

Chapter 11 - Simple Analysis of Variance

11.1 Retrieval of rat pups:

Descriptives

TIME				
	N	Mean	Std. Deviation	Std. Error
5 days	6	17.1667	5.1153	2.0883
20 days	6	22.1667	5.1153	2.0883
35 days	6	42.1667	5.1153	2.0883
Total	18	27.1667	12.1086	2.8540

$$\begin{aligned}
 SS_{total} &= \sum (X_{ij} - \bar{X}_{..})^2 = (15 - 27.1667)^2 + \\
 &\quad (10 - 27.1667)^2 + \dots + (40 - 27.1667)^2 \\
 &= 2492.5
 \end{aligned}$$

$$\begin{aligned}
 SS_{treat} &= n \sum (\bar{X}_j - \bar{X}_{..})^2 = 6 \left((17.1667 - 27.1667)^2 + (22.1667 - 27.1667)^2 + \right. \\
 &\quad \left. (42.1667 - 27.1667)^2 \right) \\
 &= 6(350.00) = 2100
 \end{aligned}$$

$$SS_{error} = SS_{total} - SS_{treat} = 2492.50 - 2100.00 = 392.50$$

ANOVA

TIME					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	2100.000	2	1050.000	40.127	.000
Within Groups	392.500	15	26.167		
Total	2492.500	17			

11.3 Recall in Eysenck (1974) for four Age/Levels of Processing groups:

Descriptives

RECALL				
	N	Mean	Std. Deviation	Std. Error
1.00	10	6.5000	1.4337	.4534
2.00	10	19.3000	2.6687	.8439
3.00	10	7.0000	1.8257	.5774
4.00	10	12.0000	3.7417	1.1832
Total	40	11.2000	5.7699	.9123

a.

$$SS_{total} = \sum (X_{ij} - \bar{X}_{..})^2 = (8 - 11.2)^2 + (6 - 11.2)^2 + \dots + (11 - 11.2)^2 = 1298.4$$

$$SS_{treat} = n \sum (\bar{X}_j - \bar{X}_{..})^2 = 10 \left((6.5 - 11.2)^2 + (19.3 - 11.2)^2 + (7.0 - 11.2)^2 + (12.0 - 11.2)^2 \right) = 10(105.98) = 1059.8$$

$$SS_{error} = SS_{total} - SS_{treat} = 1298.4 - 1059.8 = 238.6$$

ANOVA

RECALL					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1059.800	3	353.267	53.301	.000
Within Groups	238.600	36	6.628		
Total	1298.400	39			

b. Groups 1 and 3 combined versus 2 and 4 combined:

Descriptives

RECALL				
	N	Mean	Std. Deviation	Std. Error
Low	20	6.7500	1.6182	.3618
High	20	15.6500	4.9019	1.0961
Total	40	11.2000	5.7699	.9123

$$SS_{total} = \sum (X_{ij} - \bar{X}_{..})^2 = (8 - 11.20)^2 + (6 - 11.20)^2 + \dots + (11 - 11.20)^2$$

$$= 1298.40$$

$$SS_{treat} = n \sum (\bar{X}_j - \bar{X}_{..})^2 = 20 \left((6.75 - 11.20)^2 + (15.65 - 11.20)^2 \right)$$

$$= 20(39.605) = 792.1$$

$$SS_{error} = SS_{total} - SS_{treat} = 1298.4 - 792.1 = 502.3$$

ANOVA

RECALL					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	792.100	1	792.100	59.451	.000
Within Groups	506.300	38	13.324		
Total	1298.400	39			

- c. The results are somewhat difficult to interpret because the error term now includes variance between younger and older participants. Notice that this is roughly double what it was in part a. In addition, we do not know whether the level of processing effect is true for both age groups, or if it applies primarily to one group.

11.5 Rerun of Exercise 11.2 with additional subjects:

The following is abbreviated printout from SPSS

a.

Descriptives^a

RECALL				
	N	Mean	Std. Deviation	Std. Error
Younger	12	18.4167	3.2039	.9249
Older	10	12.0000	3.7417	1.1832
Total	22	15.5000	4.6980	1.0016

a. PROCESS = High

ANOVA^a

RECALL					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	224.583	1	224.583	18.800	.000
Within Groups	238.917	20	11.946		
Total	463.500	21			

a. PROCESS = High

b. and c. With and without pooling variances:

Independent Samples Test^a

t-test for Equality of Means						
		t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
RECALL	Equal variances assumed	4.336	20	.000	6.4167	1.4799
	Equal variances not assumed	4.273	17.893	.000	6.4167	1.5018

a. PROCESS = High

d. The squared t for the pooled case = $4.3359^2 = 18.80$, which is the F in the analysis of variance.

