Lecture 11: Factorial Experiment with Random Effects

Montgomery, Chapter 13

Factorial Experiments with Random Effects

- Lecture 10 has focused on fixed effects

  –Always use MSE in denominator of F-test

  –Use MSE in std error of linear contrasts

- Not always correct when random factors present

  –May use interaction MS or combination of MSs

- When random effect exists, we will use $E(\text{MS.})$ as guide for tests.

A Measurement Systems Capability Study

A typical gauge R&R experiment is shown below. An instrument or gauge is used to measure a critical dimension of certain part. Twenty parts have been selected from the production process, and three randomly selected operators measure each part twice with this gauge. The order in which the measurements are made is completely randomized, so this is a two-factor factorial experiment with design factors parts and operators, with two replications. Both parts and operators are random factors.

Parts	Operator 1		Operator 2		Operator 3	
1	21	20	20	20	19	21
2	24	23	24	24	23	24
3	20	21	19	21	20	22
.	.	.	.	.	.	.
19	25	26	25	24	25	25
20	19	19	18	17	19	17

Variance components equation: $\sigma_y^2 = \sigma_\tau^2 + \sigma_\beta^2 + \sigma_{\tau\beta}^2 + \sigma^2$

Total variability=Parts + Operators + Interaction + Experimental Error

=Parts + Reproducibility + Repeatability

Statistical Model with Two Random Factors

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \epsilon_{ijk} \quad \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, b \\ k = 1, 2, \ldots, n \end{cases}$$

$$\tau_i \sim \mathrm{N}(0, \sigma_\tau^2) \qquad \beta_j \sim \mathrm{N}(0, \sigma_\beta^2) \qquad (\tau\beta)_{ij} \sim \mathrm{N}(0, \sigma_{\tau\beta}^2)$$

- Var(y_{ijk}) = $\sigma^2 + \sigma_\tau^2 + \sigma_\beta^2 + \sigma_{\tau\beta}^2$
- Expected MS's similar to one-factor random model (can be obtained by random statement in glm procedure)

  E(MS$_\mathrm{E}$)=σ^2; E(MS$_\mathrm{A}$) = $\sigma^2 + bn\sigma_\tau^2 + n\sigma_{\tau\beta}^2$

  E(MS$_\mathrm{B}$) = $\sigma^2 + an\sigma_\beta^2 + n\sigma_{\tau\beta}^2$; E(MS$_\mathrm{AB}$) = $\sigma^2 + n\sigma_{\tau\beta}^2$
- EMS determine what MS to use in denominator

  $H_0 : \sigma_\tau^2 = 0 \rightarrow \mathrm{MS_A}/\mathrm{MS_{AB}}$

  $H_0 : \sigma_\beta^2 = 0 \rightarrow \mathrm{MS_B}/\mathrm{MS_{AB}}$

  $H_0 : \sigma_{\tau\beta}^2 = 0 \rightarrow \mathrm{MS_{AB}}/\mathrm{MS_E}$

Estimating Variance Components

- Using ANOVA method

$$\hat{\sigma}^2 = \mathrm{MS_E}$$

$$\hat{\sigma}_\tau^2 = (\mathrm{MS_A} - \mathrm{MS_{AB}})/bn$$

$$\hat{\sigma}_\beta^2 = (\mathrm{MS_B} - \mathrm{MS_{AB}})/an$$

$$\hat{\sigma}_{\tau\beta}^2 = (\mathrm{MS_{AB}} - \mathrm{MS_E})/n$$

- Sometimes results in negative estimates
- Proc Varcomp and Proc Mixed compute estimates
- Can use different estimation procedures

  ANOVA method - Method = type1

  RMLE method - Method = reml (default)

- Proc Mixed

  Variance component estimates

  Hypothesis tests and confidence intervals

Gauge Capability Example in Text 13-2

```
options nocenter ls=75;

data randr;
 input part operator resp @@;
 cards;
1 1 21 1 1 20 1 2 20 1 2 20 1 3 19 1 3 21
2 1 24 2 1 23 2 2 24 2 2 24 2 3 23 2 3 24
3 1 20 3 1 21 3 2 19 3 2 21 3 3 20 3 3 22
4 1 27 4 1 27 4 2 28 4 2 26 4 3 27 4 3 28
..........................................
20 1 19 20 1 19 20 2 18 20 2 17 20 3 19 20 3 17
;

proc glm;
 class operator part;
 model resp=operator|part;
 random operator part operator*part / test;
 test H=operator E=operator*part;
 test H=part E=operator*part;
```

```
proc mixed cl maxiter=20 covtest method=type1;
 class operator part;
 model resp = ;
 random operator part operator*part;

proc mixed cl maxiter=20 covtest;
 class operator part;
 model resp = ;
 random operator part operator*part;
run;
quit;
```

Dependent Variable: resp

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	59	1215.091667	20.594774	20.77	<.0001
Error	60	59.500000	0.991667		
CorreTotal	119	1274.591667			

