Lecture 10: Factorial Designs with Random Factors

Montgomery, Section 13.2 and 13.3

Factorial Experiments with Random Effects

- Lecture 9 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 MS's
- Will now use EMS as guide for tests

Two-Factor Mixed Effects Model

- One factor random and one factor fixed (aka Model III)
- Assume A fixed and B random
- Mixed Factor Effects Model:

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

  - $\sum_i \tau_i = 0$ and $\beta_j \overset{iid}{\sim} \mathrm{N}(0, \sigma^2_\beta)$
  - $(\tau\beta)_{ij} \sim \mathrm{N}(0, (a-1)\sigma^2_{\tau\beta}/a)$ subject to the restrictions $\sum_i(\tau\beta)_{ij} = 0$ for each j
  - $\varepsilon_{ijk} \overset{iid}{\sim} N(0, \sigma^2)$
  - $\{\beta_j\}$, $\{(\tau\beta)_{ij}\}$ and $\{\varepsilon_{ijk}\}$ are pairwise independent

- Known as **restricted** mixed effects model

- Not all $(\tau\beta)_{ij}$ are independent

$$\text{Cov}((\tau\beta)_{ij}, (\tau\beta)_{i'j}) = -\frac{1}{a}\sigma^2_{\tau\beta}, \quad i \neq i'$$

  - If $X_i \overset{iid}{\sim} N(\mu, \sigma^2)$ then

$$X_i - \overline{X} \sim N(0, \frac{n-1}{n}\sigma^2)$$
$$\text{Cov}(X_i - \overline{X}, X_j - \overline{X}) = -\frac{1}{n}\sigma^2$$

- The $(a-1)/a$ simplifies the EMS
  - E(MSE)=σ^2
  - E(MSA) = $\sigma^2 + bn\sum_i \tau_i^2/(a-1) + n\sigma^2_{\tau\beta}$
  - E(MSB) = $\sigma^2 + an\sigma^2_{\beta}$
  - E(MSAB) = $\sigma^2 + n\sigma^2_{\tau\beta}$

Hypotheses Testing and Diagnostics

- Hypothesis tests require different MS terms in the denominators

$$H_0 : \tau_1 = \tau_2 = \ldots = 0 \rightarrow \mathrm{MSA}/\mathrm{MSAB}$$
$$H_0 : \sigma^2_\beta = 0 \rightarrow \mathrm{MSB}/\mathrm{MSE}$$
$$H_0 : \sigma^2_{\tau\beta} = 0 \rightarrow \mathrm{MSAB}/\mathrm{MSE}$$

- Variance Estimates (Using ANOVA method)

$$\hat{\sigma}^2 = \mathrm{MSE}$$
$$\hat{\sigma}^2_\beta = (\mathrm{MSB} - \mathrm{MSE})/(an)$$
$$\hat{\sigma}^2_{\tau\beta} = (\mathrm{MSAB} - \mathrm{MSE})/n$$

- Diagnostics
  - Histogram or QQplot

    Normality or Unusual Observations
  - Residual Plots

    Constant variance or Unusual Observations

Estimates & Multiple Comparisons for Fixed Effects

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

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

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

- $\hat{\tau}_i = \bar{y}_{i..} - \bar{y}_{...}$

  $var(\hat{\tau}_i) =?$

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

  $Var(\hat{\tau}_i - \hat{\tau}_{i'}) = 2\sigma^2_{\tau\beta}/b + 2\sigma^2/bn = 2(n\sigma^2_{\tau\beta} + \sigma^2)/bn$

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

The Measurement Systems Capability Experiment (Example 13.3 in Text)

```
options nocenter ls=75;
data gaugerr;
  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 data=gaugerr;
  class operator part;
  model resp=operator|part;
  means operator;
  run; quit;
```

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		
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

Level of operator	N	------------resp------------ Mean	Std Dev
1	40	22.3000000	3.17199282
2	40	22.2750000	3.37401458
3	40	22.6000000	3.34203991

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

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

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: $\sigma_{\tau\beta}^2 = 0$:

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

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

- Variance components estimates:

$$\hat{\sigma}_\beta^2 = \frac{62.39 - 0.99}{(3)(2)} = 10.23, \hat{\sigma}_{\tau\beta}^2 = \frac{0.71 - 0.99}{2} = -.14(\approx 0), \hat{\sigma}^2 = 0.99$$

- Pairwise comparison for τ_1, τ_2 and τ_3

  – $(\bar{Y}_{i..} - \bar{Y}_{i'..})/\sqrt{2\text{MS}_{AB}/bn} \overset{H_0}{\sim} t_{(a-1)(b-1)}$; ($t_{1-0.05/(2*3),38} = 2.5046$)

i	i'	$\frac{\bar{Y}_{i..} - \bar{Y}_{i'..}}{\sqrt{2\text{MS}_{AB}/bn}}$	t
1	2	$(22.3 - 22.275)/\sqrt{2 \times 0.7118/(20 \times 2)}$	0.1325
1	3	$(22.3 - 22.6)/\sqrt{2 \times 0.7118/(20 \times 2)}$	-1.5898
2	3	$(22.275 - 22.6)/\sqrt{2 \times 0.7118/(20 \times 2)}$	-1.7223

  – Tukey's method uses $q_\alpha(a, (a-1)(b-1))$: $q_{0.05}(3, 38)/\sqrt{2} \approx 2.45$.

```
proc glm data=gaugerr;
  class operator part;
  model resp=operator|part;
  random part operator*part;
  test H=operator E=operator*part;
  lsmeans operator / adjust=tukey E=operator*part tdiff; /* or means */
  run; quit;
```

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		
Corrected Total	119	1274.591667			

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

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	2.61666667	1.30833333	1.84	0.1730

```
Least Squares Means
Adjustment for Multiple Comparisons: Tukey

Standard Errors and Probabilities Calculated Using the Type III MS for
operator*part as an Error Term

                                  LSMEAN
operator     resp LSMEAN         Number
1             22.3000000              1
2             22.2750000              2
3             22.6000000              3

   Least Squares Means for Effect operator
   t for H0: LSMean(i)=LSMean(j) / Pr > |t|

i/j               1             2             3
   1                     0.132514      -1.59017
                           0.9904        0.2622
   2       -0.13251                    -1.72269
             0.9904                      0.2100
   3       1.590173      1.722688
             0.2622        0.2100
```

- NOTE: `PROC GLM` does NOT use a restricted mixed model

Alternate Two-Factor Mixed Effects Model

- Reduce the restrictions on $(\tau\beta)_{ij} \Longrightarrow$ **unrestricted** mixed model
  - $\sum_i \tau_i = 0$ and $\beta_j \overset{iid}{\sim} \mathrm{N}(0, \sigma^2_\beta)$
  - $(\tau\beta)_{ij} \sim \mathrm{N}(0, \sigma^2_{\tau\beta})$
  - $\varepsilon_{ijk} \overset{iid}{\sim} N(0, \sigma^2)$
  - $\{\beta_j\}$, $\{(\tau\beta)_{ij}\}$ and $\{\varepsilon_{ijk}\}$ are pairwise independent
- SAS uses unrestricted mixed model in analysis
- Connection to Restricted Mixed Model: letting $\overline{(\tau\beta)}_{.j} = (\sum_i(\tau\beta)_{ij})/a$

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

– The above model satisfies the conditions of restricted mixed model

- Restricted mixed model is slightly more general since $cov(Y_{ij}, Y_{i'j}) \lesseqqgtr 0$

– $cov(Y_{ij}, Y_{i'j}) \geq 0$ in unrestricted mixed model.

