On-Line Appendix B: Hand Calculation of Statistical Tests

1. Working out a within-subjects *t*-test using the hypothesis-testing method

This test (also known as a correlated samples *t*-test or dependent groups *t*-test) is used when a variable has been manipulated within subjects so that pairs of scores are obtained from the same source. The procedure is as follows:

1. State the null hypothesis ($H_0 : \mu_1 = \mu_2$) and the alternative hypothesis (for a two-tailed test, $H_1 : \mu_1 \neq \mu_2$).

2. Convert each pair of scores into a *difference score* (*D*) by subtracting the second score from the first.

3. Calculate the mean, *D,* and standard deviation, s_d, of the *N* difference scores (do not ignore the sign).

4. Compute a *t*-value, using the formula:

$$t = \frac{\bar{D}}{s_D / \sqrt{N}}$$

5. Compute the *degrees of freedom* for the test from the number of participants, where $df = N - 1$.

6. Use Table 2 in On-line Appendix C to find the critical value of *t* for that number of degrees of freedom at the designated significance level (e.g., $\alpha = .05$ or $.01$).

7. If the absolute (i.e., unsigned) value of *t* that is obtained exceeds the critical value then reject the null hypothesis (H_0) and conclude that the difference between means is statistically significant. Otherwise conclude that the result is not statistically significant and make no inferences about differences between the means.

The following is a worked example of a within-subjects *t*-test where participants' pairs of scores have been obtained on two variables, A and B. The chosen alpha level is .05.

Participant	**A**	**B**	***D* (A – B)**
1	71	53	18
2	62	36	26
3	54	51	3
4	36	19	17
5	25	30	–5
6	71	52	19
7	13	20	–7
8	52	39	13
$N = 8$			$\sum D = 84$

$H_0 : \mu_1 = \mu_2$
$H_1 : \mu_1 \neq \mu_2$

$$t = \frac{\bar{D}}{s_D / \sqrt{N}}$$

$$= \frac{10.5}{120.07 / \sqrt{8}}$$

$$= \frac{10.5}{12.07 / \sqrt{8}}$$

$$= \frac{10.5}{12.07 / 2.83}$$

$$= 2.46$$

$$\text{df} = n - 1$$
$$= 7$$

The obtained t-value is greater than the tabulated value of t with 7 *df* for $p = .05$ (i.e., 2.365), so the difference between means is significant, $t(7) = 2.46$, $p < .05$.

2. Working out a between-subjects *t*-test using the hypothesis-testing method

This test (also known as an independent groups *t*-test or a two-sample *t*-test) is used when a variable has been manipulated between subjects so that every score is obtained from a different source (normally a different person). It is carried out as follows:

1. State the null hypothesis (H_0: $\mu_1 = \mu_2$) and the alternative hypothesis (for a two-tailed test, H_1: $\mu_1 \neq \mu_2$).

2. Compute means ($\bar{X}_1$ and $\bar{X}_2$) and variances (s_1^2 and s_2^2) for the scores in each experimental condition (where there are **N**1 scores in the first condition and N_2 scores in the second).

3. Compute a pooled variance estimate using the following formula:

$$s^2_{\text{pooled}} = \frac{(N_1 - 1)s_1^2 + (N_2 - 1)s_2^2}{N_1 + N_2 - 2}$$

4. Compute a *t*-value using the following formula:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{s^2_{\text{pooled}}\left(\frac{1}{N_1} + \frac{1}{N_2}\right)}}$$

5. Compute the degrees of freedom for the test from the sample sizes, where $df = N_1 + N_2 - 2$.

6. Use Table 2 in On-line Appendix C to identify the critical value of *t* for that number of degrees of freedom at the designated significance level (e.g., $\alpha = .05$).

7. If the absolute (i.e., unsigned) value of *t* that is obtained exceeds the critical value then reject the null hypothesis (H_0) and conclude that the difference between means is statistically significant. Otherwise conclude that the result is not statistically significant, and make no inferences about differences between the means.

