y through my thoughts on doing
nd pictures have helped to hold
gress. There is a delicate balance
ice that varies from one reader to
anced for you.
ansformed you into a full-blown
s given you enough information
eriments on your own. You will
fun than reading about doing

APPENDIX A

How to Do Basic Statistics

If you are a calculatophobic (see Chapter 3), this appendix is written for you. You should be able to use this simplified "cookbook" version of statistics if you have learned only basic algebra. This book is certainly not a statistics book, and the minor concession I am making in this appendix does not contradict that statement. Some teachers and students who use this book sometimes feel the need for a brief description of basic descriptive statistics and inferential statistical tests. Here I tell you how to do a few of these tests. However, I am not telling you why you are doing what you are doing, and generally I will tell you only a little bit about the conditions for choosing what to do.

It has been my observation that writing words about numbers usually confuses things. Instead, I have attempted to show you what to do with the numbers by means of worked examples. If you arrange the numbers from your data the way the numbers are arranged in the example and follow the same steps, you should have few problems.

I first mention some characteristics of numbers. Then I give you a short glossary of statistical symbols. Finally, I provide you with worked examples of each statistical operation. At the end of each worked example I indicate how you would report this result within the text of a manuscript.

■ Characteristics of Numbers

Numbers can be used in a variety of ways. Some ways convey a lot of information ("it is 28 miles to the fair") and some only a little ("the first baseman is number 28"). Some statements and statistical operations are possible with some numbers ("the theater, which is 14 miles away, is half as far as the fair"). These same statements are ludicrous with others ("the second baseman is number 14; for this reason he is only half the first baseman"). So, before you can do a statistical operation on numbers, you must determine whether the operation makes sense for the type of numbers you are using.

NOMINAL SCALE

Numbers that are simply used to name something are said to be on a **nominal scale.** Nominal scale data have no quantitative properties. The only legitimate

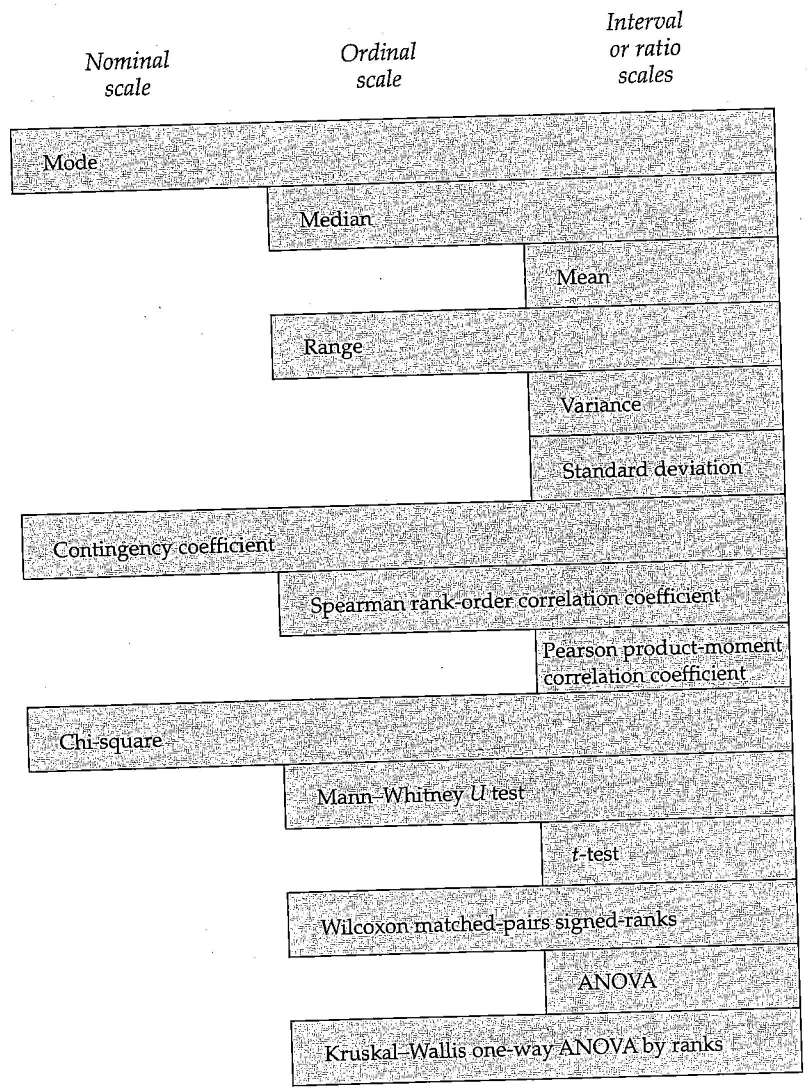

FIGURE A-1 Statistical operations and the number scales they require

statistical operation you can do with nominal data is to count the number of instances that each number occurs: How many players are there with the number 28?

