Research Design - - Topic 15
Reporting Model II in MRC with Completely Randomized Designs and a Continuous Factor

General Description

Step I. Computing statistics for the “main effects”

Step II. Computing statistics for the “two way interactions”

Step III. Computing statistics for the “three way interaction”

Statistics that should be reported

Running Model II in SPSS GLM Univariate?

1

General Description

To date, we have discussed the use of MRC for various designs. Generally, we have considered both Model I and Model II, and have discussed the direct relevance of these to the analysis of variance model. In this lecture, we will discuss presenting the results from Model II, focusing on the statistics reported and their appropriate interpretation.

This presentation includes example statistics for a three factor completely randomized design, A (with 3 levels), B (with 2 levels) and C (a centred continuous variable), where Dummy coding is used for the categorical factors. The description, however, can be generalized to any number of factors with any number of levels.

2

Sample

Data

	B1		B2	
	C	X	C	X
A1	5	31	-4	36
	-4	33	-4	37
	6	37	6	39
	3	29	3	33
	-1	37	-6	34
	-3	31		
A2	-5	32	2	36
	2	36	7	27
	4	30	-6	28
	-3	33	2	30
	-3	37		
A3	4	38	1	36
	-5	29	-1	37
	3	36	2	36
	5	32	-3	46
			-5	45
			-2	42

3

Step I. Computing statistics for the “main effects”

Compute the relevant squared multiple correlations

- Enter a1, a2, b1, c $R^2 = .31795$

  Use the regression coefficients from this run to compute intercepts and within cells slope

- Enter b1, c $R^2 = .10521$
- Enter a1, a2, c $R^2 = .26991$
- Enter a1, a2, b1 $R^2 = .30289$

4

Compute the squared semi-partial multiple correlations, and associated F-ratios

$$\hat{R}_A^2 = R_{A,B,C}^2 - R_{B,C}^2 = .21274$$

$$\hat{R}_B^2 = R_{A,B,C}^2 - R_{A,C}^2 = .04804$$

$$\hat{R}_C^2 = R_{A,B,C}^2 - R_{A,B}^2 = .01506$$

5

General Form of the F-ratios

$$F = \frac{\hat{R}^2 / v_1}{R_{error}^2 / v_2}$$

Using Model 1 error terms

$$F_A = \frac{.21274/2}{(1.-.31795)/25} = 3.899 \quad ns$$

$$F_B = \frac{.04804/1}{.68205/25} = 1.761 \qquad F_C = \frac{.01506/1}{.68205/18} = 0.552$$

Using Model 2 error terms

$$F_A = \frac{.21274/2}{1-.67404/18} = 5.873 \quad p < .05$$

$$F_B = \frac{.04804/1}{.32596/18} = 2.653 \qquad F_C = \frac{.01506/1}{.32596/18} = 0.832$$

6

Solution of the regression equation

Coefficients[a]

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	38.49660	1.43556		26.81638	0.00000
	a1	-3.10582	1.80666	-0.32993	-1.71909	0.09796
	a2	-5.25968	1.90036	-0.53132	-2.76773	0.01047
	b1	-2.02647	1.52715	-0.22336	-1.32696	0.19651
	c	-0.13984	0.18822	-0.12382	-0.74294	0.46444

a. Dependent Variable: x

The following equations yield the cell estimates on the next slide.

$$\text{int} = 38.4966 - 3.10582 * a1 - 5.25968 * a2 - 2.02647 * b1$$

$$\text{slo} = -.13984$$

7

Intercepts for the A and B factors

	B1	B2	Means
A1	33.364	35.391	34.378
A2	31.210	33.237	32.224
A3	36.470	38.497	37.483
Means	33.682	35.708	

Note: If you were to plot these intercepts, the lines would be parallel. That is, there is no AB interaction at this stage.

Main effect slope = - 0.140

8

Interpretation

The significant F-ratio for A indicates that the means of the intercepts for A differ more than expected by chance.

