

# APPENDIX A - STATISTICAL CALCULATIONS

Calculation formulas and examples are provided for descriptive and inferential statistics discussed in Chapters 7 and 9.

## DESCRIPTIVE STATISTICS

These examples use the data set ( $X$  values): 4, 3, 5, 3, 2, 4, 5, 3, 2, 1.

Statistic	Formula/ Definition	Calculations
Mean ( $M$ )	$M = \frac{\sum X}{n}$	$M = \frac{(4 + 3 + 5 + 3 + 2 + 4 + 5 + 3 + 2 + 1)}{10} = \frac{32}{10} = 3.2$
Median	Middle score or average of middle scores	Order scores from lowest to highest: 1, 2, 2, 3, 3, 3, 4, 4, 4, 5, 5 Average middle scores: $(3 + 3) / 2 = 3$
Mode	Most common score or scores	3
Range	Difference between highest and lowest scores	Range = $(5 - 1) = 4$
Variance	$\text{Variance} = \frac{\sum(X - M)^2}{n - 1}$	$\text{Variance} = \frac{[(4 - 3.2)^2 + (3 - 3.2)^2 + (5 - 3.2)^2 + (3 - 3.2)^2 + (2 - 3.2)^2 + (4 - 3.2)^2 + (5 - 3.2)^2 + (3 - 3.2)^2 + (2 - 3.2)^2 + (1 - 3.2)^2]}{(10 - 1)} = 1.73$
Standard deviation ( $SD$ )	$SD = \sqrt{\frac{\sum(X - M)^2}{n - 1}}$	$SD = \sqrt{\frac{[(4 - 3.2)^2 + (3 - 3.2)^2 + (5 - 3.2)^2 + (3 - 3.2)^2 + (2 - 3.2)^2 + (4 - 3.2)^2 + (5 - 3.2)^2 + (3 - 3.2)^2 + (2 - 3.2)^2 + (1 - 3.2)^2]}{(10 - 1)}} = 1.32$

Note:  $\Sigma$  = sum all values after this symbol;  $n$  = number of scores.

## INFERENCE STATISTICS

Test	Example
One-sample $t$ test	<p>Sample values [<math>X</math> values]: 4, 3, 5, 3, 2, 4, 5, 3, 2, 1</p> <p>Population mean (<math>\mu</math>) = 3.0, <math>n</math> is the number of scores in the sample</p> $t = \frac{(M - \mu)}{SD/\sqrt{n}} = \frac{(3.2 - 3.0)}{1.32/\sqrt{10}} = \frac{.2}{.42} = .48.$ <p>This calculated <math>t</math> value is then compared with the appropriate <math>t</math> critical value (based on alpha and <math>df</math> of [<math>n - 1</math>]) to determine significance. See Statistical Tables section of this appendix.</p>
Independent samples $t$ test	<p>Sample 1 values [<math>X_1</math> values]: 4, 3, 5, 3, 2, 4, 5, 3, 2, 1</p> <p>Sample 2 values [<math>X_2</math> values]: 2, 3, 2, 3, 2, 2, 4, 3, 2, 2</p> $t = \frac{(M_1 - M_2)}{\sqrt{s_p^2/n_1 + s_p^2/n_2}}$ <p><math>s_p^2</math> [pooled variance]—based on variance of each sample and <math>n</math> for each sample</p> $s_p^2 = \frac{(n_1 - 1)(SD_1^2) + (n_2 - 1)(SD_2^2)}{(n_1 - 1) + (n_2 - 1)}$ <p><math>SD_1^2 = 1.73</math> (from Descriptive Statistics section)</p> $M_2 = \frac{2 + 3 + 2 + 3 + 2 + 2 + 4 + 3 + 2 + 2}{10} = 2.5$ $SD_2^2 = \frac{[(2 - 2.5)^2 + (3 - 2.5)^2 + (2 - 2.5)^2 + (3 - 2.5)^2 + (2 - 2.5)^2 + (2 - 2.5)^2 + (4 - 2.5)^2 + (3 - 2.5)^2 + (2 - 2.5)^2]}{(10 - 1)} = .5$ $s_p^2 = \frac{(10 - 1)(1.73) + (10 - 1)(.5)}{(10 - 1) + (10 - 1)} = \frac{15.57 + 4.5}{18} = 1.12$ $t = \frac{(3.2 - 2.5)}{\sqrt{1.12/10 + 1.12/10}} = \frac{.7}{.47} = 1.49$ <p>This calculated <math>t</math> value is then compared with the appropriate <math>t</math> critical value (based on alpha and <math>df</math> of [<math>n_1 + n_2 - 2</math>]) to determine significance. See Statistical Tables section of this appendix.</p>

Test	Example
Related/ paired samples <i>t</i> test	<p>Condition 1 values: 4, 3, 5, 3, 2, 4, 5, 3, 2, 1</p> <p>Condition 2 values from same or matched subjects: 2, 3, 2, 3, 2, 2, 4, 3, 2, 2</p> <p><i>t</i> is calculated in this test based on difference scores for each participant or across matched pairs. Thus, difference scores are first calculated by subtracting the score in Condition 2 from the score for Condition 1 for each participant (e.g., 4 – 2 for Participant 1).</p> <p>Difference scores (<i>D</i> values): 2, 0, 3, 0, 0, 2, 1, 0, 0, –1</p> $= \frac{(2 + 0 + 3 + 0 + 0 + 2 + 1 + 0 + 0 - 1)}{10} = .7$ $SD \text{ of } D (SD_D) = \sqrt{\frac{\sum(D - M_D)^2}{n - 1}} =$ $\sqrt{\frac{[(2 - .7)^2 + (0 - .7)^2 + (3 - .7)^2 + (0 - .7)^2 + (0 - .7)^2 + (2 - .7)^2 + (1 - .7)^2 + (0 - .7)^2 + (0 - .7)^2 + (-1 - .7)^2]}{(10 - 1)}} = \sqrt{\frac{14.1}{9}} = 1.25$ $t = \frac{M_D}{SD_D/\sqrt{n}} = \frac{.7}{1.25/\sqrt{10}} = \frac{.7}{.4} = 1.75$ <p>This calculated <i>t</i> value is then compared with the appropriate <i>t</i> critical value (based on alpha and <i>df</i> of [<i>n</i> – 1]) to determine significance. See Statistical Tables section of this appendix.</p>
One-way, between- subjects ANOVA	<p>Sample 1 (<i>X</i><sub>1</sub> values): 10, 15, 20, 13 (<i>M</i><sub>1</sub> = 14.50)</p> <p>Sample 2 (<i>X</i><sub>2</sub> values): 15, 25, 30, 25 (<i>M</i><sub>2</sub> = 23.75)</p> <p>Sample 3 (<i>X</i><sub>3</sub> values): 20, 17, 25, 20 (<i>M</i><sub>3</sub> = 20.50)</p> <p>Overall mean for all samples (<i>M</i><sub><i>T</i></sub>) = 19.58</p> $F = \frac{SS_A/a - 1}{SS_{\text{error}}/a(n - 1)}$ <p><i>SS</i><sub><i>A</i></sub> = <i>n</i>∑(<i>M</i><sub><i>A</i></sub> – <i>M</i><sub><i>T</i></sub>)<sup>2</sup> Sum of squared deviations for Factor A term [<i>SS</i><sub><i>A</i></sub>], <i>n</i> = number of scores per sample, <i>M</i><sub><i>A</i></sub> indicates each sample mean</p> $SS_A = 4[(14.5 - 19.58)^2 + (23.75 - 19.58)^2 + (20.5 - 19.58)^2]$ $= 4[25.81 + 17.39 + .85] = 4[44.05] = 176.2$ <p><i>a</i> – 1 = 3 – 1 = 2, <i>a</i> = number of groups in Factor A (<i>df</i> for Factor A term)</p> <p>Sum of squared deviations for error term [<i>SS</i><sub>error</sub>], <i>X</i> is an individual score, <i>M</i><sub><i>A</i></sub> is the sample mean for that score.</p> $SS_{\text{error}} = [(10 - 14.5)^2 + (15 - 14.5)^2 + (20 - 14.5)^2 + (13 - 14.5)^2 + (15 - 23.75)^2 + (25 - 23.75)^2 + (30 - 23.75)^2 + (25 - 23.75)^2 + (20 - 20.5)^2 + (17 - 20.5)^2 + (25 - 20.5)^2 + (20 - 20.5)^2]$ $= [20.25 + .25 + 30.25 + 2.25 + 76.56 + 1.56 + 39.06 + 1.56 + .25 + 12.25 + 20.25 + .25] = 204.74$ <p><i>a</i> (<i>n</i> – 1) = 3 (4 – 1) = 9 (<i>df</i> for error term)</p> $F = \frac{176.20/2}{204.74/9} = \frac{88.1}{22.75} = 3.87$ <p>This calculated <i>F</i> value is then compared with the appropriate <i>F</i> critical value (based on alpha and <i>df</i> values) to determine significance. See Statistical Tables section of this appendix.</p>