ORDINAL SCALE

Numbers that can be ordered, or ranked, are said to lie on an **ordinal scale.** The race car driver who was champion the previous year is allowed to put the number 1 on his or her car. The driver who was second in the point standings is number 2, and so on. We know from these ordinal scale numbers that driver 1 performed better than driver 2, but we do not know by how much.

Interval or ratio scales
Mean
Variance
Standard deviation
relation coefficient
'earson product-moment orrelation coefficient
t-test
signed-ranks
ANOVA
ANOVA by ranks

ales they require

is to count the number of
players are there with the

to lie on an **ordinal scale.**
ous year is allowed to put
second in the point stand-
ordinal scale numbers that
o not know by how much.

Drivers 1 and 2 might have been 500 points apart, whereas drivers 2 and 3 might have been only 2 points apart.

INTERVAL SCALE

If the intervals between numbers are meaningful, the numbers lie on an **interval scale.** For example, temperature measured on a Fahrenheit scale is on an interval scale. It is 10 °F between 50 °F and 60 °F. It is also 10 °F between 60 °F and 70 °F.

RATIO SCALE

If you can make a ratio out of two numbers and if this ratio is meaningful, you have a **ratio scale.** Thus, although you cannot say that a temperature of 20 °F is twice as hot as 10 °F, you can say that 20 miles is twice as far as 10 miles. The big difference between an interval and a ratio scale is that the latter has an absolute zero point. On the Fahrenheit scale, "zero degrees" has no particular meaning other than the fact it is 32 arbitrary degrees below the freezing point of water. For quantities such as distance, weight, and volume, zero units is a meaningful concept.

One question you must answer before you perform a statistical operation is "What number scale am I dealing with?" Figure A-1 lists the operations we discuss in this appendix and the scales they require.

Operations that can be carried out on numbers from lower order scales such as nominal can also be used with numbers from higher order scales such as ratio. For this reason, the shaded area in the figure indicates that all the operations can legitimately be performed with interval or ratio data, but only three of them can be used with nominal data.

Symbols Used in Statistical Formulas

X = a datum or score
N = the total number of scores
Σ = sum, or add, the scores
X^2 = square X (multiply it by itself)
X^3 = cube X (multiply it by itself twice)
$\sqrt{X}$ = square root of X (What number multiplied by itself equals X?)
$|x|$ = absolute value of X (the number disregarding its sign)

Descriptive Statistics

MEASURES OF CENTRAL TENDENCY

Mode

The mode is the most frequently occurring score. Count the number of times each score occurs and pick the score with the most occurrences. The mode in the example in the next section is 2 because this number occurs twice.

Median

The median is the middle score. The scores should first be ordered by size. For an odd number of scores, the median is the middle one. For an even number of scores, the median lies halfway between the two middle scores. In the following example the median is 2.5, because the middle two scores are 2 and 3.

Mean

$$\text{Mean} = M = \overline{X} = \frac{\Sigma X}{N}$$

Example

	X
	1
	2
	2
	3
	4
	5
$\Sigma X =$	17
$N =$	6

$$\overline{X} = \frac{17}{6} = 2.8$$

Reporting in Text. $M = 2.8$

MEASURES OF DISPERSION

Range

The range is the largest score minus the smallest score. In the previous example:

Range = 5 − 1 = 4

Variance

$$\text{Variance} = S^2 = \frac{\Sigma(X - \overline{X})^2}{N}$$

Example

	X	$\overline{X}$*	$X - \overline{X}$	$(X - \overline{X})^2$
	1	3	−2	4
6 scores	2	3	−1	1
so	3	3	0	0
$N = 6$	3	3	0	0
	4	3	1	1
	5	3	2	4
	$\Sigma X = 18$			$\Sigma(X - \overline{X})^2 = 10$

*Mean = $\overline{X} = \frac{18}{6} = 3$.

$$\frac{\Sigma(X - \overline{X})^2}{N} = \frac{10}{6} = 1.67$$

Standard Deviation

$$\text{Standard deviation} = SD = \sigma = \sqrt{S^2} = \sqrt{\frac{\Sigma(X - \overline{X})^2}{N}}$$

