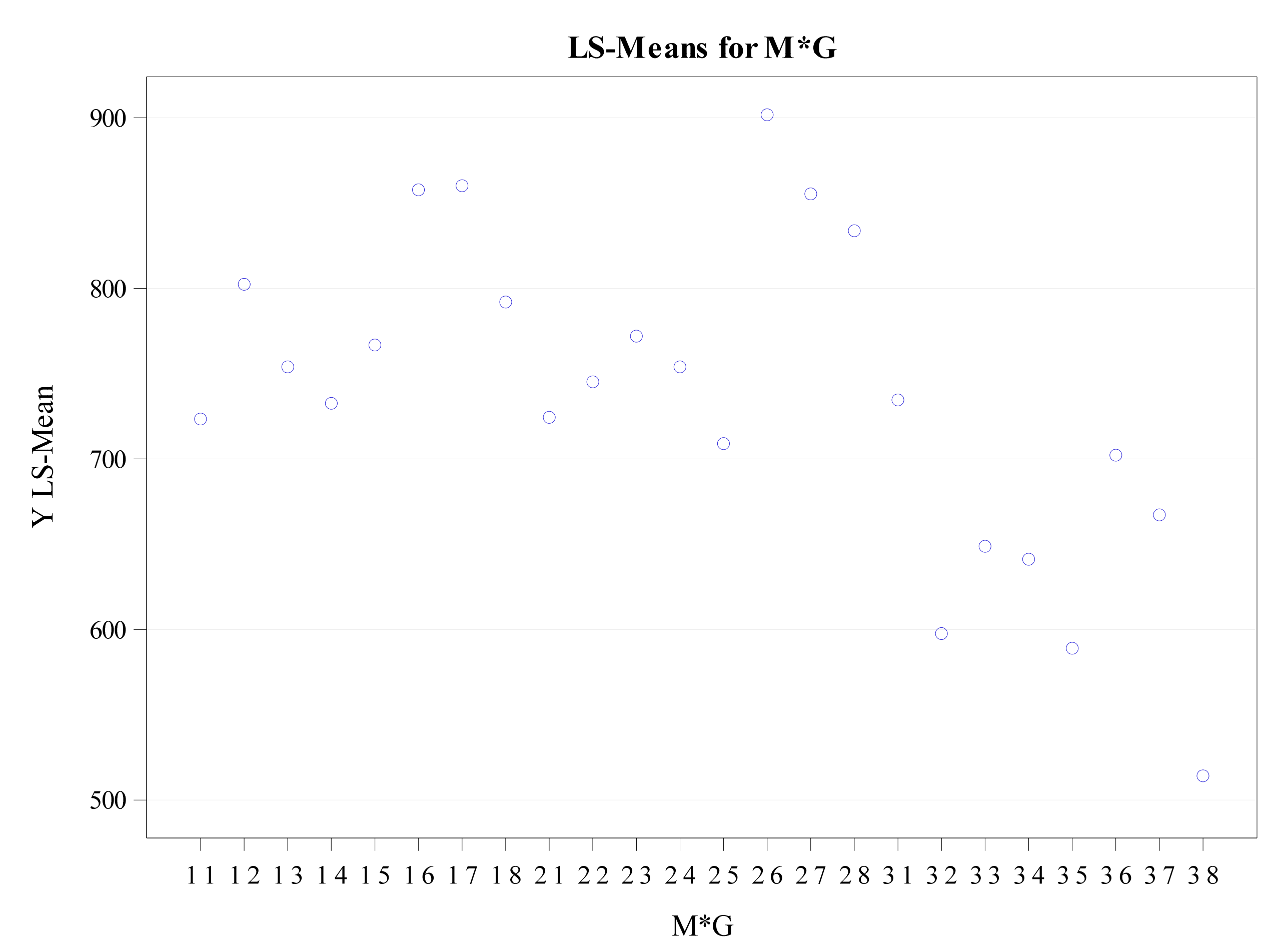


M*G Least Squares Means						
CONDENSATION METHOD	GOLD ALLOW	Estimate	Standard Error	DF	t Value	Pr > |t|
1	1	723.40	50.9920	96	14.19	<.0001
1	2	802.40	50.9920	96	15.74	<.0001
1	3	754.00	50.9920	96	14.79	<.0001
1	4	732.60	50.9920	96	14.37	<.0001
1	5	766.80	50.9920	96	15.04	<.0001
1	6	857.80	50.9920	96	16.82	<.0001
1	7	860.20	50.9920	96	16.87	<.0001
1	8	792.00	50.9920	96	15.53	<.0001
2	1	724.40	50.9920	96	14.21	<.0001
2	2	745.20	50.9920	96	14.61	<.0001
2	3	772.00	50.9920	96	15.14	<.0001
2	4	754.00	50.9920	96	14.79	<.0001

AOV - MIXED FACTOR LEVELS

The GLIMMIX Procedure

M*G Least Squares Means						
CONDENSATION METHOD	GOLD ALLOW	Estimate	Standard Error	DF	t Value	Pr > |t|
2	5	709.00	50.9920	96	13.90	<.0001
2	6	901.80	50.9920	96	17.69	<.0001
2	7	855.40	50.9920	96	16.78	<.0001
2	8	833.80	50.9920	96	16.35	<.0001
3	1	734.60	50.9920	96	14.41	<.0001
3	2	597.60	50.9920	96	11.72	<.0001
3	3	648.80	50.9920	96	12.72	<.0001
3	4	641.20	50.9920	96	12.57	<.0001
3	5	589.00	50.9920	96	11.55	<.0001
3	6	702.20	50.9920	96	13.77	<.0001
3	7	667.20	50.9920	96	13.08	<.0001
3	8	514.20	50.9920	96	10.08	<.0001



Nested Treatment Factors

Experiments in which the treatments involve nested factors arise in certain factorial arrangements of the levels of the factors or in experiments in which the randomization of Experimental Units to the treatments are restrictive. In many cases, these restrictions on the randomization are dictated by technical considerations of the implementation of the treatment assignments or methods by which the responses are obtained. For example, in an experiment in which the optimal amount of fertilizer to be applied for several different varieties of wheat is being explored, it may be possible to hand plant the different varieties in small rows in order to achieve a very high degree of precision, whereas the application of the fertilizer must be done using a machinery that is only appropriate for a large field.

Definition: Two factors F_1 with a levels and F_2 with b levels are said to be **crossed** if the physical properties of the b levels of F_2 remain the same for all levels of F_1.

Example: An experiment is designed to study the factors which affect the time to fabricate an automobile part from a specimen of metal. Three factors are identified:

1. F_1- Type of Alloy: A1, A2, A3
2. F_2- Porosity of Alloy: L, M, H
3. F_3- Amount of Lubricant Used in Cutting Machine: L1, L2, L3, L4

A complete $3 \times 3 \times 4$ factorial experiment with all 3 factors crossed would have the levels of F_2 exactly the same for all levels of both factors F_1 and F_3. Also, the levels of F_3 would be the same for all levels of F_1 and F_2. That is, the three Porosity levels would be the same for all three types of Alloys.

Definition: Factor F_2 is said to be **nested within** the levels of factor F_1 if the physical properties of the levels of factor F_2 vary depending on which level of factor F_1 is used.

Example: For many Alloys, the possible porosity levels may vary. For example, suppose that the three possible porosity levels are given as follows:

	Porosity Level		
Alloy Type	Low	Medium	High
A1	.2	.3	.4
A2	.3	.5	.7
A3	.5	.8	.95

Thus, the definition of L, M, H for the factor Porosity varies depending of the level of the factor Alloy. Thus, factor P, porosity is Nested Within the levels of factor A, alloy type. P nested within A is denoted as P(A).

Example: A company is designing a new product. There are 3 producers of raw material. The goal of a study was to evaluate the consistency of the raw material's physical properties across multiply batches of material from each producer and then variation within each batch. The experiment was designed to randomly select 6 batches of material from each of the 3 producers. From each of the 18 batches, 4 samples are taken and the physical properties are determined in the company's lab. The factors are

F_1- Producer: P1, P2, P3 F_1 - fixed

F_2- Batch: B1, B2, B3, B4, B5, B6 $F_2(F_1)$ - Random

F_3- Sample: S1, S2, S3, S4 $F_3(F_2, F_1)$ — Random

The levels of Batch only having meaning if we know from which Producer the batches are selected. Thus, the factor Batch is nested within the levels of factor Producer: B(P). Similarly, the levels of Sample only having meaning if we know from which producer and batch the sample is selected. Thus, the factor Sample is nested within both the levels of factor Producer and the factor Batch: S(B,P).

F_1

$F_2(F_1)$

Example: A social scientist is studying the use of drugs by high school students. The factors of interest are Type of High School (private-nonreligious, private-religious,public); Particular Schools within each Type of School; Individual Students from each School with each Type of School.

F_1: School Type (Fixed)
$F_2(F_1)$: school (Random)
$F_3(F_2, F_1)$: students (Random) → measured by residual var

Example: An animal scientist is studying the level of infestation of ticks on cattle. There are four breeds of cattle; five cows of each Breed; measurements are taken at four locations on each cow: head, neck, flank, and tail.

F_1 = breed (fixed)

F_2 = cows (random) $F_2(F_1)$

F_3 = location $F_3(F_1, F_2)$ (fixed)

→ will be specified in model b/c fixed.

Construction of Hierarchically Nested Designs

Suppose we have factor F_1 having fixed levels, F_2 nested within the levels of F_1 having random levels, and F_3 nested within the levels of Factor F_2 having random levels.

1. List the factors to be included in the experiment:

   F_1, F_2, F_3

2. Determine the hierarchy of the factors:

   F_2 nested within the levels of F_1

   F_3 nested within the levels of F_2

3. Determine whether factors are fixed or random

   F_1 - Fixed; F_2 - Random F_3 - Random

4. Randomly select, if possible, an equal number of levels for each factor within the levels of the factor for which it is nested

   Randomly select b levels of F_2 within each of the a levels of F_1

   Randomly select c levels of F_3 within each of the b levels of F_2

5. Randomize the run order or the assignment of the factor-level combinations of the EU's

Example The following example from Snedecor and Cochran's book, *Statistical Methods*, describes an experiment in which the variability of calcium concentration in turnip green leaves is investigated. Four turnip green plants were randomly selected. From each plant, three leaves were randomly selected. Each leaf was then processed into two 100-mg samples from which a determination of the calcium content was made. Thus, we have three factors: Plant (P) with four random levels, Leaf (L) with three random levels within each level of P, and Determination (D) with two random levels within each level of L. The data is given below:

Plants – Random
Leaf (Plant) – Random
Determination (Plant, Leaf) – Rand.

