Chapter 16—One-way Analysis of Variance

I am assuming that most people would prefer to see the solutions to these problems as computer printout. (I will use SPSS for consistency.)

16.1 Analysis of Eysenck's data:

a) The analysis of variance:

----- O N E W A Y -----

Variable RECALL
By Variable GROUP Group Membership

Analysis of Variance

Source	D.F.	Sum of Squares	Mean Squares	F Ratio	F Prob.
Between Groups	1	266.4500	266.4500	25.2294	.0001
Within Groups	18	190.1000	10.5611		
Total	19	456.5500			

Group	Count	Mean	Standard Deviation	Standard Error	95 Pct Conf Int for Mean
Grp 1	10	19.3000	2.6687	.8439	17.3909 TO 21.2091
Grp 2	10	12.0000	3.7417	1.1832	9.3234 TO 14.6766
Total	20	15.6500	4.9019	1.0961	13.3558 TO 17.9442

b) *t* test

```
t-tests for Independent Samples of GROUP    Group Membership
                         Number
 Variable               of Cases       Mean          SD   SE of Mean
 -----------------------------------------------------------------------
 RECALL
 Young                      10        19.3000       2.669      .844
 Older                      10        12.0000       3.742     1.183
 -----------------------------------------------------------------------
          Mean Difference = 7.3000
         Levene's Test for Equality of Variances: F= .383   P= .544

      t-test for Equality of Means                         95%
 Variances   t-value    df    2-Tail Sig     SE of Diff    CI for Diff
 -----------------------------------------------------------------------
 Equal       5.02       18        .000         1.453    (4.247, 10.353)
 Unequal     5.02       16.27     .000         1.453    (4.223, 10.377)
```

Notice that if you square the t value of 5.02 you obtain 25.20, which is the same as the F in the analysis of variance. Notice also that the analysis of variance procedure produces confidence limits on the means, whereas the t procedure produces confidence limits on the difference of means.

16.3 Expanding on Exercise 16.2:

a) Combine the Low groups together and the High groups together:

```
     Variable  RECALL
  By Variable  LOWHIGH
                              Analysis of Variance
                          Sum of        Mean          F      F
       Source      D.F.   Squares       Squares       Ratio  Prob.
Between Groups       1     792.1000     792.1000     59.4505  .0000
Within Groups       38     506.3000      13.3237
Total               39    1298.4000

                        Standard   Standard
Group   Count    Mean   Deviation     Error    95 Pct Conf Int for
Mean

Grp 1    20     6.7500    1.6182      .3618     5.9927  TO    7.5073
Grp 2    20    15.6500    4.9019    1.0961    13.3558  TO   17.9442

Total    40    11.2000    5.7699      .9123     9.3547  TO   13.0453
```

Here we have compared recall under conditions of Low versus High processing, and can conclude that higher levels of processing lead to significantly better recall.

b) The answer is still a bit difficult to interpret because both groups contain both younger and older subjects, and it is possible that the effect holds for one age group but not for the other.

16.5 η^2 and ω^2 for the data in Exercise 16.1:

$SS_{group} = 266.45$
$SS_{total} = 456.55$
$MS_{error} = 10.564$
$k = 2$

$$\eta^2 = \frac{SS_{group}}{SS_{total}} = \frac{266.45}{456.55} = .58$$

$$\omega^2 = \frac{SS_{group} - (k-1)MS_{error}}{SS_{total} + MS_{error}}$$

$$= \frac{266.45 - (2-1)10.564}{456.55 + 10.564} = \frac{255.886}{467.114} = .55$$

16.7 Foa *et al*. (1991) study:

Group	***n***	**Mean**	**S.D.**	**Total**	**Variance**
SIT	14	11.07	3.95	155	15.6025
PE	10	15.40	11.12	154	123.6544
SC	11	18.09	7.13	199	50.8369
WL	10	19.50	7.11	195	50.5521
Total	45	15.622		703	

$$\overline{X}.. = \frac{703}{45} = 15.622$$

$$SS_{treat} = \sum n_j\left(\overline{X}_j - \overline{X}..\right)^2$$

$$= 14(11.07 - 15.622)^2 + 10(15.40 - 15.622)^2 + 11(18.09 - 15.622)^2 + 10(19.50 - 15.622)^2$$

$$= 507.840$$

$$MS_{error} = \frac{\sum (n_i - 1)s_i^2}{\sum (n_i - 1)}$$

$$= \frac{13(15.6025) + 9(123.6544) + 10(50.8369) + 9(50.5521)}{41}$$

$$= 55.587$$

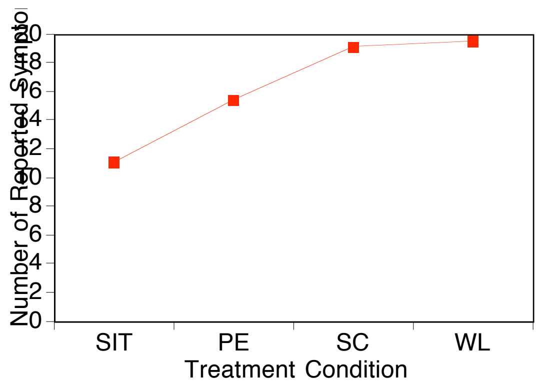
$$SS_{error} = [\Sigma(n_1 - 1)]MS_{error} = 41 * 55.587 = 2279.067$$

From these values we can fill in the complete summary table and compute the *F* value.