Two-Factor Unrestricted Mixed Model

- Reduced restrictions alter EMS
  - E(MSE)=σ^2
  - E(MSA) = $\sigma^2 + bn\sum\tau_i^2/(a-1) + n\sigma^2_{\tau\beta}$
  - E(MSB) = $\sigma^2 + an\sigma^2_\beta + n\sigma^2_{\tau\beta}$
  - E(MSAB) = $\sigma^2 + n\sigma^2_{\tau\beta}$
- `RANDOM` statement in SAS also gives these results
- Differences
  - Test $H_0 : \sigma^2_\beta = 0$ using MSAB (Note: MSE in Restricted Models)
  - Often more conservative test
  - $\hat{\sigma}^2_\beta = (\mathrm{MSB} - \mathrm{MSAB})/(an)$
  - Var($\overline{Y}_{i..}$) = $(n\sigma^2_\beta + n\sigma^2_{\tau\beta} + \sigma^2)/(bn)$ though Var($\overline{y}_{i..} - \overline{y}_{i'..}$) = $2(n\sigma^2_{\tau\beta} + \sigma^2)/bn$
- To decide which model is appropriate, suppose you ran experiment again and sampled some of the same levels of the random effect. Does this mean that the interaction effects for these levels are the same as before? Yes: Restricted No: Unrestricted

The Measurement Systems Capability Experiment (Unrestricted Model)

```
proc glm data=gaugerr;
  class operator part;
  model resp=operator|part;
  random part operator*part / test;
  means operator / tukey E=operator*part cldiff;
  /* lsmeans operator / adjust=tukey E=operator*part tdiff; */
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)
```

- Use `test` option in `random` statement to request the correct F tests for unrestricted mixed models.

```
Tests of Hypotheses for Mixed 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
```

```
Tukey's Studentized Range (HSD) Test for resp

NOTE: This test controls the Type I experimentwise error rate.

Alpha                                     0.05
Error Degrees of Freedom                    38
Error Mean Square                     0.711842
Critical Value of Studentized Range    3.44901
Minimum Significant Difference          0.4601

Comparisons significant at the 0.05 level are indicated by ***.

              Difference      Simultaneous
 operator        Between     95% Confidence
Comparison         Means         Limits
  3 - 1           0.3000    -0.1601  0.7601
  3 - 2           0.3250    -0.1351  0.7851
  1 - 3          -0.3000    -0.7601  0.1601
  1 - 2           0.0250    -0.4351  0.4851
  2 - 3          -0.3250    -0.7851  0.1351
  2 - 1          -0.0250    -0.4851  0.4351
```

```
/* DDFM = SATTERTH: Use Satterthwaite approximation procedure to compute the
                    denominator degrees of freedom for testing fixed effects*/
/* METHOD=REML: by default */
proc mixed alpha=.05 cl covtest method=reml data=gaugerr;
  class operator part;
  model resp=operator / ddfm=satterth;
  random part operator*part;
  lsmeans operator / alpha=.05 cl diff adjust=tukey;
run; quit;
```

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

Type 3 Tests of Fixed Effects

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

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	operator	_operator	Adj P	Alpha	Lower	Upper	Adj Lower	Adj Upper
operator	1	2	0.9922	0.05	-0.3920	0.4420	-0.4751	0.5251
operator	1	3	0.3308	0.05	-0.7170	0.1170	-0.8001	0.2001
operator	2	3	0.2739	0.05	-0.7420	.09201	-0.8251	0.1751

- Both `PROC VARCOMP` and `PROC MIXED` compute estimates of variance components (different estimation procedures available)
  - ANOVA method: `METHOD = TYPE1`
  - RMLE method: `METHOD = REML` (default for `PROC MIXED`)
  - MIVQUE0: default for `VARCOMP`
- `PROC MIXED` can provide hypothesis tests and confidence intervals
  - Not all outputs from `PROC MIXED` are correct!