(Continued)

(Continued)

Test	Example
Factorial, between-subjects ANOVA	<p>Samples for two factors (A and B), each with two levels (1 and 2)</p> <p>Sample A1, B1 (<math>X_{A_1, B_1}</math> values): 67, 80, 75, 77, 78  <math>(M_{A_1, B_1} = 75.40)</math></p> <p>Sample A2, B1 (<math>X_{A_2, B_1}</math> values): 65, 59, 70, 62, 60  <math>(M_{A_2, B_1} = 63.20)</math></p> <p>Sample A1, B2 (<math>X_{A_1, B_2}</math> values): 59, 60, 67, 52, 60  <math>(M_{A_1, B_2} = 59.60)</math></p> <p>Sample A2, B2 (<math>X_{A_2, B_2}</math> values): 87, 90, 82, 86, 85  <math>(M_{A_2, B_2} = 86.00)</math></p> <p>Overall means for Factor A: <math>M_{A_1} = 67.50</math>, <math>M_{A_2} = 74.6</math></p> <p>Overall means for Factor B: <math>M_{B_1} = 69.30</math>, <math>M_{B_2} = 72.80</math></p> <p>Overall mean for all scores: <math>M_T = 71.05</math></p> <p><math>a = 2</math> (number of levels for Factor A)  <math>b = 2</math> (number of levels for Factor B)  <math>n = 5</math> (number of scores per group)</p> <p>Three effects tested: Main effect of Factor A (<math>F_A</math>), Main effect of Factor B (<math>F_B</math>), and Interaction of A and B (<math>F_{A \times B}</math>)</p> $F_A = \frac{SS_A/a - 1}{SS_{error}/ab(n - 1)}$ <p><math>(a - 1)</math> is <math>df</math> for Factor A term, <math>ab(n - 1)</math> is <math>df</math> for error term</p> $F_B = \frac{SS_B/b - 1}{SS_{error}/ab(n - 1)}$ <p><math>(b - 1)</math> is <math>df</math> for Factor B term</p> $F_{A \times B} = \frac{SS_{A \times B}/(a - 1)(b - 1)}{SS_{error}/ab(n - 1)}$ <p><math>(a - 1)(b - 1)</math> is <math>df</math> for interaction <math>A \times B</math> term</p> $SS_A = 5(2) [(67.5 - 71.05)^2 + (74.6 - 71.05)^2] = 252$ $SS_{error} = [67 - 75.4]^2 + [80 - 75.4]^2 + [75 - 75.4]^2 + [77 - 75.4]^2 + [78 - 75.4]^2 + [65 - 63.2]^2 + [59 - 63.2]^2 + [70 - 63.2]^2 + [62 - 63.2]^2 + [60 - 63.2]^2 + [59 - 59.6]^2 + [60 - 59.6]^2 + [67 - 59.6]^2 + [52 - 59.6]^2 + [60 - 59.6]^2 + [87 - 86]^2 + [90 - 86]^2 + [82 - 86]^2 + [86 - 86]^2 + [85 - 86]^2 = 327.2$ $F_A = \frac{252/(2 - 1)}{327.2/[2(2)(5 - 1)]} = \frac{252}{20.45} = 12.32$ $SS_B = 5(2) [(69.3 - 71.05)^2 + (72.8 - 71.05)^2] = 61.25$ $F_B = \frac{61.25/2 - 1}{327.2/[2(2)(5 - 1)]} = \frac{61.25}{20.45} = 3.00$ <p><math>SS_{A \times B} = n \sum (M_{AB} - M_A - M_B + M_T)^2</math>, where <math>M_{AB}</math> is the <math>AB</math> condition mean, <math>M_A</math> is the mean for that level of Factor A, and <math>M_B</math> is the mean for that level of Factor B.</p> $SS_{A \times B} = 5 [(75.4 - 67.5 - 69.3 + 71.05)^2 + (86 - 74.6 - 72.8 + 71.05)^2] = 931.20$ $F_{A \times B} = \frac{931.20/[(2 - 1)]}{327.2/[2(2)(5 - 1)]} = \frac{931.20}{20.45} = 45.54$ <p>Each of these calculated <math>F</math> values is then compared with the appropriate <math>F</math> critical value (based on alpha and <math>df</math> values) to determine significance. See Statistical Tables section of this appendix.</p>