The following is a worked example of a between-subjects *t*-test where scores have been obtained from participants in an experimental and a control group. The chosen alpha level is .05.

Experimental	**Control**
21	12
34	27
24	21
37	17
27	24
25	19
29	22
33	31
17	9
24	
$N_1 = 10$	$N_2 = 9$
$M = 27.10$	$M = 20.\ 22$
$s_1^2 = 38.54$	$s_2^2 = 48.19$

$H_0 : \mu_1 = \mu_2$
$H_1 : \mu_1 \neq \mu_2$

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{s^2_{\text{pooled}}\left(\frac{1}{N_1} + \frac{1}{N_2}\right)}}$$

Where

$$s^2_{\text{pooled}} = \frac{(N_1 - 1)s_1^2 + (N_2 - 1)s_2^2}{N_1 + N_2 - 2}$$

$$= \frac{((10-1)38.54) + ((9-1)48.19)}{10+9-2}$$

$$= \frac{9*38.54 + 8*48.19}{17}$$

$$= (346.86 + 385.52)/17$$

$$= 43.09$$

So

$$t = \frac{27.10 - 20.22}{\sqrt{43.08\left(\frac{1}{10} + \frac{1}{9}\right)}}$$

$$= 6.88/\sqrt{43.08*.21}$$

$$= 6.88/\sqrt{9.05}$$

$$= 2.28$$

$$df = N_1 + N_2 - 2$$
$$= 17$$

The obtained *t*-value is greater than the tabulated value of *t* with 17 *df* for $p = .05$ (i.e., 2.110), so the difference between means is significant, $t(17) = 2.29$, $p < .05$.

3. Working out a correlation using both hypothesis-testing and effect-size methods

This procedure is used to assess the relationship between two variables where pairs of scores on each variable have been obtained from the same source. It is carried out as follows:

1. State the null hypothesis (H_0: $\rho = 0$) and the alternative hypothesis (for a two-tailed test, $H_1 : \rho \neq 0$).

2. Draw a scatterplot, representing the position of each of the *n* pairs of scores on the two variables (X and γ).

3. Create a table which contains values of X^2 , γ^2 and $X\gamma$ for each pair of scores, and sums values of X, X^2 , γ, γ^2 and X γ.

4. Compute Pearson's *r*, using the formula:

$$r = \frac{N\sum Xr - \sum Xr}{\sqrt{\left(N\sum X^2 - \left(\sum X\right)^2\right)\left(N\sum r^2 - \left(\sum r\right)^2\right)}}$$

5. Compute the *degrees of freedom* for the test from the number of participants, where $df = n - 2$.

6. Use Table 3 in On-line Appendix C to identify the critical value of *r* for that number of degrees of freedom at the designated significance level (e.g., $\alpha = .05$).

7. If the absolute (i.e., unsigned) value of *r* exceeds the critical value then reject the null hypothesis (H_0) that there is no relationship between the variables and conclude that the relationship between variables is statistically significant. Otherwise conclude that the relationship is not statistically significant.

8. Consider the statistical significance of $| r |$ (i.e., the absolute value of *r*) in the context of the effect size, noting that where $.1 \leq | r | < .3$ the correlation between variables is only weak, where $.3 \leq | r | < .5$ the correlation is moderate and where $| r | \geq .5$ the correlation is strong.

Insert Figure A.1

Figure A.1 Plot of data for worked example of correlation

The following is a worked example of a correlation where participants' pairs of scores have been obtained on two variables, X and Y (see Figure A.1). The chosen alpha level is .05.