In the previous example:

$$SD = \sqrt{1.67} = 1.29$$

Reporting in Text. $SD = 1.29$

MEASURES OF ASSOCIATION

Contingency Coefficient

The **contingency coefficient** (C) is a measure of the strength of association between two sets of numbers when nominal scale data are being considered. A chi-square (χ^2) test must first be done (see p. 315). Suppose that a chi-square test has been conducted on a two-variable experiment, and you wish to know the strength of association between these two nominal-scale variables. Also suppose that χ^2 was found to be 15, with a total number of observations of $N = 100$. Then the contingency coefficient is:

$$C = \sqrt{\frac{\chi^2}{N + \chi^2}} = \sqrt{\frac{15}{100 + 15}} = \sqrt{.130} = .36$$

No further testing for the statistical significance of the association is necessary because the chi-square test has already been computed to test for significance.

Reporting in Text. $C\ (N = 100) = .36$

Spearman Rank-Order Correlation Coefficient

A **Spearman rank-order correlation coefficient** (*rho*) is used to measure the strength of association between two ordinal scale variables. In this case, two scores, or ranks, are obtained for each participant, and the difference d is determined.

Example

Participant	*Rank on first measure*	*Rank on second measure*	d	d^2	
Bill	4	4	0	0	
Jane	1	2	−1	1	
Bob	5	5	0	0	
Pete	2	3	−1	1	$N = 5$
Mary	3	1	+2	4	
				$\Sigma d^2 = 6$	

$$rho = 1 - \frac{6\Sigma d^2}{N^3 - N} = 1 - \frac{6(6)}{125 - 5} = 1 - \frac{36}{120}$$

$$= 1 - .3 = .7$$

To determine whether the obtained *rho* is likely to have occurred because of chance variation rather than an actual association, we must consult the table of critical values for *rho* in Appendix B (Table B-1). We see that with *N* of 5, *rho* must equal 1 to be significant. It does not. We can also see from the table that the larger the number of participants, the better our chances of finding a statistically significant effect, given that there is an association present.

Reporting in Text. $\rho\ (N = 5) = .70, p > .05$

Pearson Product-Moment Correlation Coefficient

A **Pearson product-moment correlation coefficient** (r) can be used to measure the strength of association between two interval or ratio scale variables. In the following example, X represents the score on one variable, and Y the score on a second variable.

Example

Participant	X	X^2	Y	Y^2	XY	
Tom	9	81	8	64	72	
Sue	4	16	4	16	16	
Jill	4	16	6	36	24	
Dave	2	4	4	16	8	N = 8
Ken	1	1	3	9	3	
Jo	3	9	2	4	6	
Juan	7	49	8	64	56	
Al	5	25	5	25	25	
	$\Sigma X = 35$	$\Sigma X^2 = 201$	$\Sigma Y = 40$	$\Sigma Y^2 = 234$	$\Sigma XY = 210$	

$$r = \frac{N\Sigma XY - \Sigma X \Sigma Y}{\sqrt{N\Sigma X^2 - (\Sigma X)^2}\ \sqrt{N\Sigma Y^2 - (\Sigma Y)^2}} = \frac{8(210) - (35)(40)}{\sqrt{8(201) - 35^2}\ \sqrt{8(234) - 40^2}}$$

$$= \frac{1680 - 1400}{\sqrt{1608 - 1225}\ \sqrt{1872 - 1600}} = \frac{280}{\sqrt{383}\ \sqrt{272}} = \frac{280}{(19.57)(16.49)}$$

$$= \frac{280}{322.7} = .868$$

To test whether an r of this size is statistically significant with eight pairs of scores, refer to Table B-2 in Appendix B, listing critical values of r. To use this table, you must determine a quantity called the *degrees of freedom* (*df*). For this test the degrees of freedom is $N - 2$. So in the example $df = 6$. Because

to have occurred because
ion, we must consult the
3-1). We see that with N of
can also see from the table
r our chances of finding a
association present.

it

nt (r) can be used to mea-
val or ratio scale variables.
on one variable, and Y the

Y^2	XY	
64	72	
16	16	
36	24	
16	8	$N = 8$
9	3	
4	6	
64	56	
25	25	
= 234	ΣXY = 210	

$$\frac{8(210) - (35)(40)}{.01) - 35^2 \sqrt{8(234) - 40^2}}$$

$$= \frac{280}{(19.57)(16.49)}$$

72

significant with eight pairs
g critical values of r. To use
d the *degrees of freedom* (df).
the example $df = 6$. Because

r of .868 exceeds the listed value of .834, it is statistically significant at the $p < .01$ level. That is, we would expect this strength of association to occur in a sample less than 1 time in 100 because of chance selection from a single population.