Source	DF	Type III SS	Mean Square	F Value	Pr > F
operator	2	2.616667	1.308333	1.32	0.2750
part	19	1185.425000	62.390789	62.92	<.0001
operator*part	38	27.050000	0.711842	0.72	0.8614

Source	Type III Expected Mean Square
operator	Var(Error) + 2 Var(operator*part) + 40 Var(operator)
part	Var(Error) + 2 Var(operator*part) + 6 Var(part)
operator*part	Var(Error) + 2 Var(operator*part)

Tests of Hypotheses Using the Type III
MS for operator*part as an Error Term

Source	DF	Type III SS	Mean Square	F Value	Pr > F

operator	2	2.616667	1.308333	1.84	0.1730
part	19	1185.425000	62.390789	87.65	<.0001

Tests of Hypotheses for Random Model Analysis of Variance

Dependent Variable: resp

Source	DF	Type III SS	Mean Square	F Value	Pr > F
operator	2	2.616667	1.308333	1.84	0.1730
part	19	1185.425000	62.390789	87.65	<.0001
Error	38	27.050000	0.711842		

Error: MS(operator*part)

Source	DF	Type III SS	Mean Square	F Value	Pr > F
operator*part	38	27.050000	0.711842	0.72	0.8614
Error: MS(Error)	60	59.500000	0.991667		

The Mixed Procedure

Type 1 Analysis of Variance

Source	DF	Sum of Squares	Mean Square
operator	2	2.616667	1.308333
part	19	1185.425000	62.390789
operator*part	38	27.050000	0.711842
Residual	60	59.500000	0.991667

Type 1 Analysis of Variance

Source	Expected Mean Square	Error Term	Error DF
operator	Var(Residual) + 2 Var(operator*part) + 40 Var(operator)	MS(operator*part)	38
part	Var(Residual) + 2 Var(operator*part) + 6 Var(part)	MS(operator*part)	38
operator*part	Var(Residual) + 2 Var(operator*part)	MS(Residual)	60
Residual	Var(Residual)	.	

Source	F Value	Pr > F
operator	1.84	0.1730
part	87.65	<.0001
operator*part	0.72	0.8614

Covariance Parameter Estimates

Cov Parm	Estimate	Standard Error	Z Value	Pr Z	Alpha	Lower	Upper
operator	0.0149	0.0330	0.45	0.6510	0.05	-0.0497	0.0795
part	10.2798	3.3738	3.05	0.0023	0.05	3.6673	16.8924
operator*part	-0.1399	0.1219	-1.15	0.2511	0.05	-0.3789	0.0990
Residual	0.9917	0.1811	5.48	<.0001	0.05	0.7143	1.4698

The Mixed Procedure

Estimation Method REML

Iteration History

Iteration	Evaluations	-2 Res Log Like	Criterion
0	1	624.67452320	
1	3	409.39453674	0.00003340
2	1	409.39128078	0.00000004
3	1	409.39127700	0.00000000

Convergence criteria met.

Covariance Parameter Estimates

Cov Parm	Estimate	Standard Error	Z Value	Pr Z	Alpha	Lower	Upper
operator	0.0106	0.03286	0.32	0.3732	0.05	0.001103	3.7E12
part	10.2513	3.3738	3.04	0.0012	0.05	5.8888	22.1549
operator*part	0	.	.	.	.	.	.
Residual	0.8832	0.1262	7.00	<.0001	0.05	0.6800	1.1938

Confidence Intervals for Variance Components

- Can use asymptotic variance estimates to form CI
- Known as Wald's approximate CI
- Mixed: option CL=WALD or METHOD=TYPE1

  Use standard normal approximation $\rightarrow$ 95% CI uses $\pm 1.96 * s.e.$

  $\hat{\sigma}^2_\beta \pm 1.96(.0330) = (-0.05, 0.08)$

  $\hat{\sigma}^2_\tau \pm 1.96(3.3738) = (3.67, 16.89)$

- In general Proc Mixed uses Satterthwaite CI

  Default method - REML

  SAS uses Satterthwaite's Approximation:

  $\sum k_i \text{MS}_i \approx E(\sum k_i \text{MS}_i)\chi^2_\nu/\nu$.

  Will discuss this detail later on.