Rules For Expected Mean Squares (Restricted Model)

- EMS could be calculated using brute force method but may be difficult sometime
- For mixed models, good to have formal procedure
- Will take the two-factor mixed model (A fixed) as an example

1. Write each variable term in model as a row heading in a two-way table (the error term in the model as $\epsilon_{(ij..)m}$, with m for the replicate subscript)
2. Write the subscripts in the model as column headings. Over each subscript, write F for fixed factor and R for random one. Over this, write down the levels of each subscript

	F	R	R
	a	b	n
term	i	j	k
τ_i			
β_j			
$(\tau\beta)_{ij}$			
$\epsilon_{(ij)k}$			

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

	F	R	R
	a	b	n
term	i	j	k
τ_i		b	n
β_j	a		n
$(\tau\beta)_{ij}$			n
$\epsilon_{(ij)k}$			

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

	F	R	R
	a	b	n
term	i	j	k
τ_i		b	n
β_j	a		n
$(\tau\beta)_{ij}$			n
$\epsilon_{(ij)k}$	1	1	

5. 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
τ_i	0	b	n
β_j	a	1	n
$(\tau\beta)_{ij}$	0	1	n
$\epsilon_{(ij)k}$	1	1	1

6. The expected mean square for error is $E(MS_E) = \sigma^2$. For every other model term (row), the expected mean square contains σ^2 plus $\cdots$

– Cover the entries in the columns that contain non-bracketed subscript letters in this term;

– For rows including the same subscripts, multiply the remaining numbers to get coefficient for corresponding term in the model;

– A fixed effect is represented by the sum of squares of the model components associated with that factor divided by its degrees of freedom: $\tau_i \Longrightarrow Q(\tau) = \sum_{i=1}^{a} \tau_i^2/(a-1)$

Two-Factor Mixed Model (Restricted Model)

- Consider $E[MS_A] = \sigma^2 + \cdots$,
  - Ignore the first column (under i);
  - Model terms (rows) including subscript i: τ_i and $(\tau\beta)_{ij}$

    Term $\tau_i \Longrightarrow b \times n \sum_{i=1}^{a} \tau_i^2/(a-1)$

    Term $(\tau\beta)_{ij} \Longrightarrow 1 \times n\sigma^2_{\tau\beta}$
  - $E[MS_A] = \sigma^2 + bn \sum_{i=1}^{a} \tau_i^2/(a-1) + n\sigma^2_{\tau\beta}$

	F	R	R	
	a	b	n	
term	i	j	k	EMS
τ_i	0	b	n	$\sigma^2 + bn \sum_{i=1}^{a} \tau_i^2/(a-1) + n\sigma^2_{\tau\beta}$
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

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).

– Two-factor mixed model (A Fixed):

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

- Hasse Diagrams (Oehlert, 2000) can also be used to calculate the expected mean squares for balanced data (both restricted and unrestricted models).

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

Statistical Model with Two Random Factors

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

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

- Var(y_{ijk}) = $\sigma^2 + \sigma_\tau^2 + \sigma_\beta^2 + \sigma_{\tau\beta}^2$
- 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 (`METHOD=TYPE1`)

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

- May results in negative estimates, `PROC MIXED` uses `METHOD=REML` by default

The Measurement Systems Capability Experiment (Random-Effects Model)

- Assume the three operators were randomly selected $\Longrightarrow$ Random Factor

```
proc glm data=gaugerr;
  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;
run; quit;
```

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 for Random Model Analysis of Variance

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		

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

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

```
                              Sum of
Source              DF       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

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
                      Standard   Z
Cov Parm     Estimate    Error   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
```