Reporting in Text. $r\,(6) = .87, p < .01$

Inferential Statistical Tests

CHI-SQUARE

The **chi-square (χ^2) test** is used to determine whether the observed frequency of occurrence of scores is statistically different from the expected frequency.

Example

	Number of participants predicting heads after a string of tails	*Number of participants predicting tails after a string of tails*
Observed	60	40
Expected	50	50
$O - E$	+10	−10
$(O - E)^2$	100	100
$\frac{(O - E)^2}{E}$	2	2

$$\chi^2 = \Sigma\frac{(O - E)^2}{E} = 2 + 2 = 4$$

The expected frequency can be the frequency based either on a set of previous observations or on a theoretical prediction. Usually, the theoretical prediction is that the observed frequency will be that expected by chance. For instance, in the example the expectation is that the participants' predictions will show no bias (no gambler's fallacy)—half the time they will predict heads and half the time, tails.

The final step in doing an inferential statistical test is to compare the final result of your computation with a table of critical values. You will find a table for chi-square values in Appendix B (Table B-3). To find the appropriate number in the table, you must first determine the number of degrees of freedom as follows:

> df = The number of $O - E$ being considered, minus 1, which in this case equals $2 - 1 = 1$

In the table we find that with $df = 1$, χ^2 must exceed 3.84 to be significant at the $p < .05$ level of significance. Thus, the data in our example are statistically different from chance at the .05 level. If we had been testing at the $p < .01$

level, $\chi^2 = 4$ would not have exceeded 6.64, and the test would have failed to reach significance.

Reporting in Text. χ^2 (1, N = 100) = 4.00, $p < .05$

t-TEST FOR UNCORRELATED MEASURES

There are two forms of the ***t*-test,** one for uncorrelated measures and the other for correlated measures. The *t*-test for uncorrelated measures is used to determine the probability that an observed difference between two independent groups of participants occurred by chance. The underlying distributions are assumed to be normal.

Example

Group 1			
X_1	$\overline{X}_1$	$X_1 - \overline{X}_1$	$(X_1 - \overline{X}_1)^2$
9	7	2	4
8	7	1	1
7	7	0	0
7	7	0	0
4	7	−3	9
$\Sigma X_1 = 35$			$\Sigma(X_1 - \overline{X}_1)^2 = 14$

$N_1 = 5$

$$M_1 = \overline{X}_1 = \frac{\Sigma X_1}{N_1} = \frac{35}{5} = 7$$

$$\sigma_1 = \sqrt{\frac{\Sigma(X_1 - \overline{X}_1)^2}{N_1}} = \sqrt{\frac{14}{5}} = \sqrt{2.8} = 1.67$$

Group 2			
X_2	$\overline{X}_2$	$X_2 - \overline{X}_2$	$(X_2 - \overline{X}_2)^2$
5	3	2	4
4	3	1	1
3	3	0	0
2	3	−1	1
1	3	−2	4
$\Sigma X_2 = 15$			$\Sigma(X_2 - \overline{X}_2)^2 = 10$

$N_2 = 5$

he test would have failed to

;

ated measures and the other
ed measures is used to deter-
e between two independent
underlying distributions are

$(X_1 - \bar{X}_1)^2$
4
1
0
0
9
$- \bar{X}_1)^2 = 14$

$(X_2 - \bar{X}_2)^2$
4
1
0
1
4
$- \bar{X}_2)^2 = 10$

$$M_2 = \bar{X}_2 = \frac{\Sigma X_2}{N_2} = \frac{15}{5} = 3$$

$$\sigma_2 = \sqrt{\frac{\Sigma(X_2 - \bar{X}_2)^2}{N_2}} = \sqrt{\frac{10}{5}} = \sqrt{2} = 1.41$$

$$t = \frac{M_1 - M_2}{\sqrt{\left(\frac{\sigma_1}{\sqrt{N_1 - 1}}\right)^2 + \left(\frac{\sigma_2}{\sqrt{N_2 - 1}}\right)^2}} = \frac{7 - 3}{\sqrt{\left(\frac{1.67}{\sqrt{5 - 1}}\right)^2 + \left(\frac{1.41}{\sqrt{5 - 1}}\right)^2}}$$

$$= \frac{4}{\sqrt{\left(\frac{1.67}{2}\right)^2 + \left(\frac{1.41}{2}\right)^2}} = \frac{4}{\sqrt{.697 + .497}} = \frac{4}{\sqrt{1.194}} = \frac{4}{1.09} = 3.67$$

