TYPE I, TYPE II, AND TYPE III SUMS OF SQUARES

DAVID C. HOWELL
1/12/2010

The data file:

These are the data from Howell (2010) Table 16.5 **except that two additional observations were added to cell(2,4).** The example in the book had the unfortunate feature that both levels of B had the same number of observations. That is no longer true, and makes it a bit easier to see the differences in the analyses.

I have added two extra columns to my SPSS data set. Agrp is just the coding of the levels of A as 1 and 2, while Bgrp is the coding of the levels of B as 1, 2, 3, 4. The other columns (except for dv) represent contrast coding.

5	1	1	0	0	1	0	0	1.00
7	1	1	0	0	1	0	0	1.00
9	1	1	0	0	1	0	0	1.00
8	1	1	0	0	1	0	0	1.00
2	1	0	1	0	0	1	0	1.00
5	1	0	1	0	0	1	0	1.00
7	1	0	1	0	0	1	0	1.00
3	1	0	1	0	0	1	0	1.00
9	1	0	1	0	0	1	0	1.00
8	1	0	0	1	0	0	1	1.00
11	1	0	0	1	0	0	1	1.00
12	1	0	0	1	0	0	1	1.00
14	1	0	0	1	0	0	1	1.00
11	1	-1	-1	-1	-1	-1	-1	1.00
15	1	-1	-1	-1	-1	-1	-1	1.00
16	1	-1	-1	-1	-1	-1	-1	1.00
10	1	-1	-1	-1	-1	-1	-1	1.00
9	1	-1	-1	-1	-1	-1	-1	1.00
7	-1	1	0	0	-1	0	0	2.00
9	-1	1	0	0	-1	0	0	2.00
10	-1	1	0	0	-1	0	0	2.00
9	-1	1	0	0	-1	0	0	2.00
3	-1	0	1	0	0	-1	0	2.00
8	-1	0	1	0	0	-1	0	2.00
9	-1	0	1	0	0	-1	0	2.00
11	-1	0	1	0	0	-1	0	2.00
9	-1	0	0	1	0	0	-1	2.00
12	-1	0	0	1	0	0	-1	2.00

14	-1	0	0	1	0	0	-1	2.00
8	-1	0	0	1	0	0	-1	2.00
7	-1	0	0	1	0	0	-1	2.00
11	-1	-1	-1	-1	1	1	1	2.00
14	-1	-1	-1	-1	1	1	1	2.00
10	-1	-1	-1	-1	1	1	1	2.00
12	-1	-1	-1	-1	1	1	1	2.00
13	-1	-1	-1	-1	1	1	1	2.00
11	-1	-1	-1	-1	1	1	1	2.00
12	-1	-1	-1	-1	1	1	1	2.00

The following are the different analyses as they are computed by SPSS version 15.

Type I A then B then AB

Tests of Between-Subjects Effects

Dependent Variable: DV

Source	Type I Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	216.017(a)	7	30.860	5.046	.001
Intercept	3410.526	1	3410.526	557.709	.000
Agrp	9.579	1	9.579	1.566	.220
Bgrp	186.225	3	62.075	10.151	.000
Agrp * Bgrp	20.212	3	6.737	1.102	.364
Error	183.457	30	6.115		
Total	3810.000	38			
Corrected Total	399.474	37			

a R Squared = .541 (Adjusted R Squared = .434)

Type I B then A then AB

Tests of Between-Subjects Effects

Dependent Variable: DV

Source	Type I Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	216.017(a)	7	30.860	5.046	.001
Intercept	3410.526	1	3410.526	557.709	.000
Bgrp	193.251	3	64.417	10.534	.000
Agrp	2.553	1	2.553	.418	.523
Bgrp * Agrp	20.212	3	6.737	1.102	.364
Error	183.457	30	6.115		
Total	3810.000	38			

Corrected Total	399.474	37			

a R Squared = .541 (Adjusted R Squared = .434)

Type II

Tests of Between-Subjects Effects

Dependent Variable: DV

Source	Type II Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	216.017(a)	7	30.860	5.046	.001
Intercept	3410.526	1	3410.526	557.709	.000
Agrp	2.553	1	2.553	.418	.523
Bgrp	186.225	3	62.075	10.151	.000
Agrp * Bgrp	20.212	3	6.737	1.102	.364
Error	183.457	30	6.115		
Total	3810.000	38			
Corrected Total	399.474	37			

a R Squared = .541 (Adjusted R Squared = .434)

Type III

Tests of Between-Subjects Effects

Dependent Variable: DV

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	216.017(a)	7	30.860	5.046	.001
Intercept	3163.841	1	3163.841	517.370	.000
Agrp	3.464	1	3.464	.566	.458
Bgrp	185.172	3	61.724	10.093	.000
Agrp * Bgrp	20.212	3	6.737	1.102	.364
Error	183.457	30	6.115		
Total	3810.000	38			
Corrected Total	399.474	37			

a R Squared = .541 (Adjusted R Squared = .434)

Now I will run a whole set of regressions. Bear with me.