Two-Factor Mixed Effects Model

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \epsilon_{ijk}$$

- Assume A fixed and B random
  1 $\sum \tau_i = 0$ and $\beta \sim \mathrm{N}(0, \sigma^2_\beta)$ usual assumptions
  2 $(\tau\beta)_{ij} \sim \mathrm{N}(0, (a-1)\sigma^2_{\tau\beta}/a)$ $(a-1)/a$ simplifies EMS
  3 $\sum_i (\tau\beta)_{ij} = 0$ for β level j added restriction
- Due to added restriction
  - Not all $(\tau\beta)_{ij}$ indep, Cov$((\tau\beta)_{ij}, (\tau\beta)_{i'j}) = -\frac{1}{a}\sigma^2_{\tau\beta}$
  - Cov$(y_{ijk}, y_{i'jk'}) = \sigma^2_\beta - \frac{1}{a}\sigma^2_{\tau\beta}, i \neq i'$.
- Known as **restricted** mixed effects model
- This model coincides with EMS algorithm

  E(MS$_{\mathrm{E}}$)=σ^2

  E(MS$_{\mathrm{A}}$) = $\sigma^2 + bn\sum \tau_i^2/(a-1) + n\sigma^2_{\tau\beta}$

  E(MS$_{\mathrm{B}}$) = $\sigma^2 + an\sigma^2_\beta$

  E(MS$_{\mathrm{AB}}$) = $\sigma^2 + n\sigma^2_{\tau\beta}$

Hypotheses Testing and Diagnostics

- Testing hypotheses:

$$H_0 : \tau_1 = \tau_2 = \ldots = 0 \rightarrow \mathrm{MS_A}/\mathrm{MS_{AB}}$$

$$H_0 : \sigma^2_\beta = 0 \rightarrow \mathrm{MS_B}/\mathrm{MS_E}$$

$$H_0 : \sigma^2_{\tau\beta} = 0 \rightarrow \mathrm{MS_{AB}}/\mathrm{MS_E}$$

- Variance Estimates (Using ANOVA method)

$$\hat{\sigma}^2 = \mathrm{MS_E}$$

$$\hat{\sigma}^2_\beta = (\mathrm{MS_B} - \mathrm{MS_E})/an$$

$$\hat{\sigma}^2_{\tau\beta} = (\mathrm{MS_{AB}} - \mathrm{MS_E})/n$$

- Diagnostics
  - Histogram or QQplot

    Normality or Unusual Observations
  - Residual Plots

    Constant variance or Unusual Observations

Multiple Comparisons for Fixed Effects

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \epsilon_{ijk}$$

$$\overline{y}_{i..} = \mu + \tau_i + \overline{\beta}_{.} + \overline{(\tau\beta)}_{i.} + \overline{\epsilon}_{i..}$$

$$\text{Var}(\overline{y}_{i..}) = \sigma^2_\beta/b + (a-1)\sigma^2_{\tau\beta}/ab + \sigma^2/bn$$

$$\overline{y}_{i..} - \overline{y}_{i'..} = \tau_i - \tau_{i'} + \overline{(\tau\beta)}_{i.} - \overline{(\tau\beta)}_{i'.} + \overline{\epsilon}_{i..} - \overline{\epsilon}_{i'..}$$

$$\begin{aligned}\text{Var}(\overline{y}_{i..} - \overline{y}_{i'..}) &= 2\sigma^2_{\tau\beta}/b + 2\sigma^2/bn \\ &= 2(n\sigma^2_{\tau\beta} + \sigma^2)/bn\end{aligned}$$

- Need to plug in variance estimates to compute Var($\overline{y}_{i..}$)
- What are the DF?
- For pairwise comparisons, use estimate $2\text{MS}_{AB}/bn$
- Use df$_{AB}$ for t-statistic or Tukey's method.

Gauge Capability Example in Text 12-3

```
options nocenter ls=75;

data randr;
 input part operator resp @@;
 cards;
1 1 21 1 1 20 1 2 20 1 2 20 1 3 19 1 3 21
2 1 24 2 1 23 2 2 24 2 2 24 2 3 23 2 3 24
3 1 20 3 1 21 3 2 19 3 2 21 3 3 20 3 3 22
4 1 27 4 1 27 4 2 28 4 2 26 4 3 27 4 3 28
.........................................
20 1 19 20 1 19 20 2 18 20 2 17 20 3 19 20 3 17
;
proc glm;
 class operator part;
 model resp=operator|part;
run;
==================================================
Dependent Variable: resp
                              Sum of
```

Source	DF	Squares	Mean Square	F Value	Pr > F
Model	59	1215.091667	20.594774	20.77	<.0001
Error	60	59.500000	0.991667		
CorrTotal	119	1274.591667			

Source	DF	Type I SS	Mean Square	F Value	Pr > F
operator	2	2.616667	1.308333	1.32	0.2750
part	19	1185.425000	62.390789	62.92	<.0001
operator*part	38	27.050000	0.711842	0.72	0.8614

Gauge Capability Example

- $H_0 : \tau_1 = \tau_2 = \tau_3 = 0$:

$$F_0 = \frac{MS_A}{MS_{AB}} = \frac{1.308}{0.712} = 1.84$$

P-value based on $F_{2,38}$: 0.173.