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

```
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
                   Standard  Z
Cov Parm     Estimate  Error    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
- `PROC MIXED` with `METHOD=TYPE1`: Use standard normal $\rightarrow$ 95% CI uses 1.96

  $\hat{\sigma}_\tau^2 \pm 1.96(.0330) = (-0.05, 0.08), \hat{\sigma}_\beta^2 \pm 1.96(3.3738) = (3.67, 16.89)$
- `PROC MIXED` with `METHOD=REML` (by default): Satterthwaite's Approximation
- Satterthwaite's Approximation (Lec 4): Testing the Significance of σ_0^2
  - $\sigma_0^2 = E[(MS_r + \cdots + MS_s) - (MS_u + \cdots + MS_v)]/k$
  - Estimate $\hat{\sigma}_0^2 = [(MS_r + \cdots + MS_s) - (MS_u + \cdots + MS_v)]/k$
  - $f_i MS_i/\sigma_i^2 \overset{ind}{\sim} \chi^2_{f_i}$
  - Approximate $(1-\alpha) \times 100\%$ CI of σ_0^2

$$r\hat{\sigma}_0^2/\chi^2_{\alpha/2,r} \leq \sigma_0^2 \leq r\hat{\sigma}_0^2/\chi^2_{1-\alpha/2,r}$$

$$r = \frac{[(MS_r + \cdots + MS_s) - (MS_u + \cdots + MS_v)]^2}{\frac{MS_r^2}{f_r} + \cdots + \frac{MS_s^2}{f_s} + \frac{MS_u^2}{f_u} + \cdots + \frac{MS_v^2}{f_v}}$$

Example: The Measurement Systems Capability Experiment (Random-Effects Model)

- $\hat{\sigma}_\tau^2 = \frac{MS_A - MS_{AB}}{bn} = (1.31 - 0.71)/40 = 0.015$

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

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

- $\hat{\sigma}_\beta^2 = \frac{MS_B - MS_{AB}}{an} = (62.39 - 0.71)/6 = 10.28$

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

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

Three-Factor Mixed Model (A Fixed): Restricted Model

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

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

Three-Factor Mixed Model (A Fixed): Restricted Model

- Construct test statistics based on EMS
  - $H_0 : \tau_1 = \cdots = \tau_a = 0 \to ?$
  - $H_0 : \sigma^2_\beta = 0 \to \mathrm{MS_B}/\mathrm{MS_{BC}}$
  - $H_0 : \sigma^2_\gamma = 0 \to \mathrm{MS_C}/\mathrm{MS_{BC}}$
  - $H_0 : \sigma^2_{\tau\beta} = 0 \to \mathrm{MS_{AB}}/\mathrm{MS_{ABC}}$
  - $H_0 : \sigma^2_{\tau\gamma} = 0 \to \mathrm{MS_{AC}}/\mathrm{MS_{ABC}}$
  - $H_0 : \sigma^2_{\beta\gamma} = 0 \to \mathrm{MS_{BC}}/\mathrm{MS_E}$
  - $H_0 : \sigma^2_{\tau\beta\gamma} = 0 \to \mathrm{MS_{ABC}}/\mathrm{MS_E}$

Three-Factor Mixed Model (A Fixed): Unrestricted Model

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

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

Three-Factor Mixed Model (A Fixed): Unrestricted Model

- Construct test statistics based on EMS
  - $H_0 : \tau_1 = \cdots = \tau_a = 0 \rightarrow ?$
  - $H_0 : \sigma^2_\beta = 0 \rightarrow ?$
  - $H_0 : \sigma^2_\gamma = 0 \rightarrow ?$
  - $H_0 : \sigma^2_{\tau\beta} = 0 \rightarrow \mathrm{MS_{AB}}/\mathrm{MS_{ABC}}$
  - $H_0 : \sigma^2_{\tau\gamma} = 0 \rightarrow \mathrm{MS_{AC}}/\mathrm{MS_{ABC}}$
  - $H_0 : \sigma^2_{\beta\gamma} = 0 \rightarrow \mathrm{MS_{BC}}/\mathrm{MS_{ABC}}$
  - $H_0 : \sigma^2_{\tau\beta\gamma} = 0 \rightarrow \mathrm{MS_{ABC}}/\mathrm{MS_E}$

Satterthwaite's Approximate F-test

- For H_0 : effect $= 0$, no exact test exists.
- Suppose $E(MS') - E(MS'')$ is a multiple of the effect

  – Two linear combinations of mean squares MS' and MS''

  $MS' = \mathrm{MS}_r + \cdots + \mathrm{MS}_s$

  $MS'' = \mathrm{MS}_u + \cdots + \mathrm{MS}_v$

  MS' and MS'' do not share common mean squares

- Approximate test statistic $F = \frac{MS'}{MS''} = \frac{\mathrm{MS}_r + \cdots + \mathrm{MS}_s}{\mathrm{MS}_u + \cdots + \mathrm{MS}_v} \approx F_{p,q}$

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

  – f_i is the degrees of freedom associated with MS$_i$

- Need interpolation when p or q are not be integers. SAS can handle noninteger dfs.

Example: Restricted Three-Factor Mixed Model (A Fixed)

- Based on EMS, no exact test for $H_0: \tau_1 = \cdots = \tau_a = 0$ or equivalently $H_0: \sum \tau_i^2 = 0$
- Assume $a = 3, b = 2, c = 3, n = 2$

Source	DF	MS	EMS	F	P
A	2	0.7866	$12Q(A) + 6\sigma^2_{AB} + 4\sigma^2_{AC}$ $+2\sigma^2_{ABC} + \sigma^2$	?	?
B	1	0.0010	$18\sigma^2_B + 6\sigma^2_{BC} + \sigma^2$	0.33	.622
C	2	0.0560	$12\sigma^2_C + 6\sigma^2_{BC} + \sigma^2$	18.87	.051
AB	2	0.0056	$6\sigma^2_{AB} + 2\sigma^2_{ABC} + \sigma^2$	2.24	.222
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		

- $H_0 : \tau_1 = \tau_2 = \tau_3 = 0 \Longleftrightarrow H_0 : Q(A) = \sum_{i=1}^{a} \tau_i^2/(a-1) = 0$
  - $MS' = \mathrm{MS}_A + \mathrm{MS}_{ABC}$
  - $MS'' = \mathrm{MS}_{AB} + \mathrm{MS}_{AC}$
  - $E(MS' - MS'') = 12Q(A) = 12\frac{\Sigma\tau_i^2}{3-1}$

$$
\begin{aligned}
F &= \frac{\mathrm{MS}_A + \mathrm{MS}_{ABC}}{\mathrm{MS}_{AB} + \mathrm{MS}_{AC}} = \frac{.7866 + .0025}{.0107 + .0056} = 48.41 \\
p &= \frac{(.7866 + .0025)^2}{.7866^2/2 + .0025^2/4} = 2.01, \\
q &= \frac{(.0107 + .0056)^2}{.0107^2/4 + .0056^2/2} = 6.00
\end{aligned}
$$

- Interpolation needed

$$\mathrm{P}(F_{2,6} > 48.41) = .0002, \qquad \mathrm{P}(F_{3,6} > 48.41) = .0001$$

$$\mathrm{P}(F_{2.01,6} > 48.41) = .99(.0002) + .01(.0001) = .0002$$

- 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(48.41,2.01,6);
    f=finv(.95,2.01,6);
    c1=cinv(.025,18.57);  /* For Page 33 */
    c2=cinv(.975,18.57);  /* For Page 33 */
proc print data=one; run; quit;
```