As described in earlier lectures, the regression coefficient for a1 indicates that:

$$I_{A1B2} - I_{A3B2} = 35.391 - 38.497 = -3.106$$

But since there is no interaction, this is equivalent to:

$$I_{A1} - I_{A3} = 34.378 - 37.483 = -3.105$$

Similarly: $I_{A2} - I_{A3} = 32.224 - 37.483 = -5.259$

If the F-ratio for B is significant, a similar logic applies.

If the F-ratio for C was significant, the within cells slope differs from 0.

9

Step II. Computing the two-way interaction statistics.

Compute the relevant squared multiple correlations

- Enter a1, a2, b1, c, a1b1, a2b1, a1c, a2c, b1c

  $R^2 = .51808$

  Use the regression coefficients from this run to compute intercepts and two-way slopes

- Enter a1, a2, b1, c, a1c, a2c, b1c $R^2 = .37051$
- Enter a1, a2, b1, c, a1b1, a2b1, b1c $R^2 = .50727$
- Enter a1, a2, b1, c, a1b1, a2b1, a1c, a2c $R^2 = .49766$

10

Compute the squared semi-partial multiple correlations, and associated F-ratios

$$\hat{R}^2_{AB} = R^2_{A,B,C,AB,AC,BC} - R^2_{A,B,C,AC,BC} = .14757$$

$$\hat{R}^2_{AC} = R^2_{A,B,C,AB,AC,BC} - R^2_{A,B,C,AB,BC} = .01081$$

$$\hat{R}^2_{BC} = R^2_{A,B,C,AB,AC,BC} - R^2_{A,B,C,AB,AC} = .02042$$

11

F-ratios for Step II:

Using Model 1 error terms

$$F_{AB} = \frac{.14757/2}{(1-.51806)/20} = 3.062 \quad ns$$

$$F_{AC} = \frac{.01081/2}{.48194/20} = 0.449 \qquad F_{BC} = \frac{.02042/1}{.48194/20} = 0.847$$

Using Model 2 error terms

$$F_{AB} = \frac{.14757/2}{(1-.67404)/18} = 4.074 \quad p < .05$$

$$F_{AC} = \frac{.01081/2}{.32596/18} = 0.298 \qquad F_{BC} = \frac{.02042/1}{.32596/18} = 1.128$$

12

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	39.83013	1.69157		23.54627	0.00000
	a1	-4.08616	2.41441	-0.43407	-1.69240	0.10609
	a2	-9.33586	2.62664	-0.94309	-3.55430	0.00199
	b1	-6.02426	2.77130	-0.66399	-2.17380	0.04190
	c	-0.37740	0.46357	-0.33416	-0.81411	0.42517
	a1b1	2.99084	3.66649	0.26372	0.81572	0.42427
	a2b1	9.28004	3.85859	0.76238	2.40503	0.02597
	a1c	0.32138	0.48529	0.18697	0.66223	0.51538
	a2c	0.18198	0.51630	0.09148	0.35246	0.72818
	b1c	0.34547	0.37529	0.21172	0.92054	0.36826

a. Dependent Variable: x

$$\text{int} = 39.83013 - 4.08616 * a1 - 9.33586 * a2 - 6.02426 * b1 + 2.99084 * a1 * b1 + 9.28004 * a2 * b1$$

$$slo = -.37740 + .32138 * a1 + .18198 * a2 + .34547 * b1$$

13

Intercepts

	B1	B2
A1	32.711	35.744
A2	33.750	30.494
A3	33.806	39.830

Note: Plotting these intercepts would show the influence of the AB interaction.

Slopes

	B1	B2
A1	.289	-.056
A2	.150	-.195
A3	-.032	-.377

Note: Plotting these slopes would produce parallel lines because the three-way interaction is excluded.