Using contrast coding for variable A alone: (i.e. $Y = b_1A$)

ANOVA(b)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	9.579	1	9.579	.884	.353(a)
	Residual	389.894	36	10.830		
	Total	399.474	37			

a Predictors: (Constant), A
b Dependent Variable: DV

Using contrast coding for variable B alone: (i.e. $Y = b_1B_1 + b_2B_2 + b_3B_3$)

ANOVA(b)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	193.251	3	64.417	10.621	.000(a)
	Residual	206.222	34	6.065		
	Total	399.474	37			

a Predictors: (Constant), B3, B2, B1
b Dependent Variable: DV

Using contrast coding for the interaction terms alone (i.e. $Y = b_1AB_1 + b_2AB_2 + b_3AB_3$)

ANOVA(b)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	20.538	3	6.846	.614	.610(a)
	Residual	378.936	34	11.145		
	Total	399.474	37			

a Predictors: (Constant), AB3, AB2, AB1
b Dependent Variable: DV

Using contrast terms for A and B but not AB (i.e. $Y = b_1A + b_2B_1 + b_3B_2 + b_4B_3$)

ANOVA(b)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	195.805	4	48.951	7.931	.000(a)
	Residual	203.669	33	6.172		
	Total	399.474	37			

a Predictors: (Constant), B3, A, B2, B1

Using contrasts for A and AB but not B ((i.e. $Y = b_1A + b_2AB1 + b_3AB_2 + b_4AB_3$)

ANOVA(b)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	30.845	4	7.711	.690	.604(a)
	Residual	368.629	33	11.171		
	Total	399.474	37			

a Predictors: (Constant), AB3, A, AB2, AB1
b Dependent Variable: DV

Using contrasts for B and AB but not A ((i.e. $Y = b_1B_1 + b_2AB_2 + b_3AB_3\ \ b_4AB_1 + b_5AB_2 + b_6AB_3$)

ANOVA(b)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	212.552	6	35.425	5.875	.000(a)
	Residual	186.921	31	6.030		
	Total	399.474	37			

a Predictors: (Constant), AB3, B2, AB2, B3, B1, AB1
b Dependent Variable: DV

Now contrast codes for A, B, and AB (**This is the full model**)

ANOVA(b)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	216.017	7	30.860	5.046	.001(a)
	Residual	183.457	30	6.115		
	Total	399.474	37			

a Predictors: (Constant), AB3, B2, A, AB2, B3, B1, AB1
b Dependent Variable: DV

Type I sums of squares

For Type I each term is adjusted only for the terms that were entered before it. I will just use the orders A, B, AB and B, A, AB, but you could, if you wanted, us AB, B, A or A, AB, B, etc. I can't think why.

Let's start with Type I SS where the order of entry is A, B, AB

SSreg(A) = 9.579
 This will be SS(A) in the Anova

SSreg(A,B) = 195.805
 Gain is SSreg = 195.805 – 9.579 = 186.225
 This will be SS(B) in the Anova

SSreg(A, B, AB) = 216.017
 Gain from previous model is 216.017 – 195.805 = 20.212
 This is SS(AB) in the Anova

Now Type I SS where the order of entry is B, A, AB

SSreg(B) = 193.251
 This is SS(B) in that Anova

SSreg(A,B) – 195.805
 Gain is SS(Reg) is 195.805 – 193.251 = 2.554
 This is SS(B)

SSreg(A, B, AB) = 216.017
 Gain is 216.017 – 195.805 = 20.212
 This is SS(AB)

For Type II you adjust B for A, A for B, and then AB for A and B. (If we had a three way, we would adjust A for B and C, B for A and C, C for A and B, AB for A, B, C, BC for A, B, C, AC for A, B, C, and ABC for A, B, C, AB, AC, BC. I have never tried that.)

SSreg(A,B) = 195.805

SSreg(A) = 9.579
 Difference is 195.805 – 9.579 = 186.226

SSreg(B) = 193.251

Difference from A,B is 195.805 – 193.251 = 2.554

These are SS(A) and SS(B) in Anova.

For AB we take the full model
SSreg(A,B,AB) = 216.017
SSreg(A,B) = 195.805
Difference is 216.017 – 195.805 = 20.212
This is SS(AB)

Type III
Here everything is adjusted for everything else.

SSreg(A,B,AB) = 216.017
SSreg(A,B) = 195.805
Difference is 216.017 – 195.805 = 20.212
This is SS(AB)

SSreg(A,AB) = 30.845
Diff from full model is 216.017 – 30.845 = 185.172

SSreg(B,AB) = 212.552
Diff from full model is 216.017 – 212.552 = 3.465

MS(Error)
In all of these analyses the error term would come from MSresidual of the full model, which is 6.115

References

Howell, D. C. (2010). Statistical Methods for Psychology. Wadsworth:Cengage.