Test	Example
Pearson $r$ correlation	<p>Measure 1 values (<math>X</math> values): 4, 3, 5, 3, 2, 4, 5, 3, 2, 1 (<math>M_x = 3.2</math>)</p> <p>Measure 2 values (<math>Y</math> values): 2, 3, 2, 3, 2, 2, 4, 3, 2, 2 (<math>M_y = 2.5</math>)</p> $r = \frac{\Sigma[(X - M_x)(Y - M_y)]}{\sqrt{\Sigma[(X - M_x)^2(Y - M_y)^2]}}$ $= \frac{[(4 - 3.2)(2 - 2.5) + (3 - 3.2)(3 - 2.5) + (5 - 3.2)(2 - 2.5) + (3 - 3.2)(3 - 2.5) + (2 - 3.2)(2 - 2.5) + (4 - 3.2)(2 - 2.5) + (5 - 3.2)(4 - 2.5) + (3 - 3.2)(3 - 2.5) + (2 - 3.2)(2 - 2.5) + (1 - 3.2)(2 - 2.5)]}{\sqrt{[(4 - 3.2)^2 + (3 - 3.2)^2 + (5 - 3.2)^2 + (3 - 3.2)^2 + (2 - 3.2)^2 + (4 - 3.2)^2 + (5 - 3.2)^2 + (3 - 3.2)^2 + (2 - 3.2)^2 + (1 - 3.2)^2]}}$ $= \frac{[-.4 + -.1 + -.9 + -.1 + .6 + -.4 + 2.7 + -.1 + .6 + 1.1]}{\sqrt{[0.64 + 0.04 + 3.24 + 0.04 + 1.44 + 0.64 + 3.24 + 0.04 + 1.44 + 4.84]}}$ $= \frac{3}{\sqrt{[15.6][4.5]}} = \frac{3}{8.38} = .36$ <p>This calculated <math>r</math> value is then compared with the appropriate <math>r</math> critical value (based on alpha and <math>df</math> of <math>[n - 2]</math>) to determine significance. See Statistical Tables section of this appendix.</p>
Linear regression	<p>Predictor variable values (<math>X</math> values): 4, 3, 5, 3, 2, 4, 5, 3, 2, 1 (<math>M_x = 3.2</math>)</p> <p>Outcome variable values (<math>Y</math> values): 2, 3, 2, 3, 2, 2, 4, 3, 2, 2 (<math>M_y = 2.5</math>)</p> <p><math>\hat{Y} = b - X + a</math>: regression line to predict <math>Y</math> value (<math>\hat{Y}</math>) from the value of <math>X</math>, must calculate <math>b</math> (slope of the line) and <math>a</math> (the intercept of the line which is the value of <math>Y</math> when <math>X = 0</math>)</p> $b = \frac{\Sigma(X - M_x)(Y - M_y)}{\Sigma(X - M_x)^2}$ <p>the numerator of this equation is equivalent to the numerator of the Pearson <math>r</math> calculation shown in that section (3.00); the denominator is the sum of squared deviations for the predictor variable <math>X</math> (<math>SS_x</math>).</p> $SS_x = [(4 - 3.2)^2 + (3 - 3.2)^2 + (5 - 3.2)^2 + (3 - 3.2)^2 + (2 - 3.2)^2 + (4 - 3.2)^2 + (5 - 3.2)^2 + (3 - 3.2)^2 + (2 - 3.2)^2 + (1 - 3.2)^2] = 15.6$ <p><math>b = \frac{3}{15.6} = .19</math>. This is the slope of the "best fit" line.</p> <p><math>a = M_y - b(M_x) = 2.5 - .19(3.2) = 1.89</math>. This is the intercept for the "best fit" line.</p> <p>Thus, the regression equation is</p> $\hat{Y} = .19X + 1.89.$ <p>This equation can be used to predict a value of the outcome variable <math>Y</math> from the predictor variable <math>X</math>.</p>

(Continued)

(Continued)

Test	Example																
Pearson chi-square test	Frequency data from Example in Chapter 9:																
	<table border="1"> <thead> <tr> <th></th> <th>Older</th> <th>Younger</th> <th>Total</th> </tr> </thead> <tbody> <tr> <th>Yes</th> <td>8</td> <td>4</td> <td>12</td> </tr> <tr> <th>No</th> <td>2</td> <td>6</td> <td>8</td> </tr> <tr> <th>Total</th> <td>10</td> <td>10</td> <td></td> </tr> </tbody> </table>		Older	Younger	Total	Yes	8	4	12	No	2	6	8	Total	10	10	
		Older	Younger	Total													
	Yes	8	4	12													
	No	2	6	8													
	Total	10	10														
	The first step is to calculate expected frequencies for each combination of these factors if there is no relationship between these factors:																
	$f_e = \frac{f_{\text{column}}(f_{\text{row}})}{n}$																
	Expected older (yes): $f_e = \frac{10(12)}{20} = 6$																
	Expected older (no): $f_e = \frac{10(8)}{20} = 4$																
Expected younger (yes): $f_e = \frac{10(12)}{20} = 6$																	
Expected younger (no): $f_e = \frac{10(8)}{20} = 4$																	
Expected frequencies:																	
<table border="1"> <thead> <tr> <th></th> <th>Older</th> <th>Younger</th> </tr> </thead> <tbody> <tr> <th>Yes</th> <td>6</td> <td>6</td> </tr> <tr> <th>No</th> <td>4</td> <td>4</td> </tr> </tbody> </table>		Older	Younger	Yes	6	6	No	4	4								
	Older	Younger															
Yes	6	6															
No	4	4															
The chi-square statistic is calculated from the difference between the observed and expected frequencies:																	
$\chi^2 = \sum \frac{(f_{\text{observed}} - f_{\text{expected}})^2}{f_{\text{expected}}} = \frac{(8-6)^2}{6} + \frac{(4-6)^2}{6} + \frac{(2-4)^2}{4} + \frac{(6-4)^2}{4}$ $= .67 + .67 + 1 + 1 = 3.34$																	
This calculated $\chi^2$ value is then compared with the appropriate $\chi^2$ critical value (based on alpha and $df$ of number of columns $\times$ number of rows) to determine significance. See Statistical Tables section of this appendix.																	

## STATISTICAL TABLES

The tables in this section provide critical statistical values for the calculations in the section Inferential Statistics.

### Critical $t$ Values

<i>One-Tailed Test</i>						
	0.25	0.10	0.05	0.025	0.01	0.005
<i>Two-Tailed Test</i>						
<i>df</i>	0.50	0.20	0.10	0.05	0.02	0.01
1	1.000	3.078	6.314	12.706	31.821	63.657
2	0.816	1.886	2.920	4.303	6.965	9.925
3	0.765	1.638	2.353	3.182	4.541	5.841
4	0.741	1.533	2.132	2.776	3.747	4.604
5	0.727	1.476	2.015	2.571	3.365	4.032
6	0.718	1.440	1.943	2.447	3.143	3.707
7	0.711	1.415	1.895	2.365	2.998	3.499
8	0.706	1.397	1.860	2.306	2.896	3.355
9	0.703	1.383	1.833	2.262	2.821	3.250
10	0.700	1.372	1.812	2.228	2.764	3.169
11	0.697	1.363	1.796	2.201	2.718	3.106
12	0.695	1.356	1.782	2.179	2.681	3.055
13	0.694	1.350	1.771	2.160	2.650	3.012
14	0.692	1.345	1.761	2.145	2.624	2.977
15	0.691	1.341	1.753	2.131	2.602	2.947
16	0.690	1.337	1.746	2.120	2.583	2.921
17	0.689	1.333	1.740	2.110	2.567	2.898
18	0.688	1.330	1.734	2.101	2.552	2.878
19	0.688	1.328	1.729	2.093	2.539	2.861
20	0.687	1.325	1.725	2.086	2.528	2.845
21	0.686	1.323	1.721	2.080	2.518	2.831