14

Interpretation of Intercepts

The regression coefficients for the ab terms on Slide 13 are equal to the following contrast-contrast interactions based on the intercepts in Slide 14:

$$b_{a1b1} = 32.711 - 35.744 - 33.806 + 39.830 = 2.991$$

$$b_{a2b1} = 33.750 - 30.494 - 33.806 + 39.830 = 9.280$$

15

Interpretation of Slopes

The regression coefficients for the ac and bc terms on Slide 13 are equal to contrasts of slopes in Step 14:

$$b_{a1c} = -.056 - (-.377) = .321$$

$$b_{a2c} = -.195 - (-.377) = .182$$

$$b_{b1c} = -.032 - (-.377) = .345$$

Note, however, that because the three way interaction is defined as 0 in this step, these values are the same as contrasts involving the corresponding marginal means of the slopes.

16

Step III. Computing the three-way interaction statistics.

Compute the relevant squared multiple correlations

- Enter a1, a2, b1, c, a1b1, a2b1, a1c, a2c, b1c, a1b1c, a2b1c

$R^2 = .67404$

Use the regression coefficients from this run to compute intercepts and three-way slopes

17

Compute the squared semi-partial multiple correlation, and associated F-ratio

$$\hat{R}^2_{ABC} = R^2_{A,B,C,AB,AC,BC,ABC} - R^2_{A,B,C,AB,AC,BC} = .15592$$

For a completely randomized design, compute the residual R^2

$$R^2_{error} = 1 - R^2_{A,B,C,AB,AC,BC,ABC} = .32596$$

For studies involving repeated measures, other error terms will be required, as discussed in earlier topics.

18

General Form of the F-ratio

$$F = \frac{\hat{R}^2 / v_1}{R^2_{error} / v_2}$$

$$F_{ABC} = \frac{.15592/2}{.32596/18} = 4.305 \quad p < .05$$

19

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	38.20000	1.56840		24.35601	0.00000
	a1	-2.25185	2.19081	-0.23921	-1.02786	0.31763
	a2	-7.97522	2.33594	-0.80564	-3.41414	0.00309
	b1	-5.53068	2.40851	-0.60959	-2.29631	0.03388
	c	-1.60000	0.57917	-1.41668	-2.76257	0.01282
	a1b1	2.53808	3.18253	0.22380	0.79751	0.43555
	a2b1	8.73348	3.35062	0.71748	2.60653	0.01785
	a1c	1.74815	0.66255	1.01706	2.63853	0.01669
	a2c	1.62017	0.68142	0.81443	2.37765	0.02871
	b1c	2.21753	0.71668	1.35899	3.09418	0.00626
	a1b1c	-2.32123	0.86104	-0.90960	-2.69584	0.01478
	a2b1c	-2.41012	0.91395	-0.76480	-2.63704	0.01674

a. Dependent Variable: x

Note. 9 of the regression coefficients are significant but only the last two are used in Model II. See slide 23 for the values that should be reported for Model II.

20

$Int = 38.20 - 2.25185*a1 - 7.97522*a2 - 5.53068*b1 + 2.53808*a1*b1 + 8.73348*a2*b1$

Intercepts

	B1	B2
A1	32.956	35.948
A2	33.428	30.225
A3	32.669	38.200

$slo = -1.60 + 1.74815*a1 + 1.62017*a2 + 2.21753*b1 - 2.32123*a1*b1 - 2.41012*a2*b1$

Slopes

	B1	B2
A1	.044	.148
A2	-.172	.020
A3	.618	-1.600

Note: Plotting these values shows the complete results.

21

The meaning of the regression coefficients for abc in Slide 20 based on the statistics in Slide 21 are:

$$b_{a1b1c} = .044 - .148 - .618 + (-1.600) = -2.322$$

$$b_{a2b1c} = -.172 - .020 - .618 + (-1.600) = -2.410$$

Note. Each of these describe a contrast-contrast interaction of slopes. The first is (A1-A3 with B1-B2)C and the second is (A2-A3 with B1-B2)C. They do not involve any pairwise contrasts of slopes.