Participant	**X**	**X**	**γ**	**γ^2**	**Xγ**
1	16	56	33	1089	528
2	9	841	1	144	348
3	14	196	4	1764	588
4	6	676	16	256	416
5	33	1089	11	121	363
6	5	6 5	17	289	425
7	16	56	9	841	464
8	17	89	30	900	510
9	9	81	40	1600	360
10	1	441	1	441	441
N = 10	$\Sigma X = 06$	$\Sigma X^2 = 4750$	$\Sigma\gamma = 251$	$\Sigma\gamma^2 = 7445$	$\Sigma X\gamma = 4443$

$H_0 : \rho = 0$

$H_1 : \rho = 0$

$$r = \frac{N\sum Xr - \sum X \sum r}{\sqrt{\left(N\sum X^2 - \left(\sum X\right)^2\right)\left(N\sum r^2 - \left(\sum r\right)^2\right)}}$$

$$= \frac{(10*4443) - (206*251)}{\sqrt{\left(10*4750 - (206)^2\right)\left(10*7445 - (251)^2\right)}}$$

$$= \frac{44430 - 51706}{\sqrt{(47500 - 42436)(74450 - 63001)}}$$

$$= \frac{-7276}{\sqrt{5064*11449}}$$

$$= \frac{-7276}{7614.31}$$

$$= -.96$$

The absolute value of r is greater than the tabulated value of r with 8 *df* for $p = .05$ (i.e., .632 — and for $p = .01$; i.e., .765) and $r > .5$, so there is a significant strong negative correlation between the variables, $r = -.96$, $p < .01$.

4. Procedures for conducting one-way ANOVA with equal cell sizes

This is a rough-and-ready guide to analyzing experiments that involve more than two experimental conditions. As you study further in statistics you will discover more sophisticated strategies for analysis but the following procedures will get you started and, importantly, teach you no bad habits. At the same time they should help you to understand some very important statistical principles. Please bear in mind, though, that the following procedures can only be followed if cell sizes are equal. More advanced texts show the procedures to use where cells sizes are unequal. Note also that many of the steps cannot be performed by a computer package. We have marked these here with an asterisk.

1. Draw up a table with one row and as many columns as there are conditions in the study. Leave a space for each cell mean, squared deviation from grand mean, variance, and for the grand mean (note that it is customary to report the standard deviation — the square root of the variance — and not the variance when results are submitted for publication and in lab reports). In the top lef*t*-hand corner of each cell write a number to identify it. Work out the sample size *N*, the cell size and the number of cells *k* and write these down.

*2. From your notes, work out which cells (if any) you wished to compare before you collected the data (these are your planned comparisons). Write down these comparisons below the table in the form ‘Cell 2 > Cell 1’ etc. As a rule of thumb, with four or fewer cells and a sample size of 50 or less allow yourself no more planned comparisons than the number of cells.

*3. If adopting the hypothesis-testing approach, write down the null hypothesis for each planned comparison and for the overall analysis. These should take the form:

H_0: $\mu_1 = \mu_2 = \mu_3$ etc.

H_1: not all μs are equal

4. Calculate the cell means, grand mean, squared deviations from grand mean, and variances and fill them in the places provided.

5. Draw up an ANOVA table with columns for *Source, SS, df, U, F, ‘p <’* and R^2. Include a line in the ANOVA table for between cells, error (within cells) and total.

6. Calculate MS_W by pooling the within-cells variances. Add the variance for all cells and divide by the number of cells. Calculate df_w as *N-k.* Write these values in the ANOVA table.

7. Calculate df_B by subtracting one from the number of cells. Calculate MS_B by summing the total squared deviations of the cell means from the grand mean. Multiply this value by the cell size then divide by df_B. Write both values in the ANOVA table.

8. Calculate SS_B by multiplying MS_B by df_B. Calculate SS_w by multiplying MS_W by df_W. Write both values in the ANOVA table.

9. Calculate SS_T by adding SS_B and SS_w together. Write the value and $df_T = df_B + df_w$ in the ANOVA table.

10. Calculate the F-ratio by dividing MS_B by MS_W.

11. If $F < 1$ then write ns (non-significant) in the 'p <' column. Consult Table 4 in Online Appendix C and compare the F-value with the critical value for F with (df_B, df_W) degrees of freedom at the α = .05 level. If the obtained F is larger than this value then continue to compare it to other critical values for F at the .01 and .001 levels. Identify the smallest value of X for which the statement p < X is true.