(Continued)

(Continued)

<i>One-Tailed Test</i>						
	0.25	0.10	0.05	0.025	0.01	0.005
<i>Two-Tailed Test</i>						
<i>df</i>	0.50	0.20	0.10	0.05	0.02	0.01
22	0.686	1.321	1.717	2.074	2.508	2.819
23	0.685	1.319	1.714	2.069	2.500	2.807
24	0.685	1.318	1.711	2.064	2.492	2.797
25	0.684	1.316	1.708	2.060	2.485	2.787
26	0.684	1.315	1.706	2.056	2.479	2.779
27	0.684	1.314	1.703	2.052	2.473	2.771
28	0.683	1.313	1.701	2.048	2.467	2.763
29	0.683	1.311	1.699	2.045	2.462	2.756
30	0.683	1.310	1.697	2.042	2.457	2.750
40	0.681	1.303	1.684	2.021	2.423	2.704
60	0.679	1.296	1.671	2.000	2.390	2.660
120	0.677	1.289	1.658	1.980	2.358	2.617
$\infty$	0.674	1.282	1.645	1.960	2.326	2.576

### Critical *F* Values

$p = .05$  in bold

$p = .01$  in italics

Degrees of Freedom for Denominator	Degrees of Freedom for Numerator									
	1	2	3	4	5	6	7	8	9	10
1	<b>161.4</b>	<b>199.5</b>	<b>215.7</b>	<b>224.6</b>	<b>230.2</b>	<b>234.0</b>	<b>236.8</b>	<b>240.5</b>	<b>241.9</b>	<b>243.9</b>
	<i>4052</i>	<i>5000</i>	<i>5403</i>	<i>5625</i>	<i>5764</i>	<i>5859</i>	<i>5928</i>	<i>5981</i>	<i>6022</i>	<i>6056</i>
2	<b>18.51</b>	<b>19.00</b>	<b>19.16</b>	<b>19.25</b>	<b>19.30</b>	<b>19.33</b>	<b>19.35</b>	<b>19.37</b>	<b>19.38</b>	<b>19.40</b>
	<i>98.5</i>	<i>99.0</i>	<i>99.2</i>	<i>99.2</i>	<i>99.3</i>	<i>99.3</i>	<i>99.4</i>	<i>99.4</i>	<i>99.4</i>	<i>99.4</i>
3	<b>10.3</b>	<b>9.55</b>	<b>9.28</b>	<b>9.12</b>	<b>9.01</b>	<b>8.94</b>	<b>8.89</b>	<b>8.85</b>	<b>8.81</b>	<b>8.79</b>
	<i>34.1</i>	<i>30.8</i>	<i>29.5</i>	<i>28.7</i>	<i>28.2</i>	<i>27.9</i>	<i>27.7</i>	<i>27.5</i>	<i>27.3</i>	<i>27.2</i>



Degrees of Freedom for Denominator	Degrees of Freedom for Numerator									
	1	2	3	4	5	6	7	8	9	10
4	<b>7.71</b>	<b>6.94</b>	<b>6.59</b>	<b>6.39</b>	<b>6.26</b>	<b>6.16</b>	<b>6.09</b>	<b>6.04</b>	<b>6.00</b>	<b>5.96</b>
	21.2	18.0	16.7	16.0	15.5	15.2	15.0	14.8	14.7	14.5
5	<b>6.61</b>	<b>5.79</b>	<b>5.41</b>	<b>5.19</b>	<b>5.05</b>	<b>4.95</b>	<b>4.88</b>	<b>4.82</b>	<b>4.77</b>	<b>4.74</b>
	16.3	13.3	12.1	11.4	11.0	10.7	10.5	10.3	10.2	10.1
6	<b>5.99</b>	<b>5.14</b>	<b>4.76</b>	<b>4.53</b>	<b>4.39</b>	<b>4.28</b>	<b>4.21</b>	<b>4.15</b>	<b>4.10</b>	<b>4.06</b>
	13.7	10.9	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87
7	<b>5.59</b>	<b>4.74</b>	<b>4.35</b>	<b>4.12</b>	<b>3.97</b>	<b>3.87</b>	<b>3.79</b>	<b>3.73</b>	<b>3.68</b>	<b>3.64</b>
	12.2	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62
8	<b>5.32</b>	<b>4.46</b>	<b>4.07</b>	<b>3.84</b>	<b>3.69</b>	<b>3.58</b>	<b>3.50</b>	<b>3.44</b>	<b>3.39</b>	<b>3.35</b>
	11.3	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81
9	<b>5.12</b>	<b>4.26</b>	<b>3.86</b>	<b>3.63</b>	<b>3.48</b>	<b>3.37</b>	<b>3.29</b>	<b>3.23</b>	<b>3.18</b>	<b>3.14</b>
	10.6	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26
10	<b>4.96</b>	<b>4.10</b>	<b>3.71</b>	<b>3.48</b>	<b>3.33</b>	<b>3.22</b>	<b>3.14</b>	<b>3.07</b>	<b>3.02</b>	<b>2.98</b>
	10.0	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85
11	<b>4.84</b>	<b>3.98</b>	<b>3.59</b>	<b>3.36</b>	<b>3.20</b>	<b>3.09</b>	<b>3.01</b>	<b>2.95</b>	<b>2.90</b>	<b>2.85</b>
	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54
12	<b>4.75</b>	<b>3.89</b>	<b>3.49</b>	<b>3.26</b>	<b>3.11</b>	<b>3.00</b>	<b>2.91</b>	<b>2.85</b>	<b>2.80</b>	<b>2.75</b>
	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30
13	<b>4.67</b>	<b>3.81</b>	<b>3.41</b>	<b>3.18</b>	<b>3.03</b>	<b>2.92</b>	<b>2.83</b>	<b>2.77</b>	<b>2.71</b>	<b>2.67</b>
	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19	4.10
14	<b>4.60</b>	<b>3.74</b>	<b>3.34</b>	<b>3.11</b>	<b>2.96</b>	<b>2.85</b>	<b>2.76</b>	<b>2.70</b>	<b>2.65</b>	<b>2.60</b>
	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03	3.94
15	<b>4.54</b>	<b>3.68</b>	<b>3.29</b>	<b>3.06</b>	<b>2.90</b>	<b>2.79</b>	<b>2.71</b>	<b>2.64</b>	<b>2.59</b>	<b>2.54</b>
	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80
16	<b>4.49</b>	<b>3.63</b>	<b>3.24</b>	<b>3.01</b>	<b>2.85</b>	<b>2.74</b>	<b>2.66</b>	<b>2.59</b>	<b>2.54</b>	<b>2.49</b>
	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69
17	<b>4.45</b>	<b>3.59</b>	<b>3.20</b>	<b>2.96</b>	<b>2.81</b>	<b>2.70</b>	<b>2.61</b>	<b>2.55</b>	<b>2.49</b>	<b>2.45</b>
	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68	3.59