Source	***df***	***SS***	***MS***	***F***
Treatment	3	507.840	169.280	3.04
Error	41	2279.067	55.587	
Total	44	2786.907		

[$F_{.05}(3,41) = 2.84$] We can reject the null hypothesis and conclude that there are significant differences between groups. Some treatments are more effective than others.

b)



c) It would appear that the more interventionist treatments lead to fewer symptoms than the less interventionist ones, although we would have to run multiple comparisons to tell exactly which groups are different from which other groups.

16.9 If the sample sizes in Exercise 16.7 were twice as large, that would double the SS_{treat} and MS_{treat}. However it would have no effect on MS_{error}, which is simply the average of the group variances. The result would be that the *F* value would be doubled.

16.11 Effect size for tests in Exercise 16.10.
It only makes sense to calculate an effect size for significant comparisons in this study, so we will deal with SIT vs SC.

$$\hat{d} = \frac{\bar{X}_{SC} - \bar{X}_{SIT}}{\sqrt{MS_{error}}} = \frac{18.09 - 11.07}{\sqrt{55.579}} = \frac{7.02}{7.455} = 0.94$$

The SIT group is nearly a full standard deviation lower in symptoms when compared to the SC group, which is a control group.

16.13 ANOVA on GPAs for the ADDSC data:

Variable GPA
By Variable Group

Source	D.F.	Sum of Squares	Mean Squares	F Ratio	F Prob.
Between Groups	2	22.5004	11.2502	22.7362	.0000
Within Groups	85	42.0591	.4948		
Total	87	64.5595			

Group	Count	Mean	Standard Deviation	Standard Error	95 Pct Conf Int for Mean
Grp 1	14	3.2536	.5209	.1392	2.9528 TO 3.5543
Grp 2	49	2.5920	.6936	.0991	2.3928 TO 2.7913
Grp 3	25	1.7436	.8020	.1604	1.4125 TO 2.0747
Total	88	2.4563	.8614	.0918	2.2737 TO 2.6388

There is a significant difference between the groups, telling us that there is a relationship between ADDSC score in elementary school and the GPA the student has in 9th grade. From the means it is clear that the GPA declines as the ADDSC score increases.

These are real data, and they tell us that a teacher in elementary school can already pick out those students who will do well and badly in high school. I have always found these results depressing and worrisome, even though psychologists are supposed to like to be able to predict. There are some things I wish weren't so predictable.

16.15 Analysis of Darley and Latané data:

Group	***n***	**Mean**	**Total**
1	13	0.87	11.31
2	26	0.72	18.72
3	13	0.51	6.63
Total	52		36.66

$$SS_{treat} = \Sigma n_j \left(\overline{X}_j - \overline{X}_{..}\right)^2$$
$$= 13\left(0.87 - 0.705\right)^2 + 26\left(0.72 - 0.705\right)^2 + 13\left(0.51 - 0.705\right)^2$$
$$= 0.8541$$
$$MS_{error} = 0.053 \quad \text{(given in text)}$$
$$SS_{error} = [\Sigma(n_1 - 1)]MS_{error} = 49 * 0.053 = 2.597$$

From these values we can fill in the complete summary table and compute the *F* value.

Source	***df***	***SS***	***MS***	***F***
Treatment	2	0.854	0.427	8.06
Error	49	2.597	0.053	
Total	51	3.451		

$[F_{.05}(2,49) = 3.18]$ We can reject the null hypothesis and conclude that subjects are less likely to summon help quickly if there are other bystanders around.

16.17 Bonferroni test on data in Exercise 16.2:

Both of these comparisons will be made using *t* tests. The means are given in Exercise 16.15 above.

$$t = \frac{\overline{X}_i - \overline{X}_j}{\sqrt{\frac{MS_{error}}{n_i} + \frac{MS_{error}}{n_j}}}$$

For Young/Low versus Old/Low:

$$t = \frac{6.5 - 7.0}{\sqrt{\frac{6.6278}{10} + \frac{6.6278}{10}}} = \frac{-0.5}{1.151} = -0.434$$

For Young/High versus Old/High:

$$t = \frac{19.3 - 12.0}{\sqrt{\frac{6.6278}{10} + \frac{6.6278}{10}}} = \frac{7.3}{1.151} = 6.34$$

For 36 *df* for error and for 2 comparisons at a familywise error rate of $\alpha = .05$, the critical value of $t = 2.34$. There is clearly not a significant difference between young and old subjects on tasks requiring little cognitive processing, but there is a significant difference for tasks requiring substantial cognitive processing. The probability that *at least* one of these statements represents a Type I error is at most .05.