		Determination				
Plant	Leaf	1	2	$\bar{y}_{ij.}$	$\bar{y}_{i..}$	$\bar{y}_{...}$
1	1	3.28	3.09	3.185		
	2	3.52	3.48	3.500		
	3	2.88	2.80	2.845	3.175	
2	1	2.46	2.44	2.450		
	2	1.87	1.92	1.895		
	3	2.19	2.19	2.190	2.178	
3	1	2.77	2.66	2.715		
	2	3.74	3.44	3.590		
	3	2.55	2.55	2.550	2.952	
4	1	3.78	3.87	3.825		
	2	4.07	4.12	4.095		
	3	3.31	3.31	3.310	3.743	3.012

The formulas for the sum of squares decomposition in a nested treatment structure with three hierarchically nested factors with levels $F_1 : i = 1, \ldots, a;\quad F_2(F_1) : j = 1, \ldots, r_i;\quad F_3(F_1, F_2) : k = 1, \ldots, n_{ij}$ (In many cases, $F_3(F_1, F_2)$ is the error term.) are given here:

1. Total SS: $SS_{TOT} = \sum_{i=1}^{a} \sum_{j=1}^{r_i} \sum_{k=1}^{n_{ij}} (y_{ijk} - \bar{y}_{...})^2$

2. SS due to F_1: $SS_{F_1} = \sum_{i=1}^{a} \sum_{j=1}^{r_i} \sum_{k=1}^{n_{ij}} (\bar{y}_{i..} - \bar{y}_{...})^2 = \sum_{i=1}^{a} n_{i.}(\bar{y}_{i..} - \bar{y}_{...})^2$

   with $df = (a-1)$

3. SS due to F_2 nested within F_1: $SS_{F_2(F_1)} = \sum_{i=1}^{a} \sum_{j=1}^{r_i} n_{ij}(\bar{y}_{ij.} - \bar{y}_{i..})^2$

   with $df = \sum_{i=1}^{a}(r_i - 1)$

4. SS due to F_3 nested within $F_2(F_1)$: $SS_{F_3(F_1,F_2)} = \sum_{i=1}^{a} \sum_{j=1}^{r_i} \sum_{k=1}^{n_{ij}} (y_{ijk} - \bar{y}_{ij.})^2$

   with $df = \sum_{i=1}^{a} \sum_{j=1}^{r_i} (n_{ij} - 1)$

STOP friday 4/8/22 (Week 11, lecture 30)

For our example, we have the following calculations: ($a = 4, r_i = r = 3, n_{ij} = n = 2$)

$$\begin{aligned}
SS_{TOT} &= \sum_{i=1}^{a}\sum_{j=1}^{r}\sum_{k=1}^{n}(y_{ijk} - \bar{y}_{...})^2 \\
&= \sum_{i=1}^{4}\sum_{j=1}^{3}\sum_{k=1}^{2}(y_{ijk} - 3.012)^2 \\
&= [(3.28 - 3.012)^2 + (3.09 - 3.012)^2 + (3.52 - 3.012)^2 + \cdots + (3.31 - 3.012)^2] \\
&= 10.27035 \quad \text{with} \quad df = 24 - 1 = 23
\end{aligned}$$

$$\begin{aligned}
SS_{F_1} &= \sum_{i=1}^{a} n(\bar{y}_{i..} - \bar{y}_{...})^2 \\
&= (3)(2)[(3.175 - 3.012)^2 + (2.178 - 3.012)^2 + (2.952 - 3.012)^2 + (3.743 - 3.012)^2] \\
&= 7.5603 \quad \text{with} \quad df = 4 - 1 = 3
\end{aligned}$$

$$\begin{aligned}
SS_{F_2(F_1)} &= \sum_{i=1}^{a}\sum_{j=1}^{r} n(\bar{y}_{ij.} - \bar{y}_{i..})^2 \\
&= 2[(3.185 - 3.175)^2 + (3.50 - 3.175)^2 + (2.845 - 3.175)^2 \\
&\quad +(2.45 - 2.178)^2 + (1.895 - 2.178)^2 + (2.19 - 2.178)^2 \\
&\quad +(2.715 - 2.952)^2 + (3.59 - 2.952)^2 + (2.55 - 2.952)^2 \\
&\quad +(3.825 - 3.743)^2 + (4.095 - 3.743)^2 + (3.31 - 3.743)^2] \\
&= 2.6302 \quad \text{with} \quad df = 4(3 - 1) = 8
\end{aligned}$$

$$\begin{aligned}
SS_{F_3(F_1,F_2)} &= \sum_{i=1}^{a}\sum_{j=1}^{r}\sum_{k=1}^{n}(y_{ijk} - \bar{y}_{ij.})^2 \\
&= [(3.28 - 3.185)^2 + (3.09 - 3.185)^2 + (3.52 - 3.50)^2 + (3.48 - 3.50)^2 \\
&\quad +(2.88 - 2.84)^2 + (2.80 - 2.84)^2 + \cdots + (4.07 - 4.095)^2 + (4.12 - 4.095)^2 \\
&\quad +(3.31 - 3.31)^2 + (3.31 - 3.31)^2] \\
&= 0.07985 \quad \text{with} \quad df = (4)(3)(2 - 1) = 12
\end{aligned}$$

Which tests of hypotheses and multiply comparisons to conduct depend on whether the factors are fixed or random.

START Monday 4/11/22 (Week 12, Lecture 31)

F_1 - Fixed.
$F_2(F_1)$ - Fixed
$F_3(F_2, F_1)$ - Random

Case 1: F_1-Fixed Levels, F_2 fixed levels within the levels of F_1, F_3 random levels within the levels of F_2

Model: $y_{ijk} = \mu + \tau_i + \beta_{j(i)} + c_{k(ij)} \quad i = 1, \ldots, a; \quad j = 1, \ldots, b; \quad k = 1, \ldots, n_{ij}$
with model conditions:

1. $\tau_a = 0; \quad \beta_{r_i(i)} = 0 \quad$ for all $\quad i = 1, \ldots, a$
2. $c_{j(i,j)}$'s are iid $N(0, \sigma^2_{F_3(F_1,F_2)})$

Expected Mean Squares are given in the following table for the balanced case ($r_i = b, n_{ij} = n$):

Source	df	Expected Mean Squares
F_1	$a-1$	$nbQ_{F_1} + \sigma^2_{F_3(F_1,F_2)}$
$F_2(F_1)$	$a(b-1)$	$nQ_{F_2(F_1)} + \sigma^2_{F_3(F_1,F_2)}$
$F_3(F_1,F_2)$	$ab(n-1)$	$\sigma^2_{F_3(F_1,F_2)}$
Total	abn-1	

From the above table, we can determine the appropriate test statistics:

1. To Test no differences across the levels of Factor F_1:

   $H_o : \mu_1 = \mu_2 = \cdots = \mu_a$

   $F = \frac{MS_{F_1}}{MS_{F_3(F_1,F_2)}}$

2. If H_o is rejected, then construct contrasts and multiple comparisons across the levels of F_1

3. To Test no differences across the levels of Factor F_2 nested within the levels of F_1:

   $H_o : \mu_{1(i)} = \mu_{2(i)} = \cdots = \mu_{b(i)}$ for all $i = 1, \ldots, a$

   $F = \frac{MS_{F_2(F_1)}}{MS_{F_3(F_1,F_2)}}$

4. If H_o is rejected, then construct contrasts and multiple comparisons across the levels of F_2 separately for each level of F_1

Case 2: F_1-Fixed Levels, F_2 random levels within the levels of F_1,
F_3 random levels within the levels of F_2

Model: $y_{ijk} = \mu + \tau_i + b_{j(i)} + c_{k(ij)} \quad i = 1, \ldots, a; \quad j = 1, \ldots, b; \quad k = 1, \ldots, n_{ij}$
with model conditions:

1. $\tau_a = 0$

2. $b_{j(i)}$ are iid $N(0, \sigma^2_{F_2(F_1)})$ r.v.'s independent of $c_{j(i,j)}$'s which are iid $N(0, \sigma^2_{F_3(F_1,F_2)})$ r.v.'s

Expected Mean Squares are given in the following table for the balanced case ($r_i = b, n_{ij} = n$):

Source	df	Expected Mean Squares
F_1	$a-1$	$nbQ_{F_1} + n\sigma^2_{F_2(F_1)} + \sigma^2_{F_3(F_1,F_2)}$
$F_2(F_1)$	$a(b-1)$	$n\sigma^2_{F_2(F_1)} + \sigma^2_{F_3(F_1,F_2)}$
$F_3(F_1, F_2)$	$ab(n-1)$	$\sigma^2_{F_3(F_1,F_2)}$
Total	abn-1	

From the above table, we can determine the appropriate test statistics:

1. To Test no differences across the levels of Factor F_1:

   $H_o : \mu_1 = \mu_2 = \cdots = \mu_a$

   $F = \frac{MS_{F_1}}{MS_{F_2(F_1)}}$

2. If H_o is rejected, then construct contrasts and multiple comparisons across the levels of F_1

3. To Test no differences across the levels of Factor F_2 nested within the levels of F_1:

   $H_o : \sigma_{F_2(F_1)} = 0$

   $F = \frac{MS_{F_2(F_1)}}{MS_{F_3(F_1,F_2)}}$

4. If H_o is rejected, no further comparisons are relevant

Case 3: F_1-Random Levels, F_2 random levels within the levels of F_1, F_3 random levels within the levels of F_2

Model: $y_{ijk} = \mu + a_i + b_{j(i)} + c_{k(ij)} \quad i = 1, \ldots, a; \quad j = 1, \ldots, b; \quad k = 1, \ldots, n_{ij}$
with model conditions:

1. a_i are iid $N(0, \sigma^2_{F_1})$ r.v.'s

2. $b_{j(i)}$ are iid $N(0, \sigma^2_{F_2(F_1)})$ r.v.'s ; $c_{j(i,j)}$'s which are iid $N(0, \sigma^2_{F_3(F_1,F_2)})$ r.v.'s
3. a_i's, $b_{j(i)}$'s; and $c_{j(i,j)}$'s are all independent