- $H_0 : \sigma_\beta^2 = 0$:

$$F_0 = \frac{MS_B}{MS_E} = \frac{62.391}{0.992} = 62.89$$

P-value based on $F_{19,60}$: 0.000

- H_0: $\sigma^2_{\tau\beta} = 0$:

$$F_0 = \frac{MS_{AB}}{MS_E} = \frac{0.712}{0.992} = 0.72$$

P-value based on $F_{38,60}$: 0.86

- Variance components estimates:

$$\hat{\sigma}_\beta^2 = \frac{62.39 - 0.99}{(3)(2)} = 10.23$$

$$\hat{\sigma}^2_{\tau\beta} = \frac{0.71 - 0.99}{2} = -.14(\approx 0)$$

$$\hat{\sigma}^2 = 0.99$$

- Pairwise comparison for τ_1, τ_2 and τ_3.

  Bonferroni:

$$CD : t_{1-0.05/2/3,38}\sqrt{MS_{AB}(2/bn)}$$

  Tukey:

$$CD : \frac{q_{0.05}(3,38)}{\sqrt{2}}\sqrt{MS_{AB}(2/bn)}$$

Unrestricted Mixed Model

- SAS uses **unrestricted mixed model** in analysis
- $y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \epsilon_{ijk}$

  $\sum \tau_i = 0$ and $\beta_j \sim \mathrm{N}(0, \sigma^2_\beta)$

  $(\tau\beta)_{ij} \sim \mathrm{N}(0, \sigma^2_{\tau\beta})$
- Expected mean squares:

  E(MS$_{\mathrm{E}}$)=σ^2

  E(MS$_{\mathrm{A}}$) = $\sigma^2 + bn\sum \tau_i^2/(a-1) + n\sigma^2_{\tau\beta}$

  E(MS$_{\mathrm{B}}$) = $\sigma^2 + an\sigma^2_\beta + n\sigma^2_{\tau\beta}$

  E(MS$_{\mathrm{AB}}$) = $\sigma^2 + n\sigma^2_{\tau\beta}$
- random statement in SAS also gives these results
- Differences
  - E(MS$_{\mathrm{B}}$)
  - Test $H_0 : \sigma^2_\beta = 0$ using MS$_{\mathrm{AB}}$ in denominator
  - Cov($y_{ijk}, y_{i'jk'}$)=σ^2_β, $i \neq i'$.

- $\hat{\sigma}_\beta^2 = (\text{MS}_B - \text{MS}_{AB})/an$.
- $Var(\bar{y}_{i..}) = (n\sigma_\beta^2 + n\sigma_{\tau\beta}^2 + \sigma^2)/(bn)$ although $Var(\bar{y}_{i..} - \bar{y}_{j..})$ keeps the same

- Connection

$$\bar{(\tau\beta)}_{.j} = (\sum_i (\tau\beta)_{ij})/a$$

$$y_{ijk} = \mu + \tau + (\beta_j + \bar{(\tau\beta)}_{.j}) + ((\tau\beta)_{ij} - \bar{(\tau\beta)}_{.j}) + \epsilon_{ijk}$$

Check the model above satisfies most of the conditions of restricted mixed model

- Restricted model is slightly more general.
- It is difficult to provide guidelines for when the restricted or unrestricted mixed model should be used, because statisticians do not fully agree on this.

```
                    Gauge Capability Example (Unrestricted Model)

options nocenter ls=75;

data randr;
 input part operator resp @@;
 cards;
1 1 21 1 1 20 1 2 20 1 2 20 1 3 19 1 3 21
2 1 24 2 1 23 2 2 24 2 2 24 2 3 23 2 3 24
3 1 20 3 1 21 3 2 19 3 2 21 3 3 20 3 3 22
4 1 27 4 1 27 4 2 28 4 2 26 4 3 27 4 3 28
.
.
;

proc glm;
 class operator part;
 model resp=operator|part;
 random part operator*part / test;
 means operator / tukey lines E=operator*part;
 lsmeans operator / adjust=tukey E=operator*part tdiff stderr;
```

```
proc mixed alpha=.05 cl covtest;
 class operator part;
 model resp=operator / ddfm=kr;
 random part operator*part;
 lsmeans operator / alpha=.05 cl diff adjust=tukey;
run;
quit;
```

Source	Type III Expected Mean Square
operator	Var(Error) + 2 Var(operator*part) + Q(operator)
part	Var(Error) + 2 Var(operator*part) + 6 Var(part)
operator*part	Var(Error) + 2 Var(operator*part)

Dependent Variable: resp

Source	DF	Type III SS	Mean Square	F Value	Pr > F
operator	2	2.616667	1.308333	1.84	0.1730
part	19	1185.425000	62.390789	87.65	<.0001
Error	38	27.050000	0.711842		

Error: MS(operator*part)

Source	DF	Type III SS	Mean Square	F Value	Pr > F
operator*part	38	27.050000	0.711842	0.72	0.8614
Error: MS(Error)	60	59.500000	0.991667		