16.19 Effect size for WL versus SIT

$$\hat{d} = \frac{\overline{X}_{WL} - \overline{X}_{SIT}}{s_{WL}} = \frac{19.50 - 11.07}{7.11} = \frac{8.43}{7.11} = 1.18$$

The two groups differ by over a standard deviation.

16.21 Spilich *et al*. data on a cognitive task:

```
Variable  ERRORS
   By Variable  SMOKEGRP
                                  Analysis of Variance

                             Sum of         Mean              F      F
      Source         D.F.    Squares        Squares          Ratio  Prob.
Between Groups         2     2643.3778     1321.6889       4.7444  .0139
Within Groups         42    11700.4000      278.5810
Total                 44    14343.7778

                       Standard   Standard
Group    Count    Mean   Deviation    Error    95 Pct Conf Int for Mean

Grp 1      15   28.8667     14.6866     3.7921     20.7335  TO  36.9998
Grp 2      15   39.9333     20.1334     5.1984     28.7838  TO  51.0828
Grp 3      15   47.5333     14.6525     3.7833     39.4191  TO  55.6476

Total      45   38.7778     18.0553     2.6915     33.3534  TO  44.2022
```

Here we have a task that involves more cognitive involvement, and it does show a difference due to smoking condition. The non-smokers performed with fewer errors than the other two groups, although we will need to wait until the next exercise to see the multiple comparisons.

16.23 Spilich *et al*. data on driving simulation:

```
Variable  ERRORS
   By Variable  SMOKEGRP
                                 Analysis of Variance

                         Sum of         Mean              F      F
    Source        D.F.    Squares        Squares          Ratio  Prob.
Between Groups      2      437.6444      218.8222       9.2584  .0005
Within Groups      42      992.6667       23.6349
Total              44     1430.3111

                       Standard   Standard
Group    Count     Mean  Deviation  Error       95 Pct Conf Int for Mean

Grp 1      15    2.3333    2.2887      .5909      1.0659  TO       3.6008
Grp 2      15    6.8000    5.4406     1.4048      3.7871  TO       9.8129
Grp 3      15    9.9333    6.0056     1.5506      6.6076  TO      13.2591

Total      45    6.3556    5.7015      .8499      4.6426  TO       8.0685
```

Here we have a case in which the active smokers again performed worse than the non-smokers, and the differences are significant.

16.25. Analysis of three groups in Merrell's study of the effect of subjecting mice to Anthrax.

Oneway

Descriptives

TIME

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
1	24	335.7500	157.78418	32.20756	269.1236	402.3764	203	999
2	24	114.2500	36.15817	7.38075	98.9817	129.5183	48.0	191
3	24	1825.5417	103.17860	21.06124	1781.9732	1869.1102	1576	2050
Total	72	758.5139	772.99503	91.09834	576.8690	940.1588	48.0	2050

ANOVA

TIME

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	41576485.0	2	20788242.5	1692.436	.000
Within Groups	847528.958	69	12283.028		
Total	42424014.0	71			

16.27 $\hat{d}$ for Merrill's study

$$\hat{d}_1 = \frac{\bar{X}_{control} - \bar{X}_{Mozart}}{\sqrt{MS_{error}}} = \frac{335.75 - 114.25}{\sqrt{12,283.028}} = \frac{221.5}{110.83} = 2.00$$

$$\hat{d}_2 = \frac{\bar{X}_{control} - \bar{X}_{Anthrax}}{\sqrt{MS_{error}}} = \frac{335.75 - 1825.54}{\sqrt{12,283.028}} = \frac{-1489.79}{110.83} = -13.44$$

Whereas the Control and Mozart groups differ by about two standard deviations (which is a lot), the Anthrax group is about 13.44 standard deviations below the Control in performance.

16.29 Measures of effect for data in Exercise 16.28.

a)

$$\eta^2 = \frac{SS_{group}}{SS_{total}} = \frac{2.170}{6.499} = 0.33$$

$$\omega^2 = \frac{SS_{group} - (k-1)MS_{error}}{SS_{total} + MS_{error}} = \frac{2.170 - (5-1)0.173}{6.499 + 0.173} = \frac{1.478}{6.672} = 0.22$$

b) The two results differ because the first is somewhat biased, and becomes more biased as the number of treatment groups increases.

c) It makes most sense to calculate $\hat{d}$ for the difference between the most extreme treatments, which, in this case, involves comparing groups 1 and 5. The square root of MS_{error} will serve as our estimate of the standard deviation.

$$\hat{d} = \frac{\bar{X}_5 - \bar{X}_1}{\sqrt{MS_{error}}} = \frac{3.26 - 2.60}{\sqrt{0.173}} = \frac{0.66}{0.4159} = 1.59$$

The two extreme groups differ by over one and a half standard deviations.