Chapter 13 - Factorial Analysis of Variance

Note: Because of severe rounding in reporting and using means, there will be visible rounding error in the following answers, when compared to standard computer solutions. I have made the final answer equal the correct answer, even if that meant that it is not exactly the answer to the calculations shown. (e.g. 3(3.3) would be shown as 10.0, not 9.9)

13.1 Mother/infant interaction for primiparous/multiparous mothers under or over 18 years of age with LBW or full-term infants:

Table of cell means

		Size/Age			
		LBW < 18	LBW > 18	NBW	
Mother's Parity	Primi-	4.5	5.3	6.4	5.40
	Multi-	3.9	6.9	8.2	6.33
		4.2	6.1	7.3	5.87

$$SS_{total} = \Sigma X^2 - \frac{(\Sigma X)^2}{N} = 2404 - \frac{352^2}{60} = 338.93$$

$$SS_{Parity} = ns\Sigma\left(\bar{X}_{i.} - \bar{X}_{..}\right)^2$$
$$= 10(3)[(5.40-5.87)^2 + (6.33-5.87)^2]$$
$$= 30(0.4356) = 13.067$$

$$SS_{size} = np\Sigma\left(\bar{X}_{.j} - \bar{X}_{..}\right)^2$$
$$= 10(2)[(4.200-5.87)^2 + (6.10-5.87)^2 + (7.30-5.87)^2]$$
$$= 20(2.79+0.05+2.04) = 20(4.89)$$
$$= 97.733$$

$$SS_{cells} = n\Sigma\left(\bar{X}_{ij} - \bar{X}_{..}\right)^2$$
$$= 10[(4.5-5.87)^2 + \ldots + (8.2-5.87)^2]$$
$$= 10(12.853) = 128.53$$

$$SS_{PS} = SS_{cells} - SS_P - SS_S = 128.53 - 13.067 - 97.733$$
$$= 17.733$$

$$SS_{error} = SS_{total} - SS_{cells} = 338.93 - 128.53$$
$$= 210.40$$

Source	*df*	SS	MS	F
Parity	1	13.067	13.067	3.354
Size/Age	2	97.733	48.867	12.541*
P x S	2	17.733	8.867	2.276
Error	54	210.400	3.896	
Total	59	338.933		

$^*p < .05 \quad F_{.05}(2,54) = 3.17$

13.3 The mean for these primiparous mothers would not be expected to be a good estimate of the mean for the population of all primiparous mothers because 50% of the population of primiparous mothers do not give birth to LBW infants. This would be important if we wished to take means from this sample as somehow representing the population means for primiparous and multiparous mothers.

13.5 Memory of avoidance of a fear-producing stimulus:

		Area of Stimulation			
		Neutral	Area A	Area B	Mean
Delay	50	28.6	16.8	24.4	23.27
	100	28.0	23.0	16.0	22.33
	150	28.0	26.8	26.4	27.07
	Mean	28.2	22.2	22.27	24.22

$\Sigma X = 1090 \quad \Sigma X^2 = 28374 \quad N = 45 \; n = 5 \quad a = 3 \quad d = 3$

$$SS_{\text{total}} = \Sigma X^2 - \frac{(\Sigma X)^2}{N} = 28374 - \frac{1090^2}{45} = 1971.778$$

$$\begin{aligned} SS_{Delay} &= na\Sigma\left(\bar{X}_{i.} - \bar{X}_{..}\right)^2 \\ &= 5(3)[(23.27-24.22)^2 + (22.33-24.22)^2 + (27.07-24.22)^2] \\ &= 5(3)(0.90+3.57+8.12) = 30(12.60) \\ &= 188.578 \end{aligned}$$

$$\begin{aligned} SS_{Area} &= nd\Sigma\left(\bar{X}_{.j} - \bar{X}_{..}\right)^2 \\ &= 5(3)[(28.20-24.22)^2 + (22.20-24.22)^2 + (22.27-24.22)^2] \\ &= 356.044 \end{aligned}$$

$$\begin{aligned} SS_{Cells} &= n\Sigma\left(\bar{X}_{ij} - \bar{X}_{..}\right)^2 \\ &= 5[(28.60-24.22)^2 + (16.80-24.22)^2 + \ldots + (26.4-24.22)^2] \\ &= 916.578 \end{aligned}$$

$$SS_{DA} = SS_{cells} - SS_D - SS_A = 916.578 - 188.578 - 356.044 = 371.956$$

$$SS_{error} = SS_{total} - SS_{cells} = 1971.778 - 916.578 = 1055.200$$

Source	*df*	SS	MS	**F**
Delay	2	188.578	94.289	3.22
Area	2	356.044	178.022	6.07*
D x A	4	371.956	92.989	3.17*
Error	36	1055.200	29.311	
Total	44	1971.778		