Expected Mean Squares are given in the following table for the balanced case ($r_i = b, n_{ij} = n$):

Source	df	Expected Mean Squares
F_1	$a-1$	$nb\sigma^2_{F_1} + n\sigma^2_{F_2(F_1)} + \sigma^2_{F_3(F_1,F_2)}$
$F_2(F_1)$	$a(b-1)$	$n\sigma^2_{F_2(F_1)} + \sigma^2_{F_3(F_1,F_2)}$
$F_3(F_1, F_2)$	$ab(n-1)$	$\sigma^2_{F_3(F_1,F_2)}$
Total	abn-1	

From the above table, we can determine the appropriate test statistics:

1. To Test no differences across the levels of Factor F_1:

   $H_o : \sigma_{F_1} = 0$

   $F = \frac{MS_{F_1}}{MS_{F_2(F_1)}}$

2. If H_o is rejected, no further comparison are relevant

3. To Test no differences across the levels of Factor F_2 nested within the levels of F_1:

   $H_o : \sigma_{F_2(F_1)} = 0$

   $F = \frac{MS_{F_2(F_1)}}{MS_{F_3(F_1,F_2)}}$

4. If H_o is rejected, no further comparisons are relevant

The SAS code needed to analyze our Calcium Content Example is given here: Plant, Leaf(Plant), and Determination are all random.

```
*nested_equalreps.sas;
* This is an example from Snedecor and Cochran of a
Nested Design involving an experiment
to investigate the variability of calcium concentration
in the leaves of turnip greens.
Four plants were randomly selected
and then three leaves were randomly
selected from each plant. Two 100-mg samples
from each leaf. The amount of calcium is determined
by microchemical methods;

ods html;  ods graphics on;

options pagesize=55 linesize=80;
Title 'Nested Design - Equal Sample Sizes';
data raw;
input plant leaf sample X @@;
cards;
1 1 1  3.28  1 1 2  3.09  1 2 1  3.52  1 2 2  3.48
1 3 1  2.88  1 3 2  2.80
2 1 1  2.46  2 1 2  2.44  2 2 1  1.87  2 2 2  1.92
2 3 1  2.19  2 3 2  2.19
3 1 1  2.77  3 1 2  2.66  3 2 1  3.74  3 2 2  3.44
3 3 1  2.55  3 3 2  2.55
4 1 1  3.78  4 1 2  3.87  4 2 1  4.07  4 2 2  4.12
4 3 1  3.31  4 3 2  3.31
run;
proc mixed cl alpha=.05 COVTEST method=type1;
class plant leaf sample;
model X = /residuals;
random plant leaf(plant);
run;
ods graphics off;  ods html close;
```

use the type1 b/c hierarchical nested design.

don't list smallest level (F_3 F_2, F_1) b/c that will be the error.

Nested Design - Equal Sample Sizes

The Mixed Procedure

Model Information	
Data Set	WORK.RAW
Dependent Variable	X
Covariance Structure	Variance Components
Estimation Method	Type 1
Residual Variance Method	Factor
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Containment

Class Level Information		
Class	**Levels**	**Values**
plant	4	1 2 3 4
leaf	3	1 2 3
sample	2	1 2

Dimensions	
Covariance Parameters	3
Columns in X	1
Columns in Z	16
Subjects	1
Max Obs per Subject	24

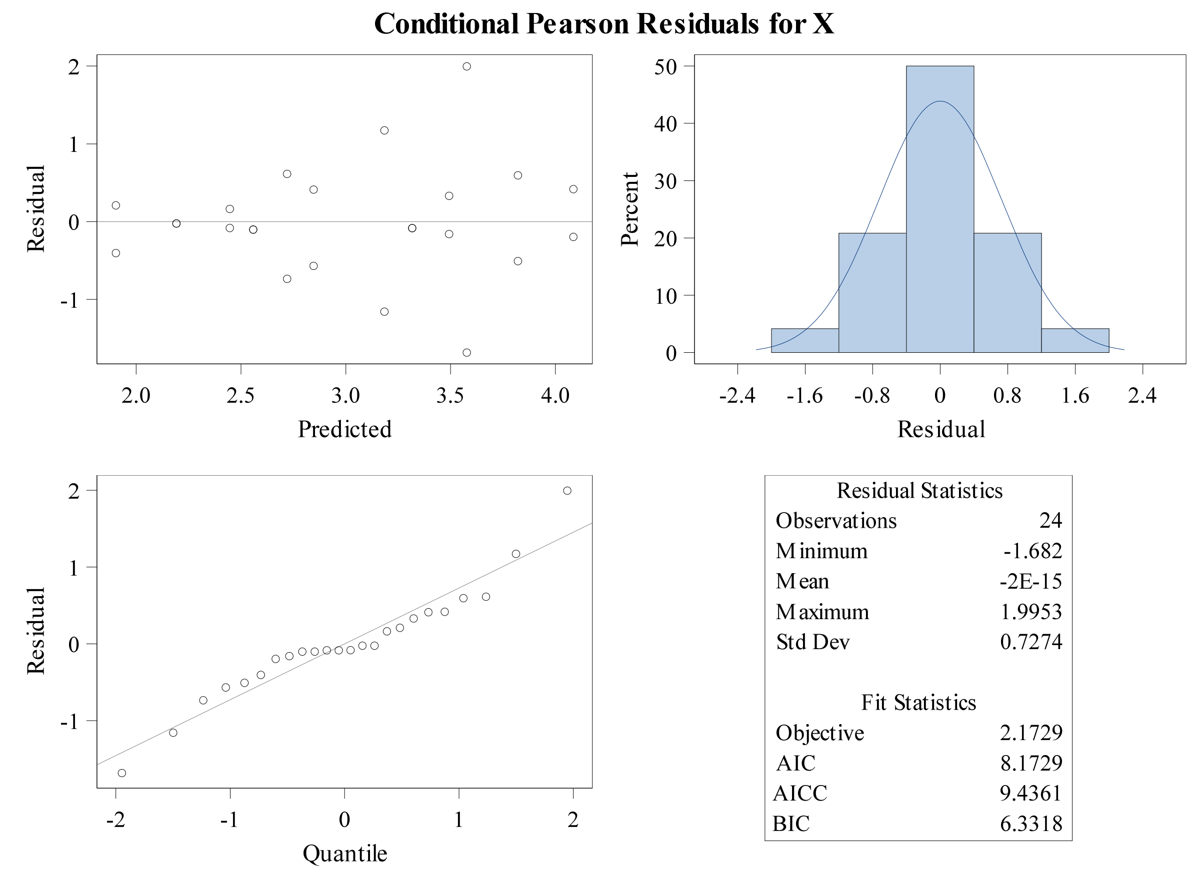
Number of Observations	
Number of Observations Read	24
Number of Observations Used	24
Number of Observations Not Used	0

Type 1 Analysis of Variance								
Source	**DF**	**Sum of Squares**	**Mean Square**	**Expected Mean Square**	**Error Term**	**Error DF**	**F Value**	**Pr > F**
plant	3	7.560346	2.520115	Var(Residual) + 2 Var(leaf(plant)) + 6 Var(plant)	MS(leaf(plant))	8	7.67	0.0097
leaf(plant)	8	2.630200	0.328775	Var(Residual) + 2 Var(leaf(plant))	MS(Residual)	12	49.41	<.0001
Residual	12	0.079850	0.006654	Var(Residual)	.	.	.	.

Nested Design - Equal Sample Sizes

The Mixed Procedure

Covariance Parameter Estimates							
Cov Parm	**Estimate**	**Standard Error**	**Z Value**	**Pr Z**	**Alpha**	**Lower**	**Upper**
plant	0.3652	0.3440	1.06	0.2884	0.05	-0.3091	1.0395
leaf(plant)	0.1611	0.08220	1.96	0.0501	0.05	-0.00006	0.3222
Residual	0.006654	0.002717	2.45	0.0072	0.05	0.003422	0.01813



Residual Statistics	
Observations	24
Minimum	-1.682
Mean	-2E-15
Maximum	1.9953
Std Dev	0.7274

Fit Statistics	
Objective	2.1729
AIC	8.1729
AICC	9.4361
BIC	6.3318

Unbalanced Nested Design

Example: A study of the variation of the tensile strength of fiber optic cable is to be designed. There are three major manufacturers (M) of the cable. From each manufacturer, large rolls (R) of wire are randomly selected. From each roll, random samples (S) of 12 inch lengths of wire are taken. The tensile strength of each of the 12 in lengths of wire is then determined in a laboratory.

The factors are M - Fixed levels, R(M) - Random levels, S(M,R) - Random levels.

The data is given here:

	Manufacturer (M)							
	M_1			M_2			M_3	
R(M)	1	2	3	1	2	3	1	2
y_{ijk}	110	130	50	130	45	120	100	130
	90	115	75	45	55	50	200	80
	120	105	85	50	65	150	90	70
			40	40			70	80
							90	150
$\bar{y}_{ij.}$	106.67	116.67	62.5	66.25	55	106.67	110	102
n_{ij}	3	3	4	4	3	3	5	5
r_i		10		10			10	
$\bar{y}_{i..}$		92.0		75.0			106.0	
$\bar{y}_{...}$				91.0				