11.7 Magnitude of effect measures for Exercise 11.3a:

$$\eta^2 = \frac{SS_{\text{group}}}{SS_{\text{total}}} = \frac{1059.8}{1298.4} = .82$$

$$\omega^2 = \frac{SS_{\text{group}} - (k-1)MS_{\text{error}}}{SS_{\text{total}} + MS_{\text{error}}} = \frac{1059.8 - (4-1)6.63}{1298.4 + 6.63} = .80$$

11.9 Magnitude of effect for Foa et al. (1991) study:

$$\eta^2 = \frac{SS_{\text{group}}}{SS_{\text{total}}} = \frac{507.840}{2786.907} = .18$$

$$\omega^2 = \frac{SS_{\text{group}} - (k-1)MS_{\text{error}}}{SS_{\text{total}} + MS_{\text{error}}} = \frac{507.840 - (4-1)55.587}{2786.907 + 55.587} = .12$$

11.11 The results are basically the same as ours, although we are presented with confidence limits on group means and r^2 (which is really η^2).

11.13 Model for Exercise 11.1:

$$X_{ij} = \mu + \tau_j + e_{ij}$$

where

μ = grand mean

τ_j = the effect of the j th treatment

e_{ij} = the unit of error for the i th subject in treatment j

11.15 Model for Exercise 11.3:

$$X_{ij} = \mu + \tau_j + e_{ij}$$

where

μ = grand mean

τ_j = the effect of the j th treatment (where a “treatment” is a particular combination of Age and Task.

e_{ij} = the unit of error for the i th subject in treatment j

11.17 Howell & Huessy (1981) study of ADD in elementary school vs. GPA in high school:

Group	Group Means	s_j^2	n_j
Never ADD	2.6774	0.9450	201
2nd only	1.6123	1.0195	13
4th only	1.9975	0.5840	12
2nd & 4th	2.0287	0.2982	8
5th only	1.7000	0.7723	14
2nd & 5th	1.9000	1.0646	9
4th & 5th	1.8986	0.0927	7
all 3 yrs	1.4225	0.3462	8
Overall	2.4444		272

$$\begin{aligned} SS_{group} &= \sum n_j (\bar{X}_j - \bar{X}_{..})^2 \\ &= 201(2.6774 - 2.4444)^2 + 13(1.6123 - 2.4444)^2 + \dots + 8(1.4225 - 2.4444)^2 \\ &= 44.5570 \end{aligned}$$

$$\begin{aligned} MS_{error} &= \text{average variance (weighted)} \\ &= \frac{200 * 0.9450 + 12 * 1.0195 + \dots + 7 * 0.3462}{200 + 12 + \dots + 7} = 0.8761 \end{aligned}$$

ANOVA

GPA

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	44.557	7	6.365	7.266	.000
Within Groups	231.282	264	.876		
Total	275.839	271			

11.19 Square Root Transformation of data in Table 11.5:

Original data:	Control	0.1	0.5	1	2
	130	93	510	229	144
	94	444	416	475	111
	225	403	154	348	217
	105	192	636	276	200
	92	67	396	167	84
	190	170	451	151	99
	32	77	376	107	44
	64	353	192	235	84
	69	365	384		284
	93	422			293
Means	109.4	258.6	390.56	248.5	156
S.D.	58.5	153.32	147.68	118.74	87.65
Var	3421.82	23506.04	21809.78	14098.86	7682.22
<i>n</i>	10	10	9	8	10

Square root transformed data:

	Control	0.1	0.5	1	2
	11.402	9.644	22.583	15.133	12.000
	9.695	21.071	20.396	21.794	10.536
	15.000	20.075	12.410	18.655	14.731
	10.247	13.856	25.219	16.613	14.142
	9.592	8.185	19.900	12.923	9.165
	13.784	13.038	21.237	12.288	9.950
	5.657	8.775	19.391	10.344	6.633
	8.000	18.788	13.856	15.330	9.165
	8.307	19.105	19.596		16.852
	9.644	20.543			17.117
Means	10.13	15.31	19.40	15.39	12.03
S.D.	2.73	5.19	4.00	3.67	3.54
Var	7.48	26.96	16.03	13.49	12.55
<i>n</i>	10	10	9	8	10

11.21 Magnitude of effect for data in Exercise 11.17:

$$\eta^2 = \frac{SS_{\text{group}}}{SS_{\text{total}}} = \frac{44.557}{275.839} = .16$$

$$\omega^2 = \frac{SS_{\text{group}} - (K - 1)MS_{\text{error}}}{SS_{\text{total}} + MS_{\text{error}}} = \frac{44.557 - (8 - 1)0.876}{275.839 + 0.876} = .1389$$

11.23 Transforming Time to Speed in Exercise 11.22 involves a reciprocal transformation. The effect of the transformation is to decrease the relative distance between large values.