$* p < .05$ $[F_{.05(2,36)} = 3.27;\ F_{.05(4,36)} = 2.64]$

13.7 In Exercise 13.5, if A refers to Area:

$\hat{\alpha}_1$ = the treatment effect for the Neutral site

$= \bar{X}_{.1} - \bar{X}_{..}$

= 28.2 – 24.22 = 3.978

13.9 The Bonferroni test to compare Site means.

$$t = \frac{\overline{N} - \overline{A}}{\sqrt{\frac{MS_{error}}{n_N} + \frac{MS_{error}}{n_A}}} = \frac{28.20 - 22.20}{\sqrt{\frac{29.311}{15} + \frac{29.311}{15}}} = 3.03 \quad (\text{Reject } H_0)$$

$$t = \frac{\overline{N} - \overline{B}}{\sqrt{\frac{MS_{error}}{n_N} + \frac{MS_{error}}{n_B}}} = \frac{28.20 - 22.27}{\sqrt{\frac{29.311}{15} + \frac{29.311}{15}}} = 3.00 \quad (\text{Reject } H_0)$$

$[t'_{.025}(2,36) = \pm 2.34]$

We can conclude that both the difference between Groups *N* and *A* and between Groups *N* and *B* are significant, and our familywise error rate will not exceed α = .05.

13.11 Rerunning Exercise 11.3 as a factorial design:

The following printout is from SPSS

Tests of Between-Subjects Effects

Dependent Variable: Recall

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	1059.800[a]	3	353.267	53.301	.000
Intercept	5017.600	1	5017.600	757.056	.000
Age	115.600	1	115.600	17.442	.000
LevelProc	792.100	1	792.100	119.512	.000
Age * LevelProc	152.100	1	152.100	22.949	.000
Error	238.600	36	6.628		
Total	6316.000	40			
Corrected Total	1298.400	39			

a. R Squared = .816 (Adjusted R Squared = .801)

[The Corrected Model is the sum of the main effects and interaction. The Intercept is the correction factor, which is $(\Sigma X)^2$. The Total (as opposed to Corrected Total) is ΣX^2. The Corrected Total is what we have called Total.]

Estimated Marginal Means

3. Age * LevelProc

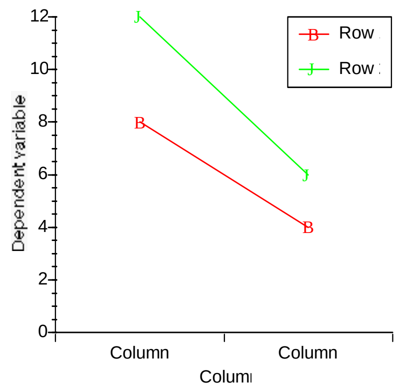
Dependent Variable: Recall

Age	LevelProc	Mean	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
1.00	1.00	6.500	.814	4.849	8.151
	2.00	19.300	.814	17.649	20.951
2.00	1.00	7.000	.814	5.349	8.651
	2.00	12.000	.814	10.349	13.651

The results show that there is a significance difference between younger and older subjects, that there is better recall in tasks which require more processing, and that there is an interaction between age and level of processing (LevelProc). The difference between the two levels of processing is greater for the younger subjects than it is for the older ones, primarily because the older ones do not do much better with greater amounts of processing.

13.13 Made-up data with main effects but no interaction:

Cell means:		
	8	12
	4	6



13.15 The interaction was of primary interest in an experiment by Nisbett in which he showed that obese people varied the amount of food they consumed depending on whether a lot or a little food was visible, while normal weight subjects ate approximately the same amount under the two conditions.