Model: $y_{ijk} = \tau_i + b_{j(i)} + e_{k(i,j)}$ $\mu +$

For our example, we have the following calculations:

$$\begin{aligned}
SS_{TOT} &= \sum_{i=1}^{a}\sum_{j=1}^{r_i}\sum_{k=1}^{n_{ij}}(y_{ijk} - \bar{y}_{...})^2 \\
&= \sum_{i=1}^{3}\sum_{j=1}^{r_i}\sum_{k=1}^{n_{ij}}(y_{ijk} - 91.0)^2 \\
&= [(110 - 91.0)^2 + (90 - 91.0)^2 + (120 - 91.0)^2 + \cdots + (150 - 91.0)^2] \\
&= 44170.0 \quad \text{with} \quad df = 30 - 1 = 29
\end{aligned}$$

$$\begin{aligned}
SS_M &= \sum_{i=1}^{a} n_{i.}(\bar{y}_{i..} - \bar{y}_{...})^2 \\
&= (10)[(92.0 - 91.0)^2] + (10)[(75.0 - 91.0)^2] + (10)[106.0 - 91.0)^2] \\
&= 4820.0 \quad \text{with} \quad df = 3 - 1 = 2
\end{aligned}$$

$$
\begin{aligned}
SS_{R(M)} &= \sum_{i=1}^{a}\sum_{j=1}^{r_i} n_{ij}(\bar{y}_{ij.} - \bar{y}_{i..})^2 \\
&= (3)(106.67 - 92.0)^2 + (3)(116.67 - 92.0)^2 + (4)(62.5 - 92.0)^2 \\
&\quad +(4)(66.25 - 75.0)^2 + (3)(55.0 - 75.0)^2 + (3)(106.67 - 75)^2 \\
&\quad +(5)(110 - 106)^2 + (5)(102 - 106)^2 \\
&= 10626.25 \quad \text{with} \quad df = (3-1) + (3-1) + (2-1) = 5
\end{aligned}
$$

$$
\begin{aligned}
SS_{S(M,R)} &= \sum_{i=1}^{a}\sum_{j=1}^{r_i}\sum_{k=1}^{n_{ij}} (y_{ijk} - \bar{y}_{ij.})^2 \\
&= [(110 - 106.67)^2 + (90 - 106.67)^2 + \cdots + (80 - 102)^2 + (150 - 102)^2 \\
&= 28723.75 \quad \text{with} \quad df = 29 - (2+5) = 22
\end{aligned}
$$

The SAS code to analyze the above experiment is given here:

```
ods html; ods graphics on;
option ls=80 ps=55 nocenter nodate;
title 'NESTED DESIGN - UNEQUAL SAMPLE SIZES';
data strength;
INPUT MAN $ ROLL SAMPLE Y @@;
LABEL MAN = 'MANUFACTURER' Y= 'STRENGTH';
cards;
M1 1 1 110 M1 1 2  90 M1 1 3 120
M1 2 1 130 M1 2 2 115 M1 2 3 105
M1 3 1  50 M1 3 2  75 M1 3 3  85 M1 3 4 40
M2 1 1 130 M2 1 2  45 M2 1 3  50 M2 1 4 40
M2 2 1  45 M2 2 2  55 M2 2 3  65
M2 3 1 120 M2 3 2  50 M2 3 3 150
M3 1 1 100 M3 1 2 200 M3 1 3  90 M3 1 4 70 M3 1 5 90
M3 2 1 130 M3 2 2  80 M3 2 3  70 M3 2 4 80 M3 2 5 150
RUN;
TITLE ANALYSIS USING PROC MIXED-TYPE1;
PROC MIXED METHOD=TYPE1 CL ALPHA=.05 COVTEST;
CLASS MAN ROLL SAMPLE;
MODEL Y = MAN/RESIDUALS;
RANDOM ROLL(MAN);
LSMEANS MAN/ADJUST=TUKEY;
RUN;
ods graphics off; ods html close;
```

ANALYSIS USING PROC MIXED-TYPE1

The Mixed Procedure

Model Information	
Data Set	WORK.STRENGTH
Dependent Variable	Y
Covariance Structure	Variance Components
Estimation Method	Type 1
Residual Variance Method	Factor
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Containment

Class Level Information		
Class	**Levels**	**Values**
MAN	3	M1 M2 M3
ROLL	3	1 2 3
SAMPLE	5	1 2 3 4 5

Dimensions	
Covariance Parameters	2
Columns in X	4
Columns in Z	8
Subjects	1
Max Obs per Subject	30

Number of Observations	
Number of Observations Read	30
Number of Observations Used	30
Number of Observations Not Used	0

Type 1 Analysis of Variance

Source	DF	Sum of Squares	Mean Square	Expected Mean Square	Error Term	Error DF	F Value	Pr > F
MAN	2	4820.000000	2410.000000	Var(Residual) + 3.9333 Var(ROLL(MAN)) + Q(MAN)	1.0806 MS(ROLL(MAN)) - 0.0806 MS(Residual)	4.5502	1.10	0.4077
ROLL(MAN)	5	10626	2125.250000	Var(Residual) + 3.64 Var(ROLL(MAN))	MS(Residual)	22	1.63	0.1943
Residual	22	28724	1305.625000	Var(Residual)	.	.	.	.

ANALYSIS USING PROC MIXED-TYPE1

The Mixed Procedure

Covariance Parameter Estimates							
Cov Parm	Estimate	Standard Error	Z Value	Pr Z	Alpha	Lower	Upper
ROLL(MAN)	225.17	415.79	0.54	0.5881	0.05	-589.76	1040.10
Residual	1305.63	398.01	3.28	0.0005	0.05	780.95	2615.45

Fit Statistics	
-2 Res Log Likelihood	279.8
AIC (Smaller is Better)	283.8
AICC (Smaller is Better)	284.3
BIC (Smaller is Better)	284.0

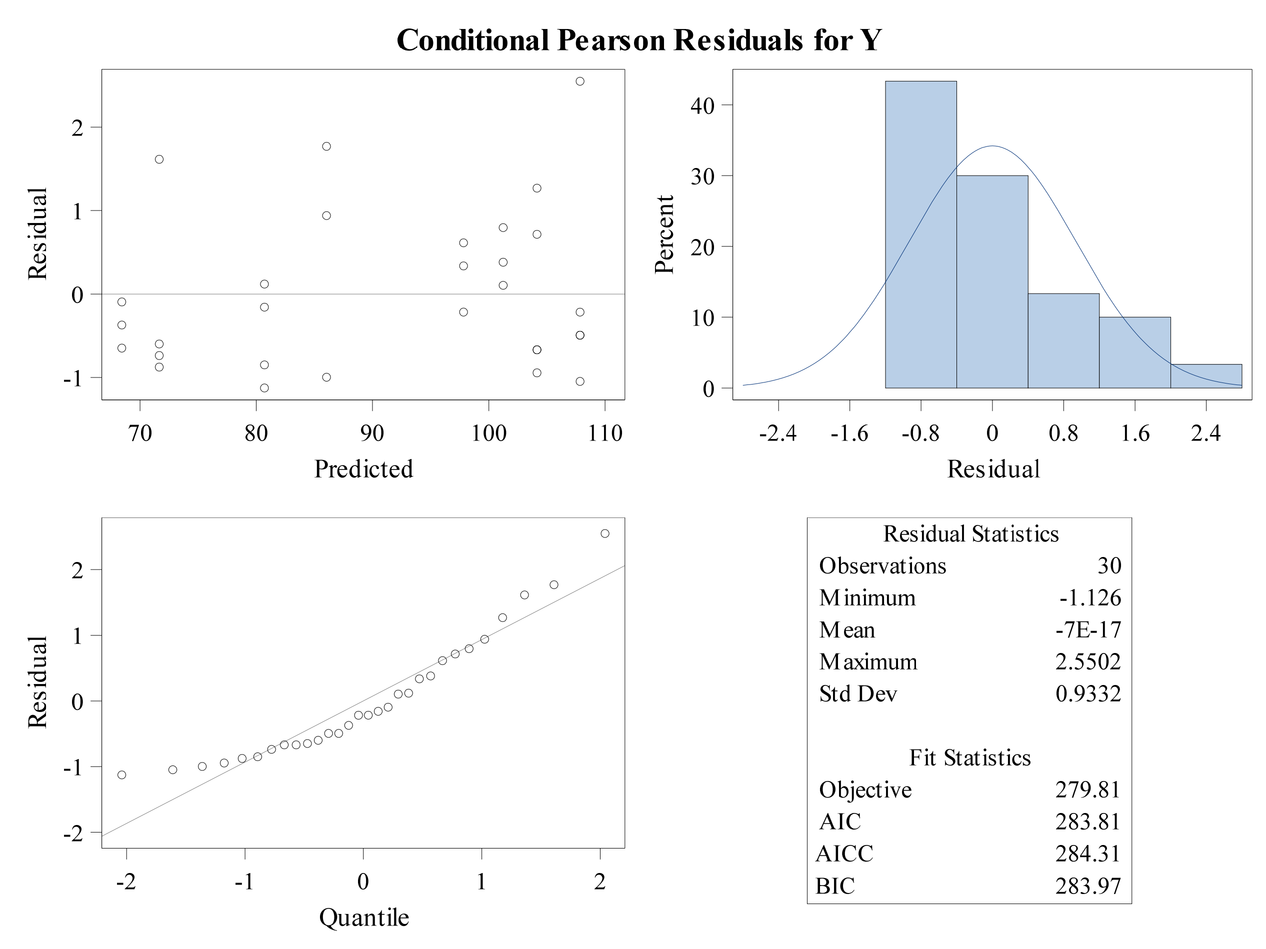
Type 3 Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
MAN	2	5	1.07	0.4116

Least Squares Means						
Effect	MANUFACTURER	Estimate	Standard Error	DF	t Value	Pr > |t|
MAN	M1	93.2555	14.3717	5	6.49	0.0013
MAN	M2	75.3724	14.3717	5	5.24	0.0033
MAN	M3	106.00	15.5932	5	6.80	0.0010

Differences of Least Squares Means									
Effect	MANUFACTURER	MANUFACTURER	Estimate	Standard Error	DF	t Value	Pr > |t|	Adjustment	Adj P
MAN	M1	M2	17.8831	20.3247	5	0.88	0.4192	Tukey-Kramer	0.6748
MAN	M1	M3	-12.7445	21.2060	5	-0.60	0.5741	Tukey-Kramer	0.8258
MAN	M2	M3	-30.6276	21.2060	5	-1.44	0.2083	Tukey-Kramer	0.3895

ANALYSIS USING PROC MIXED-TYPE1

The Mixed Procedure



Residual Statistics	
Observations	30
Minimum	-1.126
Mean	-7E-17
Maximum	2.5502
Std Dev	0.9332

Fit Statistics	
Objective	279.81
AIC	283.81
AICC	284.31
BIC	283.97

The Pearson Conditional Residuals indicate that the distribution of the residuals is right skewed with slightly unequal variances. Thus, a log-transformation of the data was done and the transformed data was then analyzed.