Tukeys Studentized Range (HSD) Test for resp

Alpha	0.05
Error Degrees of Freedom	38
Error Mean Square	0.711842
Critical Value of Studentized Range	3.44902

```
Minimum Significant Difference        0.4601

Means with the same letter are not significantly different.
Grouping Mean      N   operator
A        22.6000  40  3
A        22.3000  40  1
A        22.2750  40  2
```

Covariance Parameter Estimates

Cov Parm	Estimate	Standard Error	Z Value	Pr > Z	Alpha	Lower	Upper
part	10.2513	3.3738	3.04	0.0012	0.05	5.8888	22.1549
operator*part	0	.	.	.	.	.	.
Residual	0.8832	0.1262	7.00	<.0001	0.05	0.6800	1.1938

Effect	Num DF	Den DF	F Value	Pr > F
operator	2	98	1.48	0.2324

Least Squares Means

Effect	operator	Estimate	Standard Error	DF	t Value	Pr>|t|	Alpha	Lower	Upper
operator	1	22.3000	0.7312	20.1	30.50	<.0001	0.05	20.7752	23.8248
operator	2	22.2750	0.7312	20.1	30.46	<.0001	0.05	20.7502	23.7998
operator	3	22.6000	0.7312	20.1	30.91	<.0001	0.05	21.0752	24.1248

Differences of Least Squares Means

Effect	operator	_operator	Estimate	Error	DF	t Value	Pr > |t|	Adjustment
operator	1	2	0.02500	0.2101	98	0.12	0.9055	Tukey-Kramer
operator	1	3	-0.3000	0.2101	98	-1.43	0.1566	Tukey-Kramer
operator	2	3	-0.3250	0.2101	98	-1.55	0.1252	Tukey-Kramer

Differences of Least Squares Means

Effect	Adj P	Alpha	Lower	Upper	Adj Lower	Adj Upper
operator	0.9922	0.05	-0.3920	0.4420	-0.4751	0.5251
operator	0.3308	0.05	-0.7170	0.1170	-0.8001	0.2001
operator	0.2739	0.05	-0.7420	.09201	-0.8251	0.1751

Rules For Expected Mean Squares

- In models so far, EMS fairly straightforward
- Could calculate EMS using brute force method
- For mixed models, good to have formal procedure
- Montgomery describes procedure for **restricted** model

  0 Write the error term in the model as $\epsilon_{(ij..)m}$, where m represents the replication subscript

  1 Write each variable term in model as a row heading in a two-way table

  2 Write the subscripts in the model as column headings. Over each subscript write F if factor fixed and R if random. Over this, write down the levels of each subscript

  3 For each row, copy the number of observations under each subscript, providing the subscript does not appear in the row variable term

  4 For any bracketed subscripts in the model, place a 1 under those subscripts that are inside the brackets

  5 Fill in remaining cells with a 0 (if subscript represents a fixed factor) or a 1 (if random factor).

6 To find the expected mean square of any term (row), cover the entries in the columns that contain non-bracketed subscript letters in this term in the model. For those rows including the same subscripts, multiply the remaining numbers to get coefficient for corresponding term in the model.

For random effect, multiply $\sigma^2_{(\cdot)}$; For fixed effect, multipy $Q(\cdot) = \sum(\cdot)^2/df$

Two-Factor Fixed Model:

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \epsilon_{ijk}$$

	F	F	R	
	a	b	n	
term	i	j	k	EMS
τ_i	0	b	n	$\sigma^2 + \frac{bn\Sigma\tau_i^2}{a-1}$
β_j	a	0	n	$\sigma^2 + \frac{an\Sigma\beta_j^2}{b-1}$
$(\tau\beta)_{ij}$	0	0	n	$\sigma^2 + \frac{n\Sigma\Sigma(\tau\beta)_{ij}^2}{(a-1)(b-1)}$
$\epsilon_{(ij)k}$	1	1	1	σ^2

Two-Factor Random Model:

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \epsilon_{ijk}$$

	R	R	R	
	a	b	n	
term	i	j	k	EMS
τ_i	1	b	n	$\sigma^2 + n\sigma^2_{\tau\beta} + bn\sigma^2_{\tau}$
β_j	a	1	n	$\sigma^2 + n\sigma^2_{\tau\beta} + an\sigma^2_{\beta}$
$(\tau\beta)_{ij}$	1	1	n	$\sigma^2 + n\sigma^2_{\tau\beta}$
$\epsilon_{(ij)k}$	1	1	1	σ^2

Two-Factor Mixed Model (A Fixed):

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \epsilon_{ijk}$$

	F	R	R	
	a	b	n	
term	i	j	k	EMS
τ_i	0	b	n	$\sigma^2 + n\sigma^2_{\tau\beta} + \frac{bn\Sigma\tau_i^2}{a-1}$
β_j	a	1	n	$\sigma^2 + an\sigma^2_{\beta}$
$(\tau\beta)_{ij}$	0	1	n	$\sigma^2 + n\sigma^2_{\tau\beta}$
$\epsilon_{(ij)k}$	1	1	1	σ^2