(Continued)

(Continued)

Degrees of Freedom for Denominator	Degrees of Freedom for Numerator									
	1	2	3	4	5	6	7	8	9	10
18	<b>4.41</b>	<b>3.55</b>	<b>3.16</b>	<b>2.93</b>	<b>2.77</b>	<b>2.66</b>	<b>2.58</b>	<b>2.51</b>	<b>2.46</b>	<b>2.41</b>
	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60	3.51
19	<b>4.38</b>	<b>3.52</b>	<b>3.13</b>	<b>2.90</b>	<b>2.74</b>	<b>2.63</b>	<b>2.54</b>	<b>2.48</b>	<b>2.42</b>	<b>2.38</b>
	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43
20	<b>4.35</b>	<b>3.49</b>	<b>3.10</b>	<b>2.87</b>	<b>2.71</b>	<b>2.60</b>	<b>2.51</b>	<b>2.45</b>	<b>2.39</b>	<b>2.35</b>
	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37
22	<b>4.30</b>	<b>3.44</b>	<b>3.05</b>	<b>2.82</b>	<b>2.66</b>	<b>2.55</b>	<b>2.46</b>	<b>2.40</b>	<b>2.34</b>	<b>2.30</b>
	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26
30	<b>4.17</b>	<b>3.32</b>	<b>2.92</b>	<b>2.69</b>	<b>2.53</b>	<b>2.42</b>	<b>2.33</b>	<b>2.27</b>	<b>2.21</b>	<b>2.16</b>
	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98
40	<b>4.08</b>	<b>3.23</b>	<b>2.84</b>	<b>2.61</b>	<b>2.45</b>	<b>2.34</b>	<b>2.25</b>	<b>2.18</b>	<b>2.12</b>	<b>2.08</b>
	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80
60	<b>4.00</b>	<b>3.15</b>	<b>2.76</b>	<b>2.53</b>	<b>2.37</b>	<b>2.25</b>	<b>2.17</b>	<b>2.10</b>	<b>2.04</b>	<b>1.99</b>
	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63
120	<b>3.92</b>	<b>3.07</b>	<b>2.68</b>	<b>2.45</b>	<b>2.29</b>	<b>2.17</b>	<b>2.09</b>	<b>2.02</b>	<b>1.96</b>	<b>1.91</b>
	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47
$\infty$	<b>3.84</b>	<b>3.00</b>	<b>2.60</b>	<b>2.37</b>	<b>2.21</b>	<b>2.10</b>	<b>2.00</b>	<b>1.94</b>	<b>1.88</b>	<b>1.83</b>
	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32

### Critical $r$ Values

<i>One-Tailed Test</i>				
	0.05	0.025	0.01	0.005
<i>Two-Tailed Test</i>				
<i>df</i>	0.10	0.05	0.02	0.01
1	.988	.997	.9995	.9999
2	.900	.950	.980	.990
3	.805	.878	.934	.959
4	.729	.811	.882	.917

<i>One-Tailed Test</i>				
	0.05	0.025	0.01	0.005
<i>Two-Tailed Test</i>				
<i>df</i>	0.10	0.05	0.02	0.01
5	.669	.754	.833	.874
6	.622	.707	.789	.834
7	.582	.666	.750	.798
8	.549	.632	.716	.765
9	.521	.602	.685	.735
10	.497	.576	.658	.708
11	.476	.553	.634	.684
12	.458	.532	.612	.661
13	.441	.514	.592	.641
14	.426	.497	.574	.623
15	.412	.482	.558	.606
16	.400	.468	.542	.590
17	.389	.456	.528	.575
18	.378	.444	.516	.561
19	.369	.433	.503	.549
20	.360	.423	.492	.537
21	.352	.413	.482	.526
22	.344	.404	.472	.515
23	.337	.396	.462	.505
24	.330	.388	.453	.496
25	.323	.381	.445	.487
26	.317	.374	.437	.479
27	.311	.367	.430	.471
28	.306	.361	.423	.463
29	.301	.355	.416	.456
30	.296	.349	.409	.449

(Continued)

(Continued)

<b>One-Tailed Test</b>				
	0.05	0.025	0.01	0.005
<b>Two-Tailed Test</b>				
<i>df</i>	0.10	0.05	0.02	0.01
35	.275	.325	.381	.418
40	.257	.304	.358	.393
45	.243	.288	.338	.372
50	.231	.273	.322	.354
60	.211	.250	.295	.325
70	.195	.232	.274	.302
80	.183	.217	.256	.283
90	.173	.205	.242	.267
100	.164	.195	.230	.254

### Critical $X^2$ Values

<b>Level of Significance</b>			
<i>df</i>	0.05	0.025	0.01
1	3.84	5.02	6.64
2	5.99	7.38	9.21
3	7.81	9.35	11.34
4	9.49	11.14	13.28
5	11.07	12.83	15.09
6	12.59	14.45	16.81
7	14.07	16.01	18.48
8	15.51	17.53	20.09
9	16.92	19.02	21.67
10	18.31	20.48	23.21
11	19.68	21.92	24.72
12	21.03	23.34	26.22
13	22.36	24.74	27.69

Level of Significance			
<i>df</i>	0.05	0.025	0.01
14	23.68	26.11	29.14
15	25.00	27.49	30.58
16	26.30	28.85	32.00
17	27.59	30.19	33.41
18	28.87	31.53	34.80
19	30.14	32.85	36.19
20	31.41	34.17	37.57
21	32.67	35.48	38.93
22	33.92	36.78	40.29
23	35.17	38.08	41.64
24	36.42	39.36	42.98
25	37.65	40.65	44.31
26	38.88	41.92	45.64
27	40.11	43.19	46.96
28	41.34	44.46	48.28
29	42.56	45.72	49.59
30	43.77	46.98	50.89
40	55.76	59.34	63.69
50	67.50	71.42	76.15

## SPSS OUTPUT

Figure A.1 shows SPSS output for the example data sets discussed in Chapter 9.