The SAS code to analyze a log transformation of the tensile strength from the above experiment is given here:

```
* nested_unequalreps.sas;
*This is an example of a nested experiment with
MAN-fixed factor, ROLL and SAMPLE-random factors.
There is a unequal number of levels in the nested
factors. Three different methods of analysis are
considered.;
ods html; ods graphics on;
option ls=80 ps=55 nocenter nodate;
title 'NESTED DESIGN - UNEQUAL SAMPLE SIZES';
data strength;
INPUT MAN $ ROLL SAMPLE Y @@;
LY = LOG(Y);
LABEL MAN = 'MANUFACTURER' LY= 'LOG-STRENGTH';
cards;
M1 1 1 110 M1 1 2  90 M1 1 3 120
M1 2 1 130 M1 2 2 115 M1 2 3 105
M1 3 1  50 M1 3 2  75 M1 3 3  85 M1 3 4 40
M2 1 1 130 M2 1 2  45 M2 1 3  50 M2 1 4 40
M2 2 1  45 M2 2 2  55 M2 2 3  65
M2 3 1 120 M2 3 2  50 M2 3 3 150
M3 1 1 100 M3 1 2 200 M3 1 3  90 M3 1 4 70 M3 1 5 90
M3 2 1 130 M3 2 2  80 M3 2 3  70 M3 2 4 80 M3 2 5 150
RUN;
TITLE ANALYSIS USING PROC MIXED-TYPE1;
PROC MIXED METHOD=TYPE1  CL ALPHA=.05 COVTEST;
CLASS MAN ROLL SAMPLE;
MODEL LY = MAN/RESIDUALS;
RANDOM ROLL(MAN);
LSMEANS MAN/ADJUST=TUKEY;
RUN;
TITLE ANALYSIS USING PROC MIXED-REML;
PROC MIXED METHOD=REML CL ALPHA=.05 COVTEST;
CLASS MAN ROLL SAMPLE;
MODEL LY = MAN/RESIDUALS;
RANDOM ROLL(MAN);
LSMEANS MAN/ADJUST=TUKEY;
RUN;

ods graphics off; ods html close;
```

ANALYSIS USING PROC MIXED-TYPE1

The Mixed Procedure

Model Information	
Data Set	WORK.STRENGTH
Dependent Variable	LY
Covariance Structure	Variance Components
Estimation Method	Type 1
Residual Variance Method	Factor
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Containment

Class Level Information		
Class	**Levels**	**Values**
MAN	3	M1 M2 M3
ROLL	3	1 2 3
SAMPLE	5	1 2 3 4 5

Dimensions	
Covariance Parameters	2
Columns in X	4
Columns in Z	8
Subjects	1
Max Obs per Subject	30

Number of Observations	
Number of Observations Read	30
Number of Observations Used	30
Number of Observations Not Used	0

Type 1 Analysis of Variance

Source	DF	Sum of Squares	Mean Square	Expected Mean Square	Error Term	Error DF	F Value	Pr > F
MAN	2	0.852297	0.426148	Var(Residual) + 3.9333 Var(ROLL(MAN)) + Q(MAN)	1.0806 MS(ROLL(MAN)) - 0.0806 MS(Residual)	4.6602	1.32	0.3506
ROLL(MAN)	5	1.542976	0.308595	Var(Residual) + 3.64 Var(ROLL(MAN))	MS(Residual)	22	2.16	0.0953
Residual	22	3.136098	0.142550	Var(Residual)	.	.	.	.

ANALYSIS USING PROC MIXED-TYPE1

The Mixed Procedure

Covariance Parameter Estimates							
Cov Parm	**Estimate**	**Standard Error**	**Z Value**	**Pr Z**	**Alpha**	**Lower**	**Upper**
ROLL(MAN)	0.04562	0.05658	0.81	0.4201	0.05	-0.06529	0.1565
Residual	0.1425	0.04326	3.30	0.0005	0.05	0.08526	0.2856

Fit Statistics	
-2 Res Log Likelihood	35.0
AIC (Smaller is Better)	39.0
AICC (Smaller is Better)	39.5
BIC (Smaller is Better)	39.1

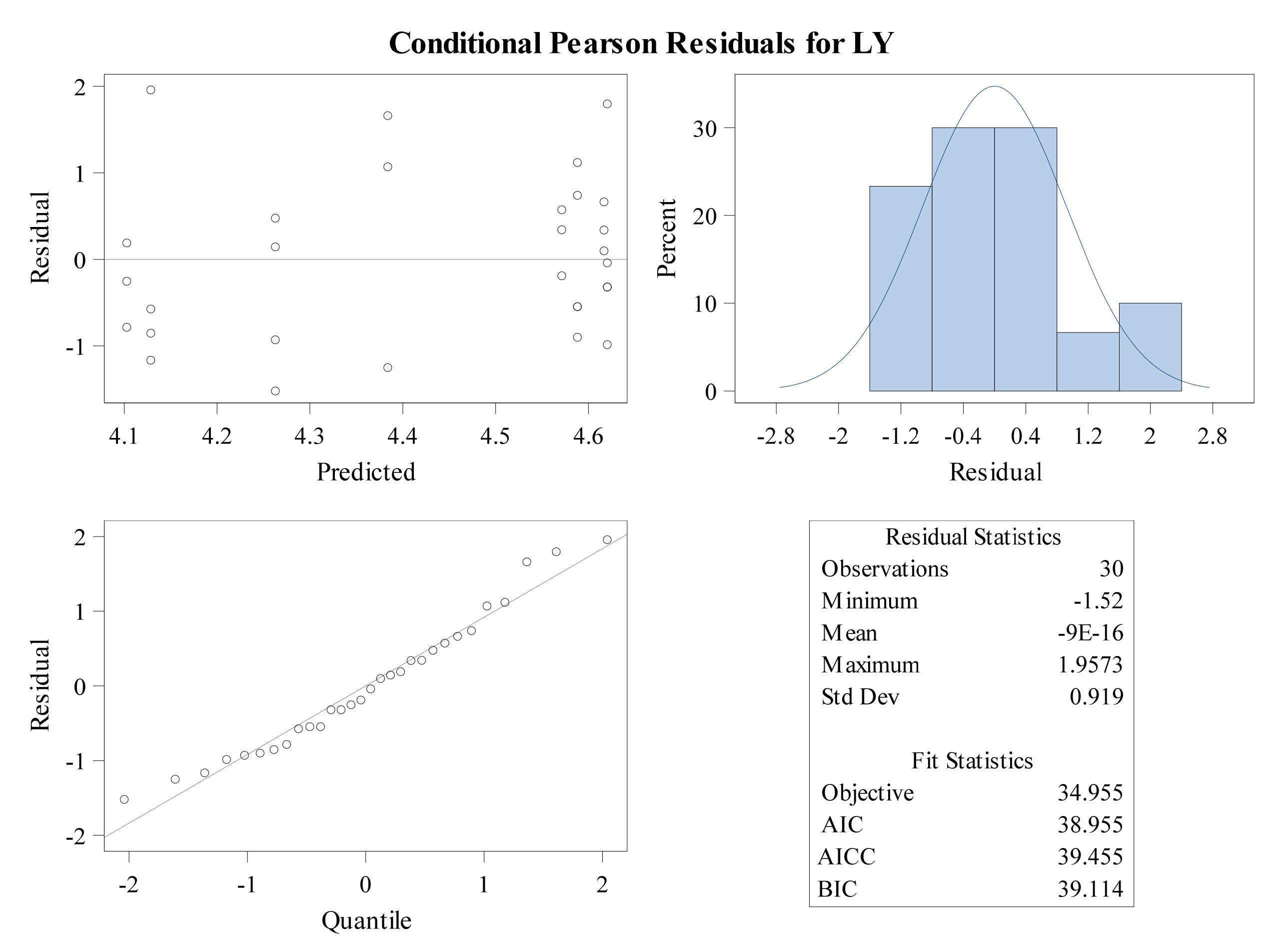
Type 3 Tests of Fixed Effects				
Effect	**Num DF**	**Den DF**	**F Value**	**Pr > F**
MAN	2	5	1.31	0.3495

Least Squares Means						
Effect	**MANUFACTURER**	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**
MAN	M1	4.4836	0.1721	5	26.06	<.0001
MAN	M2	4.2050	0.1721	5	24.44	<.0001
MAN	M3	4.6042	0.1925	5	23.92	<.0001

Differences of Least Squares Means									
Effect	**MANUFACTURER**	**MANUFACTURER**	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**	**Adjustment**	**Adj P**
MAN	M1	M2	0.2786	0.2433	5	1.14	0.3040	Tukey-Kramer	0.5309
MAN	M1	M3	-0.1206	0.2582	5	-0.47	0.6600	Tukey-Kramer	0.8892
MAN	M2	M3	-0.3992	0.2582	5	-1.55	0.1827	Tukey-Kramer	0.3483

ANALYSIS USING PROC MIXED-TYPE1

The Mixed Procedure



ANALYSIS USING PROC MIXED-REML

The Mixed Procedure

Model Information	
Data Set	WORK.STRENGTH
Dependent Variable	LY
Covariance Structure	Variance Components
Estimation Method	REML
Residual Variance Method	Profile
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Containment

Class Level Information		
Class	**Levels**	**Values**
MAN	3	M1 M2 M3
ROLL	3	1 2 3
SAMPLE	5	1 2 3 4 5

Dimensions	
Covariance Parameters	2
Columns in X	4
Columns in Z	8
Subjects	1
Max Obs per Subject	30

Number of Observations	
Number of Observations Read	30
Number of Observations Used	30
Number of Observations Not Used	0

Iteration History			
Iteration	**Evaluations**	**-2 Res Log Like**	**Criterion**
0	1	36.20654450	
1	2	34.95512505	0.00013815
2	1	34.95408590	0.00000022
3	1	34.95408430	0.00000000

ANALYSIS USING PROC MIXED-REML

The Mixed Procedure

Convergence criteria met.