11.25 The parts of speech (noun vs. verb) are fixed. But the individual items within those parts of speech may well be random, representing a random sample of nouns and a random sample of verbs.

11.27 Analysis of Davey et al. data

Report

dv

group	Mean	N	Std. Deviation
1.00	12.6000	10	6.02218
2.00	7.0000	10	2.98142

3.00	8.7000	10	2.35938
Total	9.4333	30	4.62887

ANOVA

dv

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	164.867	2	82.433	4.876	.016
Within Groups	456.500	27	16.907		
Total	621.367	29			

11.29 Analysis of Epinuneq.dat, ignoring the effect of Interval. These results come from SPSS.

Descriptives

ERRORS

	N	Mean	Std. Deviation	Std. Error
1	42	3.14	1.52	.24
2	42	4.81	1.25	.19
3	37	2.11	1.51	.25
Total	121	3.40	1.80	.16

ANOVA

ERRORS

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	147.970	2	73.985	36.197	.000
Within Groups	241.187	118	2.044		
Total	389.157	120			

11.31 Computer exercise. Repeating Exercise 11.29 using Epineq.dat. This output comes from Minitab.

a. Analysis for Interval 1:

Analysis of Variance for Errors

Source	DF	SS	MS	F	P
Dosage	2	71.72	35.86	14.93	0.000
Error	33	79.25	2.40		
Total	35	150.97			

Individual 95% CIs

For Mean

Level	N	Mean	StDev	Based on Pooled StDev	
1	12	3.167	1.801	(-----*-----)	
2	12	5.333	1.073	(-----*-----)	
3	12	1.917	1.676	(-----*-----)	
Pooled StDev =				1.5	3.0
4.5				6.0	

b. Analysis for Interval 2:

Analysis of Variance for Errors

Source	DF	SS	MS	F	P
Dosage	2	32.06	16.03	8.78	0.001
Error	33	60.25	1.83		
Total	35	92.31			

Individual 95% CIs For

Mean

Based on Pooled StDev

Level	N	Mean	StDev	Based on Pooled StDev	
1	12	2.833	1.267	(-----*-----)	
2	12	4.417	1.379	(-----*-----)	
3	12	2.167	1.403	(-----*-----)	
Pooled StDev =				2.4	3.6
4.8					

c. Analysis for Interval 3:

Analysis of Variance for Errors

Source	DF	SS	MS	F	P
Dosage	2	35.06	17.53	7.76	0.002
Error	33	74.58	2.26		
Total	35	109.64			

For Mean				Individual 95% CIs	
Level	N	Mean	StDev	Based on Pooled StDev	
1	12	3.167	1.403	-----*-----	
2	12	4.417	1.311	-----*-----	
3	12	2.000	1.758	-----*-----	
Pooled StDev =			1.503	1.2	2.4
3.6			4.8		

b. The average of the 9 variances :

$$\bar{s}^2 = \frac{1.801^2 + 1.073^2 + \dots + 1.758^2}{9} = 2.162$$

The average of the three error terms:

$$\text{average}(MS_{error}) = \frac{2.40 + 1.83 + 2.26}{3} = 2.163$$

These two values agree within minor rounding error.

11.33 Gouzoulis-Mayfrank et al. (2000) study

Descriptives

Performance								
	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Ecstasy	28	25.9286	3.88662	.73450	24.4215	27.4356	18.00	34.00
Control	28	29.6071	4.89398	.92488	27.7095	31.5048	18.00	40.00
Cannibis	28	29.3929	3.39214	.64105	28.0775	30.7082	22.00	37.00
Total	84	28.3095	4.39599	.47964	27.3555	29.2635	18.00	40.00

Test of Homogeneity of Variances

Performance			
Levene Statistic	df1	df2	Sig.
1.223	2	81	.300

ANOVA

Performance					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	238.738	2	119.369	7.082	.001
Within Groups	1365.214	81	16.854		
Total	1603.952	83			

Robust Tests of Equality of Means

Performance				
	Statistic ^a	df1	df2	Sig.
Welch	7.565	2	52.932	.001

a. Asymptotically F distributed.

(b) The pairwise differences are 3.678, 3.464, and 0.214, and the square root of MS_{error} is 4.105. This gives d values of 0.896, 0.844, and 0.05. (c) It is reasonable to tentatively conclude that Ecstasy produces lower scores than either the Control condition or the Cannabis condition, which don't differ.

11.35 There should be no effect on the magnitude of the effect size measure because η^2 is not dependent on the underlying metric of the independent variable.