### Example 1: One-Sample $t$ Test

**FIGURE A.1** ■ Output Window From the One-Sample  $t$  Test for Example 1

$t$  Test

#### One-Sample Statistics

	<i>N</i>	Mean	Std. Deviation	Std. Error Mean
Recognition Scores	10	53.5000	10.40566	3.29056

#### One-Sample Test

	Test Value = 50					
					95% Confidence Interval of the Difference	
	<i>t</i>	<i>df</i>	Sig. (two-tailed)	Mean Difference	Lower	Upper
Recognition Scores	1.064	9	.315	3.50000	-3.9438	10.9438

The output from the test contains several important values. The sample mean can be seen in the first box along with the standard deviation. These values are circled in the output in red in Figure A.1. These are the standard descriptive statistics included in the output for a  $t$  test. The  $t$  test values are included in the second box in the output. These values are circled in green in Figure A.1. The  $t$  value (1.064 for Example 1), the degrees of freedom (abbreviated  $df$ , calculated from the sample size; see discussion in Chapter 7), and the  $p$  value listed in the Sig. column (which stands for the level of significance or  $p$  value) of the box. The default test is a two-tailed test in SPSS for  $t$  tests, but you can convert the value to a one-tailed test by dividing the  $p$  value in half if the means differ in the predicted direction. (The one-tailed test has a critical region at one end of the  $t$  distribution that is twice the size of the critical region for a two-tailed test—thus, the one-tailed test has a  $p$  value that is half the  $p$  value for the two-tailed test.) Some newer versions of SPSS provide both the one-tailed and two-tailed  $p$  values in the output. The  $p$  value in the output for Example 1 is .315. In Example 1, if there is an effect, we expect the mean recognition score to be higher than 50%. In other words, a one-tailed test is warranted. Thus,

we must divide the given  $p$  value in half to obtain a  $p = .1575$  for this one-tailed  $t$  test. Because this value is *not* equal to or lower than alpha of .05 (the standard alpha level used in behavioral research), the null hypothesis cannot be rejected and must be retained. In other words, there is no evidence that participants in the sample experiment remembered the subliminal ads because their performance was not better than what is expected by chance. If you need to report the outcome of this test in APA style, you might include a statement such as, “The mean recognition score for subliminal ads ( $M = 53.50$ ) was not significantly higher than the chance value of 50%,  $t(9) = 1.06$ ,  $p = .16$ , one-tailed.” The statistical values (rounded here to two significant digits) are stated as support for a statement about the results of the study. If a two-tailed hypothesis had been made for this study, then the result would be reported as “Not significantly different.”

## Example 2: Independent-Samples $t$ Test

**FIGURE A.2** ■ Output Window From the Independent-Samples  $t$  Test for Example 2

$t$  Test

Independent Samples Test											
		Levene's Test for Equality of Variances		t-test for Equality of Means						95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper	
Recognition Scores	Equal variances assumed	2.217	.154	.107	18	.916	.6000	5.62870	-11.225	12.4255	
	Equal variances not assumed			.107	16.361	.916	.60000	5.62870	-11.311	12.5109	

### Group Statistics

	Gender	N	Mean	Std. Deviation	Std. Error Mean
Recognition Scores	Men	10	53.5000	10.40566	3.29056
	Women	10	52.9000	14.44107	4.56667

In Figure A.2, one box in the output provides descriptive statistics for each group. These values are circled in the output in red. The other box contains the test statistic and  $p$  value. These values are circled in the output in green. For Example 2, the  $t$  value is .107, and the  $p$  value (see the Sig. column) is .916. This is a two-tailed test (it is possible that either men or women could have higher recognition scores), so the  $p$  value given can be directly compared with alpha. Once again, the  $p$  value is not equal to or lower than the alpha of .05, and the null hypothesis cannot

be rejected. Thus, there is no evidence in these data that men and women differ in their memories for subliminal ads. If there was a significant difference, we could look at the means to determine which group was higher.

### Example 3: Paired-Samples $t$ Test

**FIGURE A.3** ■ Output Window From the Paired-Samples  $t$  Test for Example 3

$t$  Test

		Paired Samples Statistics			
		Mean	<i>N</i>	Std. Deviation	Std. Error Mean
Pair 1	Standard Ads	53.5000	10	10.40566	3.29056
	Emotional Ads	66.3000	10	12.68464	4.01123

#### Paired Samples Correlations

		<i>N</i>	Correlation	Sig.
Pair 1	Standard Ads and Emotional Ads	10	.118	.745

#### Paired Samples Test

		Paired Differences							
		95% Confidence Interval of the Difference							
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper	<i>t</i>	<i>df</i>	Sig. (2-tailed)
Pair 1	Standard Ads and Emotional Ads	-12.800	15.42581	4.87807	-23.835	-1.7650	-2.624	9	.028

The output in Figure A.3 indicates descriptive statistics in the first box (circled in the output in red) and the test statistic and  $p$  value in the third box (circled in green). For Example 3, the  $t$  value is  $-2.624$  with a  $p$  value of  $.028$ . For this example, the  $p$  value is lower than alpha of  $.05$ . Thus, the null hypothesis (that there is no difference between the ad types) can be rejected, and the alternative hypothesis (that there is a difference between the ad types) can be accepted. The means indicate which ad type was remembered better: In this case, the emotional ads ( $M = 66.3$ ) were remembered better than the standard ads ( $M = 53.5$ ),  $t(9) = -2.62$ ,  $p = .03$ . The second box of the output provides a test of the relationship between the two sets of scores (see the Chi-square and Pearson  $r$  tests below for more information about tests for relationships).



### Example 4: Between-Subjects ANOVA

**FIGURE A.4** ■ Output Window From the Between-Subjects ANOVA for Example 4

#### One way

#### Descriptives

Test Score - Percentage Correct

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean	
					Lower Bound	Upper Bound
Paper text	10	66.2000	13.08774	4.13871	56.8376	75.5624
Computer text	10	80.6000	11.23684	3.55340	72.5616	88.6384
Computer with voice text	10	83.0000	7.37865	2.33333	77.7216	88.2784
Total	30	76.6000	12.89106	2.35357	71.7864	81.4136

#### Descriptives

Test Score - Percentage Correct

	Minimum	Maximum
Paper text	45.00	90.00
Computer text	54.00	92.00
Computer with voice text	72.00	93.00
Total	45.00	93.00

#### ANOVA

Test Score - Percentage Correct

	Sum of Squares	df	Mean Square	F	Sig.
Between groups	1651.200	2	825.600	7.036	.003
Within groups	3168.000	27	117.333		
Total	4819.200	29			

## Post Hoc Tests

## Multiple Comparisons

Dependent Variable: Test Score - Percentage Correct

LSD (I) Text Condition	(J) Text Condition	Mean Difference (I - J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Paper text	Computer text	-14.4000*	4.84424	.006	-24.3396	-4.4604
	Computer with voice text	-16.8000*	4.84424	.002	-26.7396	-6.8604
Computer text	Paper text	14.40000*	4.84424	.006	4.4604	24.3396
	Computer with voice text	-2.40000	4.84424	.624	-12.3396	7.5396
Computer with voice text	Paper text	16.80000*	4.84424	.002	6.8604	26.7396
	Computer text	2.40000	4.84424	.624	-7.5396	12.3396

Note: \*The mean difference is significant at the .05 level.