12. Calculate R^2 for the effect by dividing SS_B by SS_T.

*13. Evaluate the planned comparisons. First set the protected alpha level by dividing your alpha level by the number of comparisons you are making. Then conduct a t-test for each planned comparison. To do this calculate the difference between the means, and divide by (the square root of MS_W multiplied by the square root of $2/n$), where n is the number of people in each cell. If the obtained t-value is greater than the critical value given by the protected alpha level conclude that it is significant.

*14. Make the statistical inferences from your ANOVA table. If the value of 'p <' in the table is less than your alpha level conclude that the effect is significant. Make some assessment of the size of the effect based on R^2.

The following is a worked example of a between-subjects one-way ANOVA where data have been obtained from four groups of four participants. Each group or cell represents a different experimental condition that a participant had been randomly assigned to.

Group 1	Group 2	Group 3	Group 4
7	5	3	2
6	7	6	3
4	3	4	5
5	4	3	3

$N = 16$, $n = 4$, $k=4$

Let us assume that we hypothesize that Group 1 will have a higher mean than Group 2 (this suggests the planned comparison Group 1 > Group 2). For the ANOVA,

$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$.
H_1 : not all as are equal

For planned comparison 1 > 2,
$H_0 : \mu_1 = \mu_2$

From these data we can construct a table showing means, squared deviations from the grand mean, and variances. Also shown is the grand mean (with equal cell sizes this is the mean of the cell means).

	Group 1	Group 2	Group 3	Group 4	Grand Mean
M	5.5	4.75	4.00	3.25	4.38
Squared deviations from grand mean	1.27	0.14	0.14	1.27	
Variance	1.67	2.92	2.00	1.58	

$MS_W = (\Sigma s^2_i)/k$
$= (1.67 + 2.92 + 2.00 + 1.58)/4$
$= 2.04$
$df_W = n - k$
$= 16 - 4 = 12$

$SS_w = MS_W * df_w = 2.04 * 12$
$= 24.50$
$df_b = k - 1$
$= 3$
$MS_B = \Sigma n(\bar{X}_i - \bar{X})^2/(k - 1)$
$= 4 *(1.25 + 0.14 + 0.14 + 1.28)/3$
$= 11 .25\ 3$
$= 3.75$

$SS_b = MS_B * df_B$
$= 11.24$

$MS_F = MS_B / MS_W$
$= 3.75/2.04$
$= 1.84$

$F_{.05}\ (3,12) = 3.49$

Source	***SS***	***df***	***MS***	***F***	***p* <**	***R*²**
Between cells (Condition)	11.25	3	3.75	1.84	ns	.31
Within cells	24.50	12	2.04			
Total	35.75	15				

The effect is of large size but the obtained *F*-value is smaller than the tabled or critical value of $F(3,12)$ for $\alpha = .05$ (i.e., 3.49) so so we cannot reject H_0 that the means are different (note that our design has limited power to detect effects due to the very small sample size). Our comparisons were planned so we proceed with them without worrying about the results of the ANOVA.

We are conducting one comparison, between Group 1 and 2, so

$$t = \left(\bar{X}_1 - \bar{X}_2\right) / \sqrt{MS_w * \left(2 / n\right)}$$
$$= \left(5.5 - 4.75\right) / \sqrt{2.04 * 2 / 4}$$
$$= 0.74$$

This is smaller than the tabulated one-tailed value of t with 12 *df* for $\alpha = .05$ (i.e., 1.782) so we cannot reject the null hypothesis that the means of the two groups are equal.

5. Procedures for conducting two-way ANOVA with equal cell sizes

1. Draw up a table with as many columns as there are levels of one variable and as many rows as there are levels of the other variable. Leave a space for each cell mean, row mean, column mean and grand mean along with the cell variance (note that here it is customary to report the standard deviation — the square root of the variance — and not the variance). In the top left-hand corner of each cell give the cell a number to identify it.

