TYPE I vs. TYPE III Sum of Squares

Type I SS

- Referred to as Sequential sum of squares
- The order that sources of variation are listed in the model impacts the value calculated for the sum of square.
- If the model has two sources of variation, for example Rep and Treatment, the calculated sum of square for the source of variation is adjusted for the term(s) that precede it in the ANOVA table.
- The last source of variation in the ANOVA table is not adjusted because all sources of variation in the model are accounted for when calculating the sum of square.
- For example, if the order of the sources of variation is Rep then Treatment in the model, the:
    - SS Replicate = SS Replicate
    - SS Treatment = SS Treatment | Replicate
- The sum of square for the second source of variation is SS Treatment | Rep (read as SS of Treatment given Replicate is already in the model).
- The order that the source of variation appears in the model impacts the calculation of sum of squares for all but the last source of variation.

Type III SS

- Referred to as the Partial Sum of Squares.
- Appropriate for use when analyzing unbalanced data, including data with missing values.
- Sum of squares are adjusted for all other effects in the model.
- That is, the sum of squares is calculated assuming that all other sources of variation are in the model.
- The order of appearance of the source of variation in the model does not impact the sum of square calculated
- For the example with the order of the sources of variation of Treatment then Rep in the model, the:
    - SS Treatment = SS Treatment | Replicate
    - SS Replicate = SS Replicate | Treatment

SAS Example

```
Options pageno=1;
Data rcbdmiss;
Input trt $ rep yield;
Datalines;
A 1 3.1
A 2 3.1
A 3 3.0
B 1 3.3
B 2 .
B 3 3.2
C 1 3.6
C 2 3.4
C 3 3.6
D 1 3.9
D 2 4.0
D 3 4.2
;;
Ods rtf file="rcbd_miss.rtf";
ods graphics off;
Proc glm;
Class rep trt;
Model yield=rep trt;
Lsmeans trt/pdiff;
Title 'ANOVA RCBD with Missing Data - Order of SOV is Rep, then TRT';
Run;
Proc glm;
Class rep trt;
Model yield=trt rep;
Lsmeans trt/pdiff;
Title 'ANOVA RCBD with Missing Data - Order of SOV is TRT, then REP;
run;
Ods rtf close;
Run;
```

'ANOVA RCBD with Missing Data - Order of SOV in Model is REP, then TRT'

The GLM Procedure

Class Level Information		
Class	**Levels**	**Values**
rep	3	1 2 3
trt	4	A B C D

Number of Observations Read	12
Number of Observations Used	11

'ANOVA RCBD with Missing Data - Order of SOV in Model is REP, then TRT'

The GLM Procedure

Dependent Variable: yield

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	1.55422980	0.31084596	20.76	0.0023
Error	5	0.07486111	0.01497222		
Corrected Total	10	1.62909091			

Calculation of Error is not impacted by the use of Type I or Type III SS.

R-Square	Coeff Var	Root MSE	yield Mean
0.954047	3.505134	0.122361	3.490909

Source	DF	Type I SS	Mean Square	F Value	Pr > F
rep	2	0.00159091	0.00079545	0.05	0.9488
trt	3	1.55263889	0.51754630	34.57	0.0009

The Type I SS for rep differs from the Type III for rep.

Source	DF	Type III SS	Mean Square	F Value	Pr > F
rep	2	0.01013889	0.00506944	0.34	0.7279
trt	3	1.55263889	0.51754630	34.57	0.0009

The Type I SS and Type III SS for the last term will have the same value

USE THE ANOVA TABLE WITH THE TYPE III SS FOR CALCULATING *F-TESTS.*

'ANOVA RCBD with Missing Data - Order of SOV in Model is REP, then TRT'

The GLM Procedure
Least Squares Means

trt	yield LSMEAN	LSMEAN Number
A	3.06666667	1
B	3.22777778	2
C	3.53333333	3
D	4.03333333	4

Least Squares Means for effect trt
Pr > |t| for H0: LSMean(i)=LSMean(j)

Dependent Variable: yield

i/j	1	2	3	4
1		0.2214	0.0055	0.0002
2	0.2214		0.0455	0.0009
3	0.0055	0.0455		0.0041
4	0.0002	0.0009	0.0041	

Note: To ensure overall protection level, only probabilities associated with pre-planned comparisons should be used.

'ANOVA RCBD with Missing Data - Order of SOV in Model is TRT, then REP'

The GLM Procedure

Class Level Information		
Class	**Levels**	**Values**
rep	3	1 2 3
trt	4	A B C D

Number of Observations Read	12
Number of Observations Used	11

'ANOVA RCBD with Missing Data - Order of SOV in Model is TRT, then REP'

The GLM Procedure

Dependent Variable: yield

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	1.55422980	0.31084596	20.76	0.0023
Error	5	0.07486111	0.01497222		
Corrected Total	10	1.62909091			

R-Square	Coeff Var	Root MSE	yield Mean
0.954047	3.505134	0.122361	3.490909

Source	DF	Type I SS	Mean Square	F Value	Pr > F
trt	3	1.54409091	0.51469697	34.38	0.0009
rep	2	0.01013889	0.00506944	0.34	0.7279

Source	DF	Type III SS	Mean Square	F Value	Pr > F
trt	3	1.55263889	0.51754630	34.57	0.0009
rep	2	0.01013889	0.00506944	0.34	0.7279

Dependent Variable: yield

trt	yield LSMEAN	LSMEAN Number
A	3.06666667	1
B	3.22777778	2
C	3.53333333	3
D	4.03333333	4

Least Squares Means for effect trt
Pr > |t| for H0: LSMean(i)=LSMean(j)

Dependent Variable: yield

i/j	1	2	3	4
1		0.2214	0.0055	0.0002
2	0.2214		0.0455	0.0009
3	0.0055	0.0455		0.0041
4	0.0002	0.0009	0.0041	

Note: To ensure overall protection level, only probabilities associated with pre-planned comparisons should be used.