Three-Factor Mixed Model (A Fixed):

$$y_{ijkl} = \mu + \tau_i + \beta_j + \gamma_k + (\tau\beta)_{ij} + (\tau\gamma)_{ik} + (\beta\gamma)_{jk} + (\tau\beta\gamma)_{ijk} + \epsilon_{ijkl}$$

	F	R	R	R	
	a	b	c	n	
term	i	j	k	l	EMS
τ_i	0	b	c	n	$\sigma^2 + cn\sigma^2_{\tau\beta} + bn\sigma^2_{\tau\gamma} + n\sigma^2_{\tau\beta\gamma} + \frac{bcn\Sigma\tau_i^2}{a-1}$
β_j	a	1	c	n	$\sigma^2 + an\sigma^2_{\beta\gamma} + acn\sigma^2_{\beta}$
γ_k	a	b	1	n	$\sigma^2 + an\sigma^2_{\beta\gamma} + abn\sigma^2_{\gamma}$
$(\tau\beta)_{ij}$	0	1	c	n	$\sigma^2 + n\sigma^2_{\tau\beta\gamma} + cn\sigma^2_{\tau\beta}$
$(\tau\gamma)_{ik}$	0	b	1	n	$\sigma^2 + n\sigma^2_{\tau\beta\gamma} + bn\sigma^2_{\tau\gamma}$
$(\beta\gamma)_{jk}$	a	1	1	n	$\sigma^2 + an\sigma^2_{\beta\gamma}$
$(\tau\beta\gamma)_{ijk}$	0	1	1	n	$\sigma^2 + n\sigma^2_{\tau\beta\gamma}$
ϵ_{ijkl}	1	1	1	1	σ^2

Rules For Expected Mean Squares (Unrestricted Model)

- Replace Step 5 with the following step

5' For any (interactive) model term (row) including a subscript for a random factor, place a 1 in the remaining cells of this row; and fill in remaining cells with a 0 (if column subscript represents a fixed factor) or a 1 (if random factor)

	F	R	R	
	a	b	n	
term	i	j	k	EMS
τ_i	0	b	n	$\sigma^2 + n\sigma^2_{\tau\beta} + \frac{bn\Sigma\tau_i^2}{a-1}$
β_j	a	1	n	$\sigma^2 + an\sigma^2_{\beta} + n\sigma^2_{\tau\beta}$
$(\tau\beta)_{ij}$	1	1	n	$\sigma^2 + n\sigma^2_{\tau\beta}$
$\epsilon_{(ij)k}$	1	1	1	σ^2

Approximate F Tests

- For some models, no exact F-test exists
- Recall 3 Factor Mixed Model (A - fixed)
- No exact test for A based on EMS

Assume $a = 3, b = 2, c = 3, n = 2$ and following MS were obtained

Source	DF	MS	EMS	F	P
A	2	0.7866	$12\phi_A + 6\sigma^2_{AB} + 4\sigma^2_{AC}$ $+2\sigma^2_{ABC} + \sigma^2$	?	?
B	1	0.0010	$18\sigma^2_B + 6\sigma^2_{BC} + \sigma^2$	0.33	.622
AB	2	0.0056	$6\sigma^2_{AB} + 2\sigma^2_{ABC} + \sigma^2$	2.24	.222
C	2	0.0560	$12\sigma^2_C + 6\sigma^2_{BC} + \sigma^2$	18.87	.051
AC	4	0.0107	$4\sigma^2_{AC} + 2\sigma^2_{ABC} + \sigma^2$	4.28	.094
BC	2	0.0030	$6\sigma^2_{BC} + \sigma^2$	10.00	.001
ABC	4	0.0025	$2\sigma^2_{ABC} + \sigma^2$	8.33	.001
Error	18	0.0003	σ^2		

- Possible approaches:
  - Could assume some variances are negligible, not recommended without “conclusive” evidence
  - Pool (insignificant) means squares with error, also risky, not recommended when df for error is already big.

Satterthwaite's Approximate F-test

- H_0 : effect $= 0$, e.g., $H_0 : \tau_1 = \cdots = \tau_a = 0$ or equivalently $H_0 : \sum \tau_i^2 = 0$.
  No exact test exists.

- Get two linear combinations of mean squares

  $$MS' = \mathrm{MS}_r \pm \ldots \pm \mathrm{MS}_s$$

  $$MS'' = \mathrm{MS}_u \pm \ldots \pm \mathrm{MS}_v$$

  such that 1) MS' and MS'' do not share common mean squares; 2) $E(MS') - E(MS'')$ is a multiple of the effect.