Covariance Parameter Estimates							
Cov Parm	**Estimate**	**Standard Error**	**Z Value**	**Pr > Z**	**Alpha**	**Lower**	**Upper**
ROLL(MAN)	0.04638	0.05793	0.80	0.2117	0.05	0.01033	11.0582
Residual	0.1434	0.04352	3.30	0.0005	0.05	0.08552	0.2888

Fit Statistics	
-2 Res Log Likelihood	35.0
AIC (Smaller is Better)	39.0
AICC (Smaller is Better)	39.5
BIC (Smaller is Better)	39.1

Type 3 Tests of Fixed Effects				
Effect	**Num DF**	**Den DF**	**F Value**	**Pr > F**
MAN	2	5	1.29	0.3531

EXAMPLE OF NESTED AND CROSSED DESIGN

An experiment was designed to investigate the cutoff times of automatic safety switches on lawnmowers produced by three manufacturers, M_1, M_2, and M_3. Three lawnmowers are randomly selected from each of three manufacturers. The lawnmowers are evaluated at two different engine speeds, S_1 and S_2. At both of these speeds, the lawnmowers are evaluated twice. The cutoff times in 10^{-2} seconds are given in the following table. The model for this experiment with y_{ijkm} the cutoff time for the mth run of mower j from manufacturer i conducted at speed k is given by

$$y_{ijkm} = \mu + \tau_i + L_{j(i)} + \beta_k + (\tau\beta)_{ik} + (L * \beta)_{j(i),k} + e_{ijkm}$$

		Speed(S)	
Manufacturer(M)	Lawnmower(L)	L	H
M_1	1	211,230	278,278
	2	184,188	249,272
	3	216,232	275,271
M_2	4	205,217	247,251
	5	169,168	239,252
	6	200,187	261,242
M_3	7	195,207	227,231
	8	199,198	259,282
	9	203,197	251,222

Source	DF	MS	EXPECTED MEAN SQUARE
M	2	MS_M	$\sigma^2_{e(L,M,S)} + 2\sigma^2_{L(M)*S} + 4\sigma^2_{L(M)} + 12Q_M$
L(M)	6	$MS_{L(M)}$	$\sigma^2_{e(L,M,S)} + 2\sigma^2_{L(M)*S} + 4\sigma^2_{L(M)}$
S	1	MS_S	$\sigma^2_{e(L,M,S)} + 2\sigma^2_{L(M)*S} + 18Q_S$
M*S	2	MS_{M*S}	$\sigma^2_{e(L,M,S)} + 2\sigma^2_{L(M)*S} + 6Q_{M*S}$
L(M)*S	6	$MS_{L(M)*S}$	$\sigma^2_{e(L,M,S)} + 2\sigma^2_{L(M)*S}$
Error	18	MSE	$\sigma^2_{e(L,M,S)}$
Total	35		

```
*nested,lawn.sas;
OPTIONS LS=90 PS=55;
DATA LAWN;
INPUT M $ L S $ SUB Y @@;
CARDS;
1 1 L 1 211  1 1 L 2 230  1 1 H 1 278  1 1 H 2 278
1 2 L 1 184  1 2 L 2 188  1 2 H 1 249  1 2 H 2 272
1 3 L 1 216  1 3 L 2 232  1 3 H 1 275  1 3 H 2 271
2 4 L 1 205  2 4 L 2 217  2 4 H 1 247  2 4 H 2 251
2 5 L 1 169  2 5 L 2 168  2 5 H 1 239  2 5 H 2 252
3 6 L 1 200  2 6 L 2 187  2 6 H 1 261  2 6 H 2 242
3 7 L 1 195  3 7 L 2 207  3 7 H 1 227  3 7 H 2 231
3 8 L 1 199  3 8 L 2 198  3 8 H 1 259  3 8 H 2 282
3 9 L 1 203  3 9 L 2 197  3 9 H 1 251  3 9 H 2 222
RUN;
PROC MIXED CL ALPHA=.05 COVTEST METHOD=TYPE3;
CLASS S M L SUB;
MODEL Y = M S M*S;
RANDOM L(M) S*L(M);
LSMEANS M S M*S/ADJUST=TUKEY;
RUN;



proc GLIMMIX data=LAWN;
CLASS S M L DEM;
MODEL Y = M S M*S/ddfm=SAT;
RANDOM L(M) S*L(M);
lsmeans M S M*S /  plot = meanplot cl;
RUN;
```

ANALYSIS USING PROC MIXED-REML

The Mixed Procedure

odel Information	
Data Set	WORK.LAWN
Dependent Variable	Y
Covariance Structure	Variance Components
Estimation Method	Type 3
Residual Variance Method	Factor
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Containment

Class Level Information		
Class	Levels	Values
S	2	H L
M	3	1 2 3
L	9	1 2 3 4 5 6 7 8 9
DEM	2	1 2

Dimensions	
Covariance Parameters	3
Columns in X	12
Columns in Z	27
Subjects	1
Max Obs per Subject	36

Number of Observations	
Number of Observations Read	36
Number of Observations Used	36
Number of Observations Not Used	0

ANALYSIS USING PROC MIXED-REML

No evidence of mech on fixed effects - now can look at individual effects

The Mixed Procedure

Type 3 Analysis of Variance								
Source	**DF**	**Sum of Squares**	**Mean Square**	**Expected Mean Square**	**Error Term**	**Error DF**	**F Value**	**Pr > F**
M	2	2971.500000	1485.750000	Var(Residual) + 2 Var(S*L(M)) + 4 Var(L(M)) + Q(M,S*M)	MS(L(M))	6	2.39	0.1722
S	1	26732	26732	Var(Residual) + 2 Var(S*L(M)) + Q(S,S*M)	MS(S*L(M))	6	73.33	0.0001
S*M	2	375.166667	187.583333	Var(Residual) + 2 Var(S*L(M)) + Q(S*M)	MS(S*L(M))	6	0.51	0.6219
L(M)	6	3726.000000	621.000000	Var(Residual) + 2 Var(S*L(M)) + 4 Var(L(M))	MS(S*L(M))	6	1.70	0.2668
S*L(M)	6	2187.333333	364.555556	Var(Residual) + 2 Var(S*L(M))	MS(Residual)	18	3.64	0.0153
Residual	18	1802.500000	100.138889	Var(Residual)	.	.	.	.

Covariance Parameter Estimates							
Cov Parm	**Estimate**	**Standard Error**	**Z Value**	**Pr Z**	**Alpha**	**Lower**	**Upper**
L(M)	64.1111	103.94	0.62	0.5374	0.05	-139.60	267.82
S*L(M)	132.21	106.55	1.24	0.2147	0.05	-76.6324	341.05
Residual	100.14	33.3796	3.00	0.0013	0.05	57.1743	219.00

Fit Statistics	
-2 Res Log Likelihood	252.8
AIC (Smaller is Better)	258.8
AICC (Smaller is Better)	259.7
BIC (Smaller is Better)	259.4

Type 3 Tests of Fixed Effects				
Effect	**Num DF**	**Den DF**	**F Value**	**Pr > F**
M	2	6	2.39	0.1722
S	1	6	73.33	0.0001
S*M	2	6	0.51	0.6219

ANALYSIS USING PROC MIXED-REML

The Mixed Procedure

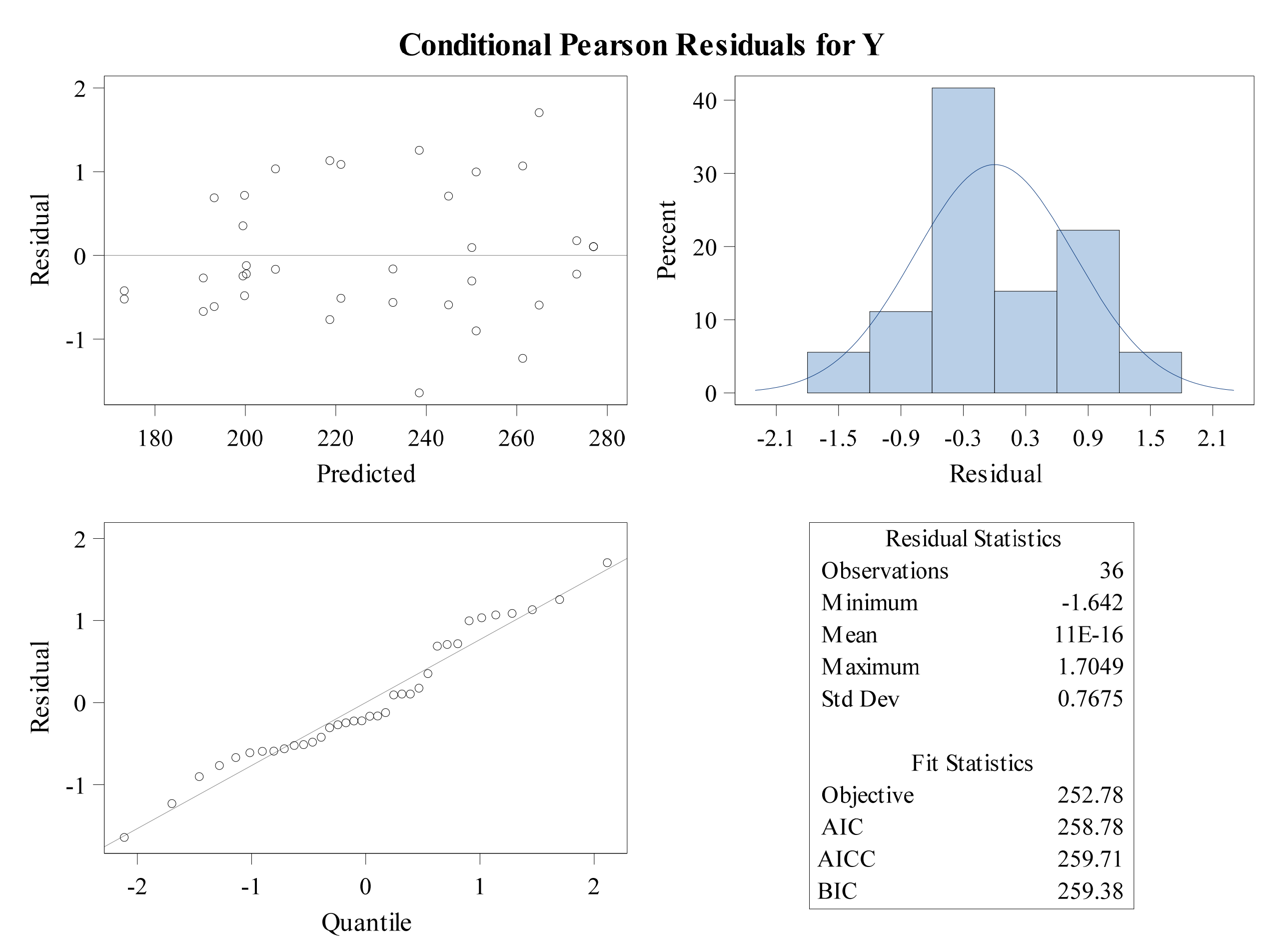
Least Squares Means							
Effect	**S**	**M**	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**
M		1	240.33	7.1937	6	33.41	<.0001
M		2	219.83	7.1937	6	30.56	<.0001
M		3	222.58	7.1937	6	30.94	<.0001
S	H		254.83	5.2323	6	48.70	<.0001
S	L		200.33	5.2323	6	38.29	<.0001
S*M	H	1	270.50	9.0625	6	29.85	<.0001
S*M	H	2	248.67	9.0625	6	27.44	<.0001
S*M	H	3	245.33	9.0625	6	27.07	<.0001
S*M	L	1	210.17	9.0625	6	23.19	<.0001
S*M	L	2	191.00	9.0625	6	21.08	<.0001
S*M	L	3	199.83	9.0625	6	22.05	<.0001