22

Statistics that should be reported drawn from the 3 steps

Step		b	S.E.	t	β	F-ratio
1	a1	-3.106	1.807	-1.719	-.330	
	a2	-5.260	1.900	-2.768**	-.531	5.873*
	b1	-2.026	1.527	-1.327	-.223	2.653
	c	-.140	0.188	-0.743	-.124	0.832
2	a1b1	2.991	3.666	.816	.264	
	a2b1	9.280	3.859	2.405*	.762	4.074*
	a1c	.321	0.485	.662	.187	
	a2c	.182	.516	.352	.091	0.298
	b1c	.345	.375	.921	.212	1.128
3	a1b1c	-2.321	.861	-2.696**	-.910	
	a2b1c	-2.410	.914	-2.637**	-.765	4.305*

23

Interpretation

For main effects, only A was significant, F(2,18) = 5.873, p< .05

- *We could thus perform post hoc tests comparing the three intercepts for A on Slide 8, 34.378, 32.224, and 37.483. The regression coefficients on Slide 7 for a1 and a2 test two of the three contrasts of interest. Only the regression coefficient for a2 was significant, b = -5.260, t(18) = -2.768, p< .02, indicating that the intercept for A2 was significantly less than that for A3. To compare A1 with A2, the analysis must be redone with A2 coded with all 0's; or some other procedure should be used.*

24

For the two way interactions, only that for AB was significant, $F(2,18) = 4.074$, $p < .05$.

- *Because AB was significant, this suggests that contrasts involving any of the cell intercepts in Slide 14 could be made, as in tests of simple main effects of means. This application of multiple regression does not provide any such contrasts.*
- *The regression coefficients for a1b1 and a2b1 describe contrast-contrast interactions. Of these, only the regression coefficient for a2b1 was significant, $b=9.280$, $t(18) = 2.405$, $p < .05$ which refers to the contrast-contrast interaction (A2-A3)*(B1-B2). If it was desired to test the (A1-A2)*(B1-B2) contrast-contrast interaction, another multiple regression analysis would be required.*

25

The three way interaction was significant, $F(2,18) = 4.305$, $p < .05$.

- *Because this F-ratio was significant, it suggests that contrasts could be made of any of the cell slopes in Slide 21. This analysis does not provide any such contrasts. One could perform specific contrasts using formulae described in earlier lectures.*
- *The regression coefficients for a1b1c and a2b1c describe contrast-contrast interactions for these slopes. Both were significant. The three way interaction of slopes (A1-A3)*(B1-B2)*C, $b = -2.32123$, $t(18)=-2.696$, $p<.02$ and (A2-A3)*(B1-B2)*C, $b = -2.41012$, $t(18)= -2.637$ were significant. As can be seen in Slide 21, this interaction occurs because the differences between the slopes for the two levels of A are different at the B1 level than at the B2 level.*

26

The meaning of the regression coefficients in Slide 20 given the intercepts and slopes in Slide 21.

		Meaning
Intercepts	Constant	INT(A3B2)
	a1	A1B2-A3B2
	a2	A2B2-A3B2
	b1	A3B1-A3B2
	a1b1	A1B1-A1B2-A3B1+A3B2
	a2b1	A2B1-A2B2-A3B1+A3B2
Slopes	c	A3B2
	a1c	A1B2-A3B2
	a2c	A2B1-A3B2
	b1c	A3B1-A3B2
	a1b1c	A1B1-A1B2-A3B1+A3B2
	a2b3c	A2B1-A2B2-A3B1+A3B2

27

Running Model II in SPSS GLM Univariate?

In previous examples, we showed how the analysis using Model I could also be run in SPSS GLM Univariate, making appropriate changes to the Design statement where required. This was true whether one of the factors was categorical or continuous. If all factors are categorical, they could also be run in SPSS GLM Univariate using Model II, however, Model II cannot be run in SPSS GLM Univariate if any factor is continuous. Again, this is because SPSS GLM solves specific equations and does not compute a series of squared multiple correlations and resulting squared semi-partial multiple correlations.

28