  – approximate test statistic F: $F = \frac{MS'}{MS''} = \frac{\mathrm{MS}_r \pm \ldots \pm \mathrm{MS}_s}{\mathrm{MS}_u \pm \ldots \pm \mathrm{MS}_v} \approx F_{p,q}$

    where $p = \frac{(\mathrm{MS}_r \pm \ldots \pm \mathrm{MS}_s)^2}{\mathrm{MS}_r^2/f_r + \ldots + \mathrm{MS}_s^2/f_s}$ and $q = \frac{(\mathrm{MS}_u \pm \ldots \pm \mathrm{MS}_v)^2}{\mathrm{MS}_u^2/f_u + \ldots + \mathrm{MS}_v^2/f_v}$

  – f_i is the degrees of freedom associated with MS_i

  – p and q may not be integers, interpolation is needed. SAS can handle noninteger dfs.

  – Caution when subtraction is used

Example: 3-Factor Mixed Model (A Fixed)

$H_0 : \tau_1 = \tau_2 = \tau_3 = 0$

$MS' = \text{MS}_A$

$MS'' = \text{MS}_{AB} + \text{MS}_{AC} - \text{MS}_{ABC}$

$E(MS' - MS'') = 12\phi_A = 12\frac{\Sigma\tau_i^2}{3-1}$

$$F = \frac{\text{MS}_A}{\text{MS}_{AB} + \text{MS}_{AC} - \text{MS}_{ABC}} = \frac{.7866}{.0107 + .0056 - .0025} = 57.0$$

$$p = 2 \qquad q = \frac{.0138^2}{.0107^2/4 + .0056^2/2 + .0025^2/4} = 4.15$$

- Interpolation needed

$$\text{P}(F_{2,4} > 57) = .0011 \qquad \text{P}(F_{2,5} > 57) = .0004$$

$$\text{P} = .85(.0011) + .15(.0004) = .001$$

- SAS can be used to compute P-values and quantile values for F and χ^2 values with noninteger degrees of freedom.

  Upper Tail Probability: probf(x,df1,df2) and probchi(x,df)

  Quantiles : finv(p,df1,df2) and cinv(p,df)

```
data one;
  p=1-probf(57,2.0,4.15);
  f=finv(.95,2.0,4.15);
  c1=cinv(.025,18.57);
  c2=cinv(.975,18.57);
proc print data=one;

OBS        P          F          C1         C2
 1      .00096     6.7156     8.61485    32.2833
```

Another Approach to Testing $H_0 : \tau_1 = \tau_2 = \tau_3 = 0$

$MS' = \mathrm{MS}_A + \mathrm{MS}_{ABC}$

$MS'' = \mathrm{MS}_{AB} + \mathrm{MS}_{AC}$

$E(MS' - MS'') =?$

$$F = \frac{\mathrm{MS}_A + \mathrm{MS}_{ABC}}{\mathrm{MS}_{AB} + \mathrm{MS}_{AC}} = \frac{.7866 + .0025}{.0107 + .0056} = 48.41$$

$$p = \frac{.7891^2}{.7866^2/2 + .0025^2/4} = 2.01 \qquad q = \frac{.0163^2}{.0107^2/4 + .0056^2/2} = 6.00$$

P-value= $P(F > 48.41) = 0.002$

- This is again found significant
- Avoid subtraction,summation should be preferred.

Approximate Confidence Intervals

Suppose we are interested in σ_x^2.

- Case 1: there exists a mean square MS_x with df_x such that $E(MS_x) = \sigma_x^2$. Then $\hat{\sigma}_x^2 = MS_x$, and

$$\frac{\text{df}_x MS_x}{\sigma_x^2} \sim \chi^2(\text{df}_x)$$

  Exact 100(1-α)% CI: $\frac{\mathrm{df}_x MS_x}{\chi^2_{\alpha/2,\mathrm{df}_x}} \leq \sigma_x^2 \leq \frac{\mathrm{df}_x MS_x}{\chi^2_{1-\alpha/2,\mathrm{df}_x}}$

- Case 2: there exist

  $MS' = \mathrm{MS}_r + \ldots + \mathrm{MS}_s$ and, $MS'' = \mathrm{MS}_u + \ldots + \mathrm{MS}_v$

  such that $E(MS' - MS'') = k\sigma_x^2$. Then

$$\hat{\sigma}_x^2 = \frac{MS' - MS''}{k}, \text{ and } \frac{\text{df}_x \hat{\sigma}_x^2}{\sigma_x^2} \approx \chi^2(\text{df}_x)$$

  where

$$\text{df}_x = \frac{(\hat{\sigma}_x^2)^2}{\sum \frac{MS_i}{k^2 f_i}} = \frac{(\mathrm{MS}_r + \cdot + \mathrm{MS}_s - \mathrm{MS}_u - \cdot - \mathrm{MS}_v)^2}{\mathrm{MS}_r^2/f_r + \cdot + \mathrm{MS}_s^2/f_s + \mathrm{MS}_u^2/f_u + \cdot + \mathrm{MS}_v^2/f_v}$$