13.17 Unequal sample sizes Klemchuk, Bond, & Howell (1990):

Cell *ns*:

		Age		
		Younger	Older	
Daycare	No	14	12	26
	Yes	10	4	14
		24	16	40 = N

$$\bar{n}_h = \frac{k}{\Sigma \frac{1}{n_i}} = \frac{4}{\frac{1}{14} + \frac{1}{12} + \frac{1}{10} + \frac{1}{4}} = 7.925$$

Cell Means:

		Age		
		Younger	Older	
Daycare	No	-1.2089	0.0750	-0.5669
	Yes	-0.5631	0.5835	0.0102
		-0.8860	0.3292	13.895

$$SS_{Daycare} = an_h\Sigma\left(\bar{X}_{i.} - \bar{X}_{..}\right)^2 = 2*7.925\left[\left(-0.5669-(-.2784)\right)^2 + \left(0.0102-(-.2784)\right)^2\right]$$
$$= 2*7.925*.0.1665 = 2.639$$

$$SS_{Age} = dn_h\Sigma\left(\bar{X}_{.j} - \bar{X}_{..}\right)^2 = 2*7.925\left[\left(-0.8860-(-.2784)\right)^2 + \left(0.3292-(-.2784)\right)^2\right]$$
$$= 2*7.925*0.7384 = 11.704$$

$$SS_{cells} = n_h\Sigma\left(\bar{X}_{ij} - \bar{X}_{..}\right)^2 = 7.925\left[\begin{array}{l}\left(-1.2089-(-.2784)\right)^2 + \left(0.0750-(-.2784)\right)^2 + \\ \left(-0.5631-(-.2784)\right)^2 + \left(0.5835-(-.2784)\right)^2\end{array}\right]$$
$$= 7.925*1.8146*14.3811$$

$$SS_{AD} = SS_{cells} - SS_D - SS_A = 14.381 - 2.639 - 11.704 = 0.038$$

Cell variances:

		Age	
		Younger	Older
Daycare	No	0.7407	0.1898
	Yes	0.9551	0.2456

$$SS_{error} = \Sigma\left(n_{ij}-1\right)s_{ij}^2 = 13(0.7407) + 11(0.1898) + 9(0.9551) + 3(0.2456) = 21.050$$

Source	*df*	SS	MS	*F*
Age	1	11.704	11.704	20.02*
Daycare	1	2.639	2.639	4.51*
A×D	1	0.038	0.038	<1
Error	36	21.050	0.585	
Total	39			

$*p < .05$ $[F_{.05,(1,36)} = 4.11]$

We can conclude that there are effects due to both Age and Daycare, but there is no interaction between the two main effect variables. (One variance is nearly 4 times another, but with such small sample sizes we can't tell if there is heterogeneity of variance.)?

13.19 Magnitude of effect for mother-infant interaction data in Exercise 13.1:

$$\eta^2_P = \frac{SS_{parity}}{SS_{total}} = \frac{13.067}{338.933} = .04$$

$$\eta^2_S = \frac{SS_{size}}{SS_{total}} = \frac{97.733}{338.933} = .29$$

$$\eta^2_{PS} = \frac{SS_{PS}}{SS_{total}} = \frac{17.733}{338.933} = .05$$

$$\omega^2_P = \frac{SS_{parity} - (p-1)MS_{error}}{SS_{total} + MS_{error}} = \frac{13.067 - (1)(3.896)}{338.933 + 3.896} = .03$$

$$\omega^2_S = \frac{SS_{size} - (s-1)MS_{error}}{SS_{total} + MS_{error}} = \frac{97.733 - (2)(3.896)}{338.933 + 3.896} = .26$$

$$\omega^2_{PS} = \frac{SS_{PS} - (p-1)(s-1)MS_{error}}{SS_{total} + MS_{error}} = \frac{17.733 - (1)(2)(3.896)}{338.933 + 3.896} = .03$$

13.21 Magnitude of effect for avoidance learning data in Exercise 13.5:

$$\eta^2_D = \frac{SS_{delay}}{SS_{total}} = \frac{188.578}{1971.778} = .10$$

$$\eta^2_A = \frac{SS_{Area}}{SS_{total}} = \frac{356.044}{1971.778} = .18$$

$$\eta^2_{DA} = \frac{SS_{DA}}{SS_{total}} = \frac{371.956}{1971.778} = .19$$

$$\omega^2_D = \frac{SS_{delay} - (d-1)MS_{error}}{SS_{total} + MS_{error}} = \frac{188.578 - (2)(29.311)}{1971.778 + 29.311} = .06$$

$$\omega^2_A = \frac{SS_{area} - (a-1)MS_{error}}{SS_{total} + MS_{error}} = \frac{356.044 - (2)(29.311)}{1971.778 + 29.311} = .15$$

$$\omega^2_{DA} = \frac{SS_{DA} - (d-1)(a-1)MS_{error}}{SS_{total} + MS_{error}} = \frac{371.956 - (2)(2)(29.311)}{1971.778 + 29.311} = .13$$

13.23 Three-way ANOVA on Early Experience x Intensity of UCS x Conditioned Stimulus (Tone or Vibration):
n = 5 in all cells SS$_{total}$ = 41,151.00