The output window in Figure A.4 contains a box with descriptive statistics (if you choose that option) circled in red, a box with the  $F$  statistic and  $p$  value (circled in green), and a box with the post hoc tests (if you choose that option). The between-groups row of the ANOVA box indicates a significant effect of text condition with  $F = 7.036$  and a  $p$  value of .003. The test is significant because the  $p$  value is lower than an alpha of .05. In other words, we can reject the null hypothesis that there is no difference between the condition means. This result might be reported as, "The effect of text condition was significant,  $F(2, 27) = 7.04$ ,  $p = .003$ ." Although this test is significant, it does not tell us which conditions are different from each other.

The post hoc tests indicate which pairs of means are significantly different from one another. Post hoc tests are different from normal sets of  $t$  tests because they adjust for an inflated alpha level. Each time a pairwise comparison is made, it becomes more likely that a Type I error (see Chapter 7) is made. In other words, the likelihood of making a Type I error is higher for multiple  $t$  tests than for a single test. Thus, the post hoc tests keep alpha at .05 for each test by adjusting the alpha level for the set of tests based on the number of tests conducted. These tests are shown in the last box in the output. The LSD test was chosen in this example. The box shows  $p$  values (in the Sig. column circled in blue) lower than .05 for the paper text versus computer text; and paper text versus computer with voice text tests; but not for the computer text versus computer with voice text test. Thus, if we examine the means shown in the first box the output of Figure A.4, we see that both of the computer conditions resulted in higher test scores than the paper text condition, but the two computer conditions do not significantly differ from each other.

## Example 5: Within-Subjects ANOVA

**FIGURE A.5** ■ Output From Output Window for the Within-Subjects ANOVA for Example 5

### General Linear Model

#### Within-Subjects Factors

Measure: MEASURE\_1

textcond	Dependent Variable
1	paper
2	comp
3	compvoice

#### Descriptive Statistics

	Mean	Std.Deviation	N
Paper text	66.2000	13.08774	10
Computer text	80.6000	11.23684	10
Computer with voice text	83.0000	7.37865	10

#### Multivariate Tests<sup>a</sup>

Effect		Value	F	Hypothesis df	Error df	Sig.
textcond	Pillai's Trace	.591	5.784 <sup>b</sup>	2.000	8.000	.028
	Wilks' Lambda	.409	5.784 <sup>b</sup>	2.000	8.000	.028
	Hotelling's Trace	1.446	5.784 <sup>b</sup>	2.000	8.000	.028
	Roy's Largest Root	1.446	5.784 <sup>b</sup>	2.000	8.000	.028

<sup>a</sup>Design: Intercept  
Within-Subjects Design: textcond

<sup>b</sup>Exact statistic

#### Mauchly's Test of Sphericity<sup>a</sup>

Measure: MEASURE\_1

Within-Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.
textcond	.975	.205	2	.902

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

**Mauchly's Test of Sphericity<sup>a</sup>**

Measure: MEASURE\_1

Within-Subjects Effect	Epsilon <sup>b</sup>		
	Greenhouse-Geisser	Huynh-Feldt	Lower bound
textcond	.975	1.000	.500

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. Design: Intercept

Within-Subjects Design: textcond

b. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

**Tests of Within-Subjects Effects**

Measure: MEASURE\_1

Source		Type III Sum of Squares	df	Mean Square
textcond	Sphericity Assumed	1651.200	2	825.600
	Greenhouse-Geisser	1651.200	1.951	846.514
	Huynh-Feldt	1651.200	2.000	825.600
	Lower-bound	1651.200	1.000	1651.200
Error(textcond)	Sphericity Assumed	2379.467	18	132.193
	Greenhouse-Geisser	2379.467	17.555	135.541
	Huynh-Feldt	2379.467	18.000	132.193
	Lower-bound	2379.467	9.000	264.385

**Tests of Within-Subjects Effects**

Measure: MEASURE\_1

Source		F	Sig.
textcond	Sphericity Assumed	6.245	.009
	Greenhouse-Geisser	6.245	.009
	Huynh-Feldt	6.245	.009
	Lower-bound	6.245	.034
Error(textcond)	Sphericity Assumed		
	Greenhouse-Geisser		
	Huynh-Feldt		
	Lower-bound		

(Continued)

**Tests of Within-Subjects Contrasts**

Measure: MEASURE\_1

Source	textcond	Type III Sum of Squares	df	Mean Square	F	Sig.
textcond	Linear	1411.200	1	1411.200	12.168	.007
	Quadratic	240.000	1	240.000	1.617	.235
Error(textcond)	Linear	1043.800	9	115.978		
	Quadratic	1335.667	9	115.978		

**Tests of Between-Subjects Effects**

Measure: MEASURE\_1

Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	176026.800	1	176026.800	2009.099	.000
Error	788.533	9	87.615		

**Estimated Marginal Means****textcond****Estimates**

Measure: MEASURE\_1

textcond	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	66.200	4.139	56.838	75.562
2	80.600	3.553	72.562	88.638
3	83.000	2.333	77.722	88.278

### Pairwise Comparisons

Measure: MEASURE\_1

(I) textcond	(J) textcond	Mean Difference (I - J)	Std. Error	Sig. <sup>a</sup>	95% Confidence Interval for Difference <sup>a</sup>	
					Lower Bound	Upper Bound
1	2	- 14.400*	5.512	.028	-26.869	- 1.931
	3	- 16.800*	4.816	.007	- 27.695	- 5.905
2	1	14.400*	5.512	.028	1.931	26.869
	3	- 2.400	5.073	.647	-13.876	9.076
3	1	16.800*	4.816	.007	5.905	27.695
	2	2.400	5.073	.647	- 9.076	13.876

Based on estimated marginal means

a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

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

### Multivariate Tests

	Value	<i>F</i>	Hypothesis <i>df</i>	Error <i>df</i>	Sig.
Pillai's trace	.591	5.784 <sup>a</sup>	2.000	8.000	.028
Wilks' lambda	.409	5.784 <sup>a</sup>	2.000	8.000	<b>.028</b>
Hotelling's trace	1.446	5.784 <sup>a</sup>	2.000	8.000	.028
Roy's largest root	1.446	5.784 <sup>a</sup>	2.000	8.000	.028

Each *F* tests the multivariate effect of textcond. These tests are based on the linearly independent pairwise comparisons among the estimated marginal means.

The output in Figure A.5 is more complex for the repeated measures test than it is for the other tests we have seen. However, the output still contains the information needed to determine whether the tests are significant. To evaluate the main effect of the text condition, look for the “Tests of Within-Subjects Effects” box. The first row of the last box in the Tests of Within-Subjects Effects section shows the  $F$  and  $p$  values (circled in green). For Example 5, the  $F = 6.245$  and  $p = .009$ . Thus, we can reject the null hypothesis that there is no difference in condition means because the main effect of text condition is significant. The post hoc tests are shown in the box of the output labeled “Pairwise Comparisons.” The conditions are indicated by code value with  $p$  values listed in the Sig. column. As in Example 4, the post hoc tests indicate that learning for the paper text condition (coded as “1”) is lower than both of the computer conditions (coded as “2” and “3”), but the two computer conditions do not significantly differ from each other.