Differences of Least Squares Means											
Effect	**S**	**M**	**_S**	**_M**	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**	**Adjustment**	**Adj P**
M		1		2	20.5000	10.1735	6	2.02	0.0905	Tukey	0.1893
M		1		3	17.7500	10.1735	6	1.74	0.1316	Tukey	0.2652
M		2		3	-2.7500	10.1735	6	-0.27	0.7960	Tukey	0.9608
S	H		L		54.5000	6.3644	6	8.56	0.0001	Tukey-Kramer	0.0001
S*M	H	1	H	2	21.8333	12.8164	6	1.70	0.1394	Tukey-Kramer	0.5736
S*M	H	1	H	3	25.1667	12.8164	6	1.96	0.0972	Tukey-Kramer	0.4502
S*M	H	1	L	1	60.3333	11.0235	6	5.47	0.0016	Tukey-Kramer	0.0116
S*M	H	1	L	2	79.5000	12.8164	6	6.20	0.0008	Tukey-Kramer	0.0062
S*M	H	1	L	3	70.6667	12.8164	6	5.51	0.0015	Tukey-Kramer	0.0112
S*M	H	2	H	3	3.3333	12.8164	6	0.26	0.8035	Tukey-Kramer	0.9997
S*M	H	2	L	1	38.5000	12.8164	6	3.00	0.0239	Tukey-Kramer	0.1459
S*M	H	2	L	2	57.6667	11.0235	6	5.23	0.0020	Tukey-Kramer	0.0144
S*M	H	2	L	3	48.8333	12.8164	6	3.81	0.0089	Tukey-Kramer	0.0599
S*M	H	3	L	1	35.1667	12.8164	6	2.74	0.0336	Tukey-Kramer	0.1953
S*M	H	3	L	2	54.3333	12.8164	6	4.24	0.0054	Tukey-Kramer	0.0381
S*M	H	3	L	3	45.5000	11.0235	6	4.13	0.0062	Tukey-Kramer	0.0428
S*M	L	1	L	2	19.1667	12.8164	6	1.50	0.1854	Tukey-Kramer	0.6796

ANALYSIS USING PROC MIXED-REML

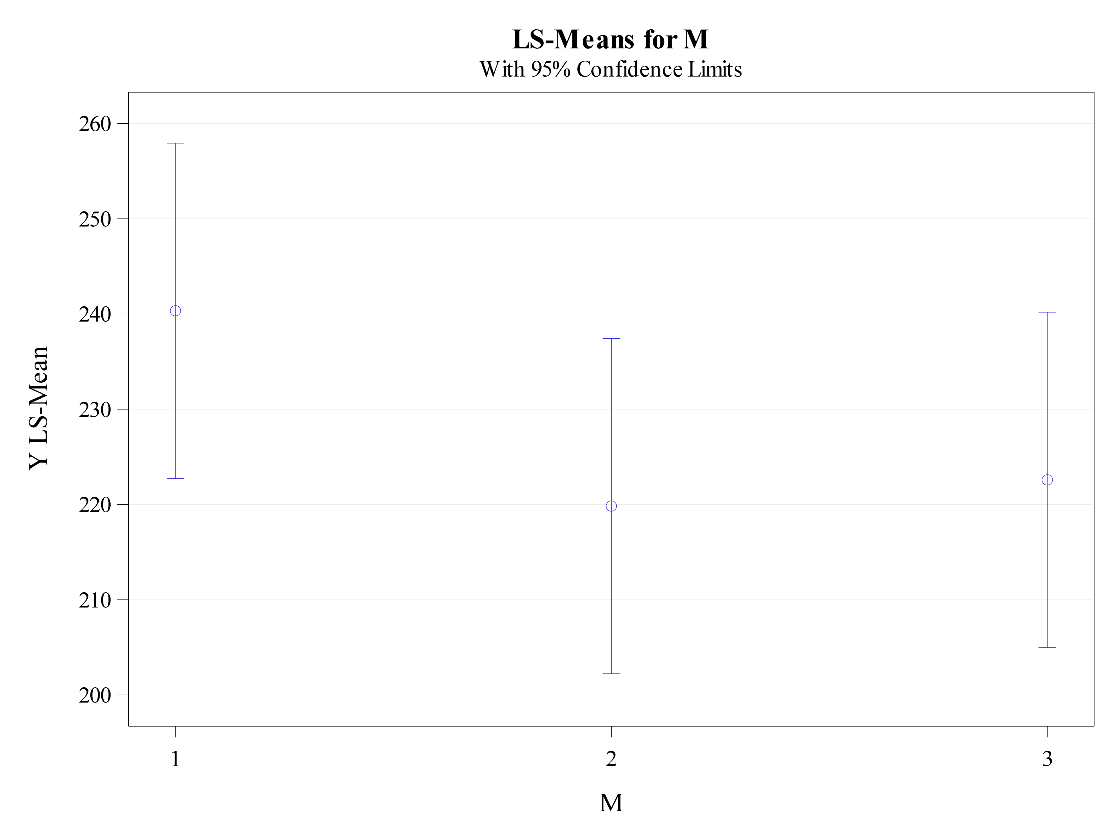
The Mixed Procedure

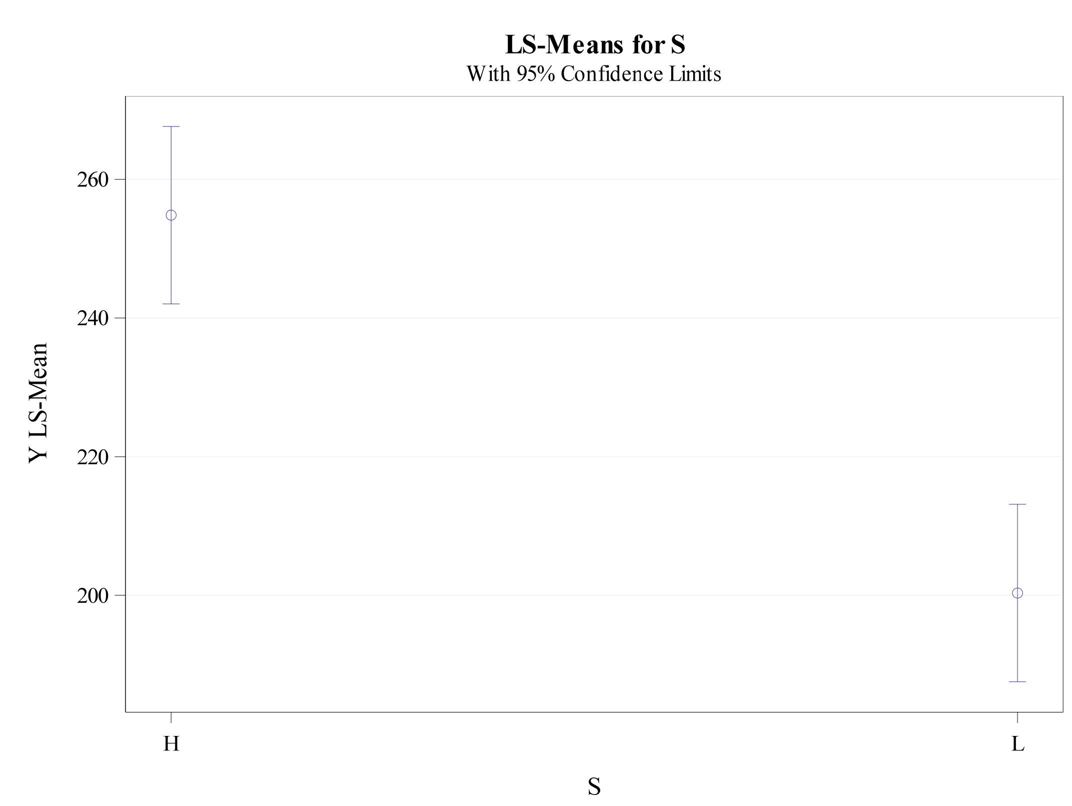
Differences of Least Squares Means											
Effect	S	M	_S	_M	Estimate	Standard Error	DF	t Value	Pr > |t|	Adjustment	Adj P
S*M	L	1	L	3	10.3333	12.8164	6	0.81	0.4509	Tukey-Kramer	0.9560
S*M	L	2	L	3	-8.8333	12.8164	6	-0.69	0.5164	Tukey-Kramer	0.9767