E×I×C Cells

	CS = Tone				CS = Vibration				
Exper:	Hi	Med	Low		Hi	Med	Low		
Control	11	16	21	12.0	19	24	29	24.00	20.00
Tone	25	28	34	29.0	21	26	31	26.00	27.50
Vib	6	13	20	13.0	40	41	52	44.33	28.67
Both	22	30	30	27.33	35	38	48	40.33	33.83
	16	21.75	105	21.33	28.75	32.25	40.00	33.66	27.50

E×I Cells

	Intensity			
Experience:	High	Med	Low	
Control	15	20	25	20.00
Tone	23	27	32.5	27.50
Vib	23	27	36	28.67
Both	28.5	34	39	33.83
	22.38	27.00	33.12	27.50

E×C Cells

	Conditioned Stimulus		
Experience:	Tone	Vib	
Control	16.00	24.00	20.00
Tone	29.00	26.00	27.50
Vib	13.00	44.33	28.67
Both	27.33	40.33	33.83
	21.33	33.66	27.50

I×C Cells

	Conditioned Stim		
Intensity:	Tone	Vib	
High	16.00	28.75	22.38
Med	21.75	32.25	27.00
Low	26.25	40.00	33.12
	21.33	33.67	27.50

$$SS_E = nic\Sigma\left(\bar{X}_{i..} - \bar{X}_{...}\right)^2 = 5(3)(2)\begin{bmatrix}(20-27.5)^2 + (27.5-27.5)^2 + \\ (28.67-27.5)^2 + (33.83-27.5)^2\end{bmatrix}$$
$$= 2931.667$$

$$SS_I = nec\Sigma\left(\bar{X}_{.j.} - \bar{X}_{...}\right)^2 = 5(4)(2)\left[(22.38-27.5)^2 + (27.00-27.5)^2 + (33.12-27.5)^2\right]$$
$$= 2326.250$$

$$SS_{cellsEI} = nc\Sigma\left(\bar{X}_{ij.} - \bar{X}_{...}\right)^2 = (5)(2)\left[(15.00-27.50)^2 + \ldots + (39.00-27.50)^2\right]$$
$$= 5325.000$$

$$SS_{E\times I} = SS_{cellsEI} - SS_E - SS_I = 5325.000 - 2931.667 - 2326.250 = 67.083$$

$$SS_C = nei\Sigma\left(\bar{X}_{..k} - \bar{X}_{...}\right)^2 = 5(4)(3)\left[(21.33-27.5)^2 + (33.66-27.5)^2\right]$$
$$= 4563.333$$

$$SS_{cellsEC} = ni\Sigma\left(\bar{X}_{i.k} - \bar{X}_{...}\right)^2 = (5)(3)\left[(16.00-27.50)^2 + \ldots + (40.33-27.50)^2\right]$$
$$= 12,110.000$$

$$SS_{E\times C} = SS_{cellsEC} - SS_E - SS_C = 12,110.000 - 2931.667 - 4563.333 = 4615.000$$

$$SS_{cellsIC} = ne\Sigma\left(\bar{X}_{ij.} - \bar{X}_{...}\right)^2 = (5)(4)\left[(15.00-27.50)^2 + \ldots + (39.00-27.50)^2\right]$$
$$= 6945.000$$

$$SS_{I\times C} = SS_{cellsIC} - SS_I - SS_C = 6945.000 - 2326.250 - 4563.333 = 55.417$$

$$SS_{cellsEIC} = n\Sigma\left(\bar{X}_{ijk} - \bar{X}_{...}\right)^2 = (5)\left[(11.00-27.50)^2 + \ldots + (48.00-27.50)^2\right]$$
$$= 14,680.000$$

$$SS_{E\times I\times C} = SS_{cellsEIC} - SS_E - SS_I - SS_C - SS_{EI} - SS_{EC} - SS_{IC}$$
$$= 14,680.000 - 2931.667 - 2326.250 - 4563.333 - 67.083 - 4615.000 - 55.417$$
$$= 121.25$$

$$SS_{error} = SS_{total} - SS_{Cells C \times E \times I} = 41{,}151.000 - 14{,}680.000 = 26.471.000$$

Source	*df*	SS	MS	*F*
Experience	3	2931.667	977.222	3.544*
Intensity	2	2326.250	1163.125	4.218*
Cond Stim	1	4563.333	4563.333	16.550*
E x I	6	67.083	11.181	<1
E x C	3	4615.000	1538.333	5.579*
I x C	2	55.417	27.708	<1
E x I x C	6	121.250	20.208	<1
Error	96	26,471.000	275.740	
Total	119	41,151.000		

$* p < .05 \quad [F_{.05(1,96)} = 3.94;\ F_{.05(2,96)} = 3.09;\ F_{.05(3,96)} = 2.70;\ F_{.05(6,96)} = 2.19]$

There are significant main effects for all variables with a significant Experience × Conditioned Stimulus interaction.