Obs	p	f	c1	c2
1	.000197687	5.13799	8.61485	32.2833

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
- `PROC MIXED` in SAS allows one to specify G and R
- G through `RANDOM`, R through `REPEATED`
- Unrestricted linear mixed model is default

Sample Size Calculations

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

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

- For a set of fixed effects,
  - Use SAS to calculate $power$=1-β=1-PROBF$(F_{\alpha,\nu_1,\nu_2}, \nu_1, \nu_2, \delta)$ with

$$\delta = \frac{E[MS_N - MS_D] \times \nu_1}{E[MS_D]}$$

  - Use OCC in Chart V: β vs. $\Phi = \sqrt{\delta/(\nu_1 + 1)}$

- For a set of random effects:
  - $\nu_1 \times MS_N/\sigma_N^2 \sim \chi^2_{\nu_1}$ and $\nu_2 \times MS_D/\sigma_D^2 \sim \chi^2_{\nu_2}$
  - $F_0/\lambda^2 \sim F_{\nu_1,\nu_2}$ with $\lambda^2 = E[MS_N]/E[MS_D]$
  - Use OCC in Chart VI: β vs. $\lambda = \sqrt{E[MS_N]/E[MS_D]}$
  - Use SAS to calculate $power$=1-β=1-PROBF$(F_{\alpha,\nu_1,\nu_2}/\lambda^2, \nu_1, \nu_2)$.

Mixed Effects Model

Restricted Mixed Model

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

Unrestricted Mixed Model

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

Random-Effects Model

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