*2. From your notes work out which cells, rows or columns (if any) you wished to compare before you collected the data (planned comparisons). Write down these comparisons below the table in the form 'Cell 2 > Cell 1' or 'Row 3 < Row 4'. As a rule of thumb, with four or fewer cells and a sample size of 50 or less allow yourself no more planned comparisons than the number of cells.

*3. If adopting the hypothesis-testing approach, state the null hypothesis for each planned comparison and each effect. For the row effect:
   H_0 : Row means are equal
   H_1 : Row means are not equal for the column effect
   H_0 : Column means are equal
   H_1 : Column means are not equal and for the interaction effect
   H_0 : Cell means are equal to the values expected from the row and column effects
   H_1 : Cell means are not equal to the values expected from the row and column effects.

4. Calculate the means and variances and fill them in the places provided.

5. Draw a line graph of the means, with one line representing the means of each column. Examine the plot. Do the lines appear to be coming closer at some point (i.e., do they deviate from the parallel)? If they do, this suggests an interaction. Is one line higher than the other? If it is, this suggests a row main effect. Are two points at one level of the column factor higher than two other points? If they are, this suggests a column main effect.

6. Draw up an ANOVA table with columns for *Source, SS, df, MS, F, p* < and R^2. Include a line in the ANOVA table for between, interaction (row by column), row, column, error and total as sources.

7. Calculate MS_W by pooling the within-cells variances. Add the variance for all cells and divide by the number of cells. Calculate df_W as $n - k$, where n is the number of observations (normally participants) and k is the number of groups. Write these values in the ANOVA table.

8. Calculate df_B by subtracting one from the number of cells. Calculate MS_B by summing the total squared deviations of the cell means from the grand mean, multiply by the cell size and then divide by df_B. Write both values in the table.

9. Calculate SS_B by multiplying MS_B by df_B. Calculate SS_W by multiplying MS_W by df_W. Write both values in the table.

10. Calculate SS_R, SS_C and SS_{RC}: SS_R is the sum of the squared deviations between the row means and the grand mean, multiplied by the cell size; SS_C is the sum of the squared deviations between the column means and the grand mean, multiplied by the cell size; SS_{RC} is given by

$$SS_{RC} = SS_B - SS_R - SS_C.$$

For the 2 x 2 design the *df*s for row, column and row by column will all be 1.

11. Calculate SS_T by adding SS_B and SS_W together. Write the value in the ANOVA table.

12. Calculate MS_R by dividing SS_R by df_R. Calculate MS_C by dividing SS_C by df_C. Calculate MS_{RC} by dividing SS_{RC} by df_{RC}. Write all values in the ANOVA table.

13. Calculate the *F*-ratio for each effect by dividing each *MS* by MS_W.

14. If any $F < 1$ then write down *ns* (indicating that the effect was non-significant) in the $p <$ column. Otherwise, consult Table 4 in On-line Appendix C and compare the *F*-value with the critical value for *F* with (df_B, df_W) degrees of freedom at the .05 level. If obtained *F* is larger than this value, then continue to compare it to other critical values for *F* at the .01 and .001, levels. Write the smallest value of *X* for which the statement $p < X$ is true. If this statement is not true for any tabled value, write *ns* in the table.

15. Calculate R^2 for every effect by dividing the relevant *SS* by SS_T.

*16. Evaluate the planned comparisons. First set the protected alpha level by dividing your alpha level by the number of comparisons you are making. Then conduct a *t*-test for each planned comparison. To do this, calculate the difference between the means, and divide by the square root of MS_W multiplied by the square root of $2/n$, where *n* is the number of people in each cell. If the obtained *t*-value is greater than the critical value given by the protected alpha level, conclude that it is significant.

*17. Make the statistical inferences from your ANOVA table. If the value of '$p <$' in the table is less than your alpha level, conclude that the effect is significant. Make some assessment of the size of the effect based on R^2.

The following is a worked example of a between-subjects two-way ANOVA.