13.25 Analysis of Epineq.dat:

Tests of Between-Subjects Effects

Dependent Variable: Trials to reversal

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	141.130[a]	8	17.641	8.158	.000
Intercept	1153.787	1	1153.787	533.554	.000
DOSE	133.130	2	66.565	30.782	.000
DELAY	2.296	2	1.148	.531	.590
DOSE * DELAY	5.704	4	1.426	.659	.622
Error	214.083	99	2.162		
Total	1509.000	108			
Corrected Total	355.213	107			

a. R Squared = .397 (Adjusted R Squared = .349)

13.27 Tukey on Dosage data from Exercise 13.25

Multiple Comparisons

Dependent Variable: Trials to reversal
Tukey HSD

(I) dosage of epinephrine	(J) dosage of epinephrine	Mean Difference (I-J)	Std. Error	Sig.
0.0 mg/kg	0.3 mg/kg	-1.67*	.35	.000
	1.0 mg/kg	1.03*	.35	.010
0.3 mg/kg	0.0 mg/kg	1.67*	.35	.000
	1.0 mg/kg	2.69*	.35	.000
1.0 mg/kg	0.0 mg/kg	-1.03*	.35	.010
	0.3 mg/kg	-2.69*	.35	.000

Based on observed means.

*. The mean difference is significant at the .05 level.

.

All of these groups differed from each other at $p \leq .05$.

13.29 Simple effects on data in Exercise 13.28.

Source	*df*	SS	MS	*F*
Condition	1	918.750	918.75	34.42*
Cond @ Inexp.	1	1014.00	1014.00	37.99*
Cond @ Exp.	1	121.50	121.50	4.55*
Cond*Exper	1	216.750	216.75	8.12*
Other Effects	9	2631.417		
Error	36	961.000	26.694	
Total	47	4727.917		

*P < .05 [$F_{.05(1.36)} = 4.12$]

13.31 Dress codes and Performance

$$SS_{total} = \Sigma\left(X - \bar{X}_{..}\right)^2$$
$$= (91-72.050)^2 + (78-72.050)^2 + ... + (56-72.050)^2$$
$$= 13554.65$$

$$SS_{Code} = nc\Sigma\left(\bar{X}_{i.} - \bar{X}..\right)^2$$
$$= 10*7[(73.929-72.050)^2 + (70.171-72.050)^2]$$
$$= 494.290$$

$$SS_{School(Yes)} = n\Sigma\left(\bar{X}_{.j} - \bar{X}..\right)^2$$
$$= 10[(79.7-73.929)^2 + (71.5-73.929)^2 + ... + (73.5-73.929)^2]$$
$$= 10(147.414) = 1474.14$$

$$SS_{School(No)} = n\Sigma\left(\bar{X}_{.j} - \bar{X}..\right)^2$$
$$= 10[(68.5-70.171)^2 + (73.7-70.171)^2 + ... + (71.1-70.171)^2]$$
$$= 10(126.314) = 1263.14$$

$$SS_{School(Code)} = SS_{School(Yes)} + SS_{School(No)} = 1474.14 + 1263.14 = 2737.28$$
$$SS_{error} = SS_{total} - SS_C - SS_{S(C)} = 13554.65 - 494.29 - 2737.28 = 10323.08$$

Source	***df***	**SS**	**MS**	***F***
Code	1	494.290	494.290	2.166
Error$_1$	12	2737.280	228.107	
School(Code)	12	2737.280	288.107	2.784*
Error$_2$	126	10323.08	81.931	
Total	139	13554.65		

* $p < .05$

The *F* for Code is not significant but the *F* for the nested effect is. But notice that the two *F* values are not all that far apart but their *p* values are very different. The reason for this is that we only have 12 *df* for error to test Code, but 126 *df* for error to test School(Code).

13.33 Analysis of Seligman et al. (1990)

If we think that males are generally more optimistic than females, then the sample sizes themselves are part of the "treatment" effect. We probably would not want to ignore that if we are looking at sex as an independent variable. In fact, the lack of independence between sample size and the effect under study is an important problem when it occurs.

13.35 This question does