You also see a box in the output in Figure A.5 for Mauchly’s Test of Sphericity (with the  $p$  value circled in blue). Sphericity is an assumption of the repeated measures test. The assumption is that pairs of scores in the population have similar variability. (See Chapter 9 for more discussion of sphericity.) If the sphericity test is significant in the repeated measures output, the  $F$  statistic needs to be adjusted in order to retain accuracy of the test. Thus, the “Tests of Within-Subjects Effects” box contains a few different corrections below the “Sphericity Assumed” row. The sphericity assumed values are used if the sphericity test is not significant. However, if the sphericity test is significant, a correction is used because violations of this assumption can increase the chance of a Type I error (Keppel & Wickens, 2004). A common correction used in psychological research is the Greenhouse-Geisser correction. A full discussion of the correction techniques is provided in Howell’s (2009) statistics text.

### Example 6: Factorial ANOVA

The output in Figure A.6 is similar to that for Example 4; however, three tests of interest appear in the “Tests of Between-Subjects Effects” box (see green circled portion). The two main effects are indicated in the rows with the variable labels (“Study” and “Test” for Example 6). The main effect of study format was significant,  $F(1, 36) = 8.91, p = .005$ ; however, the main effect of test format was not significant,  $F(1, 36) = 1.32, p = .257$ . The means in the Descriptive Statistics box (see red circled portion) indicate that studied pictures ( $M = 80.45$ ) were better remembered than studied words ( $M = 71.50$ ), regardless of test format. This is a common finding in memory studies (Paivio, 2007). However, the interaction between study format and test format was also significant,  $F(1, 36) = 31.96, p < .001$ . Note that the  $p$  value in the output for the interaction is listed as .000. This value represents a value smaller than .0005 that has been rounded to three significant digits. In fact,  $p$  can never equal 0. The convention used in reporting such values is to indicate that the  $p$  was less than .001. Newer versions of SPSS list  $p$  values smaller than .0005 as  $p < .001$ .

**FIGURE A.6** ■ Output From Output Window for the Factorial ANOVA for Example 6**Univariate Analysis of Variance****Between-Subjects Factors**

		Value Label	N
Study format	1.00	Picture	20
	2.00	Word	20
Test format	1.00	Picture	20
	2.00	Word	20

**Descriptive Statistics**

Dependent Variable: Recognition Score

Study Format	Test Format	Mean	Std. Deviation	N
Picture	Picture	87.2000	7.29992	10
	Word	73.7000	13.08986	10
	Total	80.4500	12.42440	20
Word	Picture	61.3000	8.52513	10
	Word	81.7000	7.88881	10
	Total	71.5000	13.16894	20
Total	Picture	74.2500	15.36871	20
	Word	77.7000	11.29089	20
	Total	75.9750	13.42498	40

**Tests of Between-Subjects Effects**

Dependent Variable: Recognition Score

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected model	3793.075 <sup>a</sup>	3	1264.358	14.066	.000
Intercept	230888.025	1	230888.025	2568.673	.000
Study	801.025	1	801.025	8.912	.005
Test	119.025	1	119.025	1.324	.257
Study *test	2873.025	1	2873.025	31.963	.000
Error	3235.900	36	89.886		
Total	327917.000	40			
Corrected Total	7028.975	39			

a.  $r^2 = .540$  (Adjusted  $r^2 = .501$ )



## Example 7: Chi-Square Test

**FIGURE A.7** ■ Output From Output Window for the Chi-Square Test for Example 7

### Crosstabs

#### Case Processing Summary

	Cases					
	Valid		Missing		Total	
	<i>N</i>	Percentage	<i>N</i>	Percentage	<i>N</i>	Percentage
Age* Response	20	100.0%	0	.0%	20	100.0%

#### Age\* Response Crosstabulation

##### Count

		Response		
		Yes	No	Total
Age	Older	8	2	10
	Younger	4	6	10
Total		12	8	20

#### Chi-square Tests

	Value	<i>df</i>	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson chi-square	3.333 <sup>a</sup>	1	.068		
Continuity correction <sup>b</sup>	1.875	1	.171		
Likelihood ratio	3.452	1	.063		
Fisher's exact test				.170	.085
<i>N</i> of valid cases	20				

The output from the chi-square is shown in Figure A.7. The second box in the output shows the cross-tabulation with the number of participants in each condition (Younger/yes, etc.). The values for the chi-square test are shown in the box labeled “Chi-square tests.” The Pearson chi-square value of 3.333 is shown in the first row with its  $p$  value of .068 (see the green circled portion). For this example, the relationship is not significant because the  $p$  value is greater than the alpha level of .05. In other words, the type of response (“yes” or “no”) is not significantly related to the age group of the participants.

### Example 8: Pearson $r$ Correlation Test

**FIGURE A.8** ■ Output From Output Window for the Pearson  $r$  Test for Example 8

Correlations			
		Age	Days
Age	Pearson correlation	1	-.900**
	Sig. (2-tailed)		.00
	<i>N</i>	20	20
Days	Pearson correlation	-.900**	1
	Sig. (2-tailed)	.000	
	<i>N</i>	20	20

In Figure A.8, the Correlations box in the output indicates the Pearson  $r$  value (including the direction of relationship) in the first row and the  $p$  value in the second row (see green circled portion). For Example 8, the variables are significantly related (negatively) with  $r = -.90$  and  $p < .001$ . In other words, in this study, as age increased, the number of days it took participants to remember to mail the card decreased.

### Example 9: Simple Linear Regression

Figure A.9 shows the output from the linear regression. The last box contains the output information needed to determine the regression equation and whether age significantly predicts number of days (see blue circled portion). The  $t$  test in the bottom row provides the significance test (see green circled portion). In this case, the test is significant,  $t(18) = -8.78$ ,  $p < .001$ .

**FIGURE A.9** ■ Output From Output Window for the Simple Linear Regression Test for Example 9

### Regression

#### Variables Entered/Removed<sup>a</sup>

Model	Variables Entered	Variables Removed	Method
1	age <sup>b</sup>	.	Enter

a. Dependent Variable: days

b. All requested variables entered

#### Model Summary

Model	<i>r</i>	<i>r</i> <sup>2</sup>	Adjusted <i>r</i> <sup>2</sup>	Std. Error of the Estimate
1	.900 <sup>a</sup>	.811	.800	1.59066

a. Predictors: (Constant), age

#### ANOVA<sup>a</sup>

Model		Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	Sig.
1	Regression	195.006	1	195.006	77.071	.000 <sup>b</sup>
	Residual	45.544	18	2.530		
	Total	240.550	19			

a. Dependent Variable: days

b. Predictors: (Constant), age

#### Coefficients<sup>a</sup>

Model		Unstandardized Coefficients		Standardized Coefficients		Sig.
		<i>B</i>	Std. Error	Beta	<i>t</i>	
1	(Constant)	9.209	.707		13.035	.000
	Age	-.113	.013	-.900	-8.779	.000

a. Dependent Variable: days