ANALYSIS USING PROC MIXED-REML

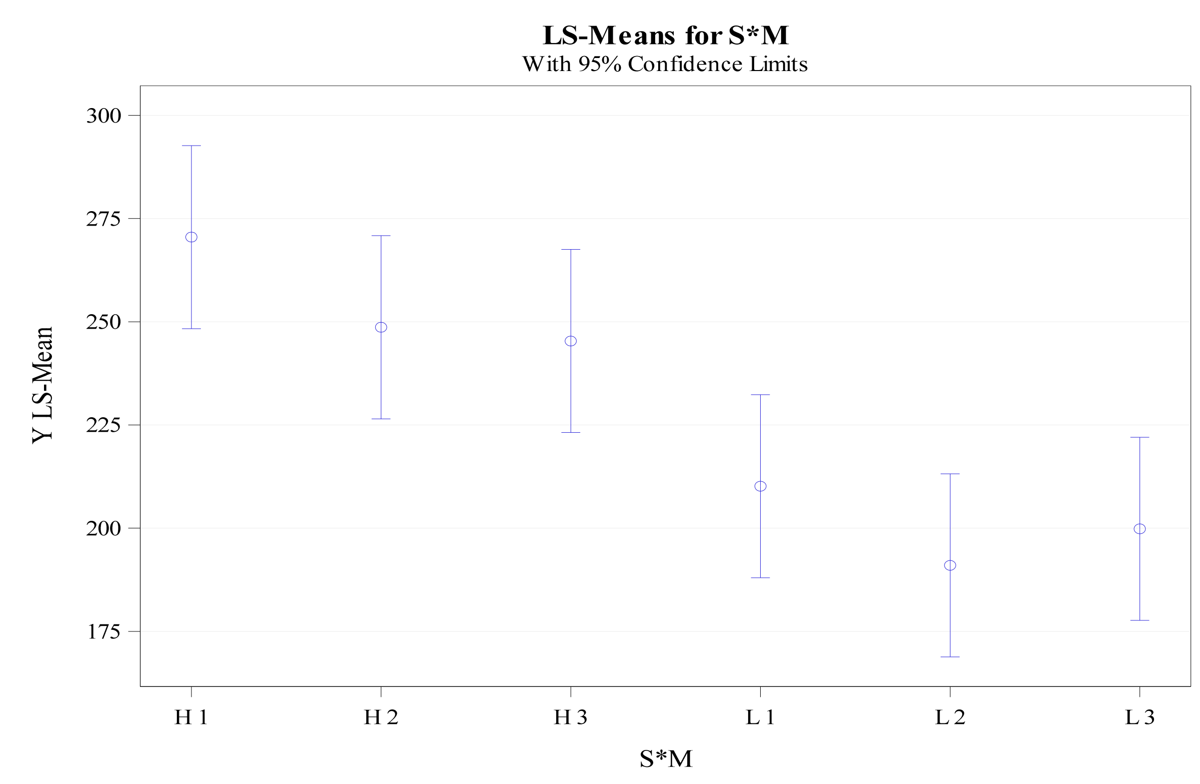
The GLIMMIX Procedure





ANALYSIS USING PROC MIXED-REML

The GLIMMIX Procedure



Lawnmower Example
Calculations of Estimated Errors of Difference in Marginal and Treatment Means

The Lawnmower example will be used to illustrate the procedures to estimate the errors.

The model is $y_{ijkm} = \mu + \tau_i + L_{j(i)} + \beta_k + (\tau\beta)_{ik} + (L*\beta)_{j(i)k} + e_{ijkm}$

where $i = 1,2,3 -$ Manufacturer; $j = 1,2,3 -$ Lawnmower within Man.; $k = 1,2 -$ Speed; $m = 1,2 -$ repeats

1. Compute the estimated standard error of the difference in the means for

   Manufacturer 1 and 2, $\widehat{SE}(\hat{\mu}_{1..} - \hat{\mu}_{2..})$:

   $y_{ijkm} = \mu + \tau_i + L_{j(i)} + \beta_k + (\tau\beta)_{ik} + (L*\beta)_{j(i)k} + e_{ijkm} \Rightarrow$

   $\overline{y}_{1...} = \mu + \tau_1 + \overline{L}_{.(1)} + \overline{\beta}_{.} + \overline{(\tau\beta)}_{1.} + \overline{(L*\beta)}_{.(1).} + \overline{e}_{1...}$

   $\overline{y}_{2...} = \mu + \tau_2 + \overline{L}_{.(2)} + \overline{\beta}_{.} + \overline{(\tau\beta)}_{2.} + \overline{(L*\beta)}_{.(2).} + \overline{e}_{2...} \Rightarrow$

   $Var[\overline{y}_{1...} - \overline{y}_{2...}] = Var\left([\overline{L}_{.(1)} - \overline{L}_{.(2)}]\right) + Var\left([\overline{(L*\beta)}_{.(1).} - \overline{(L*\beta)}_{.(2).}]\right) + Var\left([\overline{e}_{1...} - \overline{e}_{2...}]\right)$

   Therefore, we have

$$
\begin{aligned}
Var[\bar{y}_{1...} - \bar{y}_{2...}] &= 2Var\left(\overline{L}_{.(1)}\right) + 2Var\left(\overline{(L*\beta)}_{.(1).}\right) + 2Var\left(\overline{e}_{1...}\right) \\
&= \frac{2\sigma^2_{L(M)}}{3} + \frac{2\sigma^2_{L(M)*S}}{(3)(2)} + \frac{2\sigma^2_e}{(3)(2)(2)} \\
&= \frac{2}{12}\left[4\sigma^2_{L(M)} + 2\sigma^2_{L(M)*S} + \sigma^2_e\right] \\
&= \frac{1}{6}\left[EMS_{L(M)}\right]
\end{aligned}
$$

   An estimate of the standard error and the degrees of freedom of the estimate are given as follows.

   $\widehat{SE}(\hat{\mu}_{1..} - \hat{\mu}_{2..}) = \sqrt{\frac{1}{6}[MS_{L(M)}]} = \sqrt{\frac{621}{6}} = 10.1735.$

- The degrees of freedom associated with this estimate is $df_{L(M)} = 6$.

- Compute the value of Tukey-Kramer's HSD with $\alpha = .05$ that would be used to determine which pairs of means across the levels of the factor Manufacture are different:

  $HSD = (q_{.05,3,6})\widehat{SE}(\hat{\mu}_{1..} - \mu_{2..})/\sqrt{2} = (4.339)(10.1735)/\sqrt{2} = 31.213$

2. Compute the estimated standard error of the difference in the means for Speed 1 and 2, $\widehat{SE}(\hat{\mu}_{..1} - \hat{\mu}_{..2})$:

   $\bar{y}_{..1.} = \mu + \bar{\tau}_{.} + \bar{L}_{.(.)} + \beta_1 + \overline{(\tau\beta)}_{.1} + \overline{(L*\beta)}_{.(.)1} + \bar{e}_{..1.}$

   $\bar{y}_{..2.} = \mu + \bar{\tau}_{.} + \bar{L}_{.(.)} + \beta_2 + \overline{(\tau\beta)}_{.2} + \overline{(L*\beta)}_{.(.)2} + \bar{e}_{..2.} \Rightarrow$

   $Var[\bar{y}_{..1.} - \bar{y}_{..2.}] = Var\left([\bar{L}_{.(.)} - \bar{L}_{.(.)}]\right) + Var\left([\overline{(L*\beta)}_{.(.)1} - \overline{(L*\beta)}_{.(.)2}]\right) + Var\left([\bar{e}_{..1.} - \bar{e}_{..2.}]\right)$

   Therefore, we have

$$
\begin{aligned}
Var[\bar{y}_{..1.} - \bar{y}_{..2.}] &= \frac{2\sigma^2_{L(M)*S}}{(3)(3)} + \frac{2\sigma^2_e}{(3)(3)(2)} \\
&= \frac{2}{18}\left[2\sigma^2_{L(M)*S} + \sigma^2_e\right] \\
&= \frac{1}{9}\left[EMS_{L(M)*S}\right]
\end{aligned}
$$

   An estimate of the standard error and the degrees of freedom of the estimate are given as follows.

   $\widehat{SE}(\hat{\mu}_{..1} - \hat{\mu}_{..2}) = \sqrt{\frac{1}{9}[MS_{L(M)*S}]} = \sqrt{\frac{364.56}{9}} = 6.3644.$

- The degrees of freedom associated with this estimate is $df_{L(M)*S} = 6$.

- Compute the value of Tukey's HSD with $\alpha = .05$ that would be used to determine which pairs of means across the levels of the factor Speed are different:

  $HSD = (q_{.05,2,6})\widehat{SE}(\hat{\mu}_{..1} - \mu_{..2})/\sqrt{2} = (3.46)(6.3644)/\sqrt{2} = 15.571$

3. Compute the estimated standard error of the difference in the means for (Man 1,Speed 1) and (Man 2, Speed 2) $\widehat{SE}(\hat{\mu}_{1.1} - \hat{\mu}_{2.2})$:

$\bar{y}_{1.1.} = \mu + \tau_1 + \bar{L}_{.(1)} + \beta_1 + (\tau\beta)_{11} + \overline{(L * \beta)}_{.(1)1} + \bar{e}_{1.1.}$

$\bar{y}_{2.2.} = \mu + \tau_2 + \bar{L}_{.(2)} + \beta_2 + (\tau\beta)_{22} + (L * \bar{\beta})_{.(2)2} + \bar{e}_{2.2.} \Rightarrow$

$Var[\bar{y}_{1.1.} - \bar{y}_{2.2.}] = Var\left([\bar{L}_{.(1)} - \bar{L}_{.(2)}]\right) + Var\left([\overline{(L * \beta)}_{.(1)1} - \overline{(L * \beta)}_{.(2)2}]\right) + Var\left([\bar{e}_{1.1.} - \bar{e}_{2.2.}\right)$

Therefore, we have

$$\begin{aligned}
Var[\bar{y}_{1.1.} - \bar{y}_{2.2.}] &= \frac{2\sigma^2_{L(M)}}{3} + \frac{2\sigma^2_{L(M)*S}}{3} + \frac{2\sigma^2_e}{(3)(2)} \\
&= \frac{2}{6}\left[2\sigma^2_{L(M)} + 2\sigma^2_{L(M)*S} + \sigma^2_e\right] \\
&= \frac{1}{3}\left[\frac{1}{2}EMS_{L(M)} + \frac{1}{2}EMS_{L(M)*S}\right] \\
&= \frac{1}{6}\left[EMS_{L(M)} + EMS_{L(M)*S}\right]
\end{aligned}$$

An estimate of the standard error and the degrees of freedom of the estimate are given as follows.

$$\widehat{SE}(\hat{\mu}_{1.1} - \hat{\mu}_{2.2}) = \sqrt{\tfrac{1}{6}[MS_{L(M)} + MS_{L(M)*S}]} = \sqrt{\tfrac{621+364.56}{6}} = 12.8164.$$

- The degrees of freedom associated with this estimate is obtained using the Satterthwaite approximation:

$$df \approx \frac{(MS_{L(M)} + MS_{L(M)*S})^2}{\frac{(MS_{L(M)})^2}{df_{MS_{L(M)}}} + \frac{(MS_{L(M)*S})^2}{df_{MS_{L(M)*S}}}} = \frac{(621 + 364.56)^2}{\frac{(621)^2}{6} + \frac{(364.56)^2}{6}} = 11.24$$

.

- Compute the value of Tukey's HSD with $\alpha = .05$ that would be used to determine which pairs of means across the levels of the treatment combinations of (Manufacture, Speed) are different:

$$HSD = (q_{.05,6,11.24})\widehat{SE}(\hat{\mu}_{1.1} - \mu_{2.2})/\sqrt{2} = (4.804)(12.8164)/\sqrt{2} = 43.536$$

@ ≈ 45 min a week.

Finished Monday 4/11/22 (Week 12, lecture 31)