Approximate 100(1-α)% CI:

$$\frac{\mathrm{df}_x \hat{\sigma}_x^2}{\chi^2_{\alpha/2,\mathrm{df}_x}} \le \sigma_x^2 \le \frac{\mathrm{df}_x \hat{\sigma}_x^2}{\chi^2_{1-\alpha/2,\mathrm{df}_x}}$$

Gauge Capability Example (Both Factors are Random)

Dependent Variable: RESP

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	59	1215.09166667	20.594774	20.77	0.0001
Error	60	59.50000000	0.991667		
CorrecTotal	119	1274.59166667			

Source	DF	Type III SS	Mean Square	F Value	Pr > F
OPERATOR	2	2.61666667	1.308333	1.32	0.2750
PART	19	1185.42500000	62.390789	62.92	0.0001
OPERATOR*PART	38	27.05000000	0.711842	0.72	0.8614

$$\hat{\sigma}_\tau^2 = \frac{MS_A - MS_{AB}}{bn} = (62.39 - 0.71)/6 = 10.28$$

$$\text{df} = \frac{(62.39 - 0.71)^2}{62.39^2/19 + 0.71^2/38} = 18.57$$

CI: $(18.57(10.28)/32.28, 18.57(10.28)/8.61) = (5.91, 22.17)$

$\hat{\sigma}^2_\beta = \frac{MS_B - MS_{AB}}{an} = (1.31 - 0.71)/40 = 0.015$

$$\text{df} = \frac{(1.31 - 0.71)^2}{1.31^2/2 + 0.71^2/38} = .413$$

CI: $(.413(.015)/3.079, .413(.015)/2.29 \times 10^{-8}) = (.002, 270781)$

General Mixed Effect Model

- In terms of linear model

$$Y = X\beta + Z\delta + \epsilon$$

  β is a vector of fixed-effect parameters

  δ is a vector of random-effect parameters

  ϵ is the error vector

- δ and ϵ assumed uncorrelated
  - means 0
  - covariance matrices G and R (allows correlation)
- Cov(Y) = $ZGZ' + R$
- If $R = \sigma^2 I$ and $Z = 0$, back to standard linear model
- SAS Proc Mixed allows one to specify G and R
- G through RANDOM, R through REPEATED
- Unrestricted linear mixed model is default

Model Size determination

- Recall sample size calculations on a hypothesis test of a set of effects using

$$F_0 = \frac{MS_N}{MS_D} \sim^{H_0} F_{\nu_1,\nu_2},$$

- If fixed effect, the noncentral parameter:

$$\delta = \frac{E(MS_N - MS_D) * \nu_1}{E(MS_D)}$$

power = $1 - probf(F_{\alpha,\nu_1,\nu_2}, \nu_1, \nu_2, \delta)$

- If random effect:

$\nu_1 MS_N = \chi^2_{\nu_1} * E(MS_N), \nu_2 MS_D = \chi^2_{\nu_2} * E(MS_D)$

The scaling parameter $\lambda^2 = E(MS_N)/E(MS_D)$

power = $1 - probf(F_{\alpha,\nu_1,\nu_2}/\lambda^2, \nu_1, \nu_2)$

Mixed Model

restricted model

Factor	Parameter	ν_1	ν_2
A(Fixed)	$\delta = \frac{bn\sum\tau_i^2}{\sigma^2+n\sigma_{\tau\beta}^2}$	$a-1$	$(a-1)(b-1)$
B(Random)	$\lambda^2 = 1 + \frac{an\sigma_\beta^2}{\sigma^2}$	$b-1$	$ab(n-1)$
AB(Random)	$\lambda^2 = 1 + \frac{n\sigma_{\tau\beta}^2}{\sigma^2}$	$(a-1)(b-1)$	$ab(n-1)$

unrestricted model

Factor	Parameter	ν_1	ν_2
A(Fixed)	$\delta = \frac{bn\sum\tau_i^2}{\sigma^2+n\sigma_{\tau\beta}^2}$	$a-1$	$(a-1)(b-1)$
B(Random)	$\lambda^2 = 1 + \frac{an\sigma_\beta^2}{\sigma^2+n\sigma_{\tau\beta}^2}$	$b-1$	$(a-1)(b-1)$
AB(Random)	$\lambda^2 = 1 + \frac{n\sigma_{\tau\beta}^2}{\sigma^2}$	$(a-1)(b-1)$	$ab(n-1)$

Random Model

Factor	Parameter	ν_1	ν_2
A(Random)	$\lambda^2 = 1 + \frac{bn\sigma_\tau^2}{\sigma^2+n\sigma_{\tau\beta}^2}$	$a-1$	$(a-1)(b-1)$
B(Random)	$\lambda^2 = 1 + \frac{an\sigma_\beta^2}{\sigma^2+n\sigma_{\tau\beta}^2}$	$b-1$	$ab(n-1)$
AB(Random)	$\lambda^2 = 1 + \frac{n\sigma_{\tau\beta}^2}{\sigma^2}$	$(a-1)(b-1)$	$ab(n-1)$