The degrees of freedom for an uncorrelated t-test is:

$$df = N_1 + N_2 - 2$$
$$= 5 + 5 - 2 = 8$$

Table B-4 in Appendix B indicates that, with 8 df, t must exceed 3.355 for the difference to be significant at $p < .01$. Thus, our value of 3.67 is significant at that level.

Reporting in Text. $t\ (8) = 3.67, p < .01$

t-TEST FOR CORRELATED MEASURES

The t-test for correlated measures is used to determine the probability that an observed difference (D) between two conditions for the same or matched participants occurred by chance.

$$t = \frac{\bar{X}_D}{\frac{\sigma_D}{\sqrt{N - 1}}}$$

Example

Participant	*Condition 1*	*Condition 2*	*Difference (D)*	$\bar{X}_D$*	$X_D - \bar{X}_D$	$(X_D - \bar{X}_D)^2$
1	9	6	3	3	0	0
2	8	5	3	3	0	0
3	7	5	2	3	−1	1
4	8	4	4	3	1	1
5	8	5	3	3	0	0
$N = 5$	$\Sigma X_1 = 40$	$\Sigma X_2 = 25$	$\Sigma D = 15$			$\Sigma(X_D - \bar{X}_D)^2 = 2$

$*M_D = \bar{X}_D = \frac{\Sigma D}{N} = \frac{15}{5} = 3$

$$\sigma_D = \sqrt{\frac{\Sigma(X_D - \overline{X}_D)^2}{N}} = \sqrt{\frac{2}{5}} = \sqrt{.4} = .632$$

$$t = \frac{\overline{X}_D}{\frac{\sigma_D}{\sqrt{N-1}}} = \frac{3}{\frac{.632}{\sqrt{5-1}}} = \frac{3}{\frac{.632}{2}} = \frac{3}{.316} = 9.49$$

The degrees of freedom for correlated measures is:

$df = N - 1 = 5 - 1 = 4$

Table B-4 is used for either form of the t-test. In this example, t must exceed 4.604 to be significant at the $p < .01$ level. It does, so it is.

Reporting in Text. $t\ (4) = 9.49, p < .01$

MANN–WHITNEY *U* TEST

The **Mann–Whitney *U* test** is used under the same general conditions as an uncorrelated t-test but only when the assumptions of normal distributions or an interval scale cannot be met.

$$U = N_1N_2 + \frac{N_1(N_1+1)}{2} - R_1$$

or

$$U = N_1N_2 + \frac{N_2(N_2+1)}{2} - R_2$$

whichever is smaller

where

N_1 = the number of participants in the smaller group
N_2 = the number of participants in the larger group
R_1 = the sum of the ranks for the smaller group
R_2 = the sum of the ranks for the larger group

Example

	Group 1			*Group 2*	
	X_1	*Rank*		X_2	*Rank*
$N_1 = 10$	1	1	$N_2 = 10$	2	2
	3	3.5		4	5
	3	3.5		7	8
	5	6		8	9.5
	6	7		10	13.5
	8	9.5		13	16
	9	11.5		15	17
	9	11.5		16	18
	10	13.5		17	19
	12	15		18	20
		$R_1 = 82$			$R_2 = 128$

NOTE: The rankings were determined by ordering all scores regardless of which group they came from. Where there were ties in rankings, an average was used.

= 9.49

ıres is:

est. In this example, t must
.. It does, so it is.

ıme general conditions as an
ns of normal distributions or

er is smaller

ller group
er group
oup
up

	Group 2	
	X_2	Rank
	2	2
	4	5
	7	8
	8	9.5
10	10	13.5
	13	16
	15	17
	16	18
	17	19
	18	20
		R_2 = 128

s regardless of which group they
was used.

$$U = N_1N_2 + \frac{N_1(N_1+1)}{2} - R_1 = (10)(10) + \frac{10(10+1)}{2} - 82$$

$$= 100 + \frac{110}{2} - 82 = 73$$

or

$$U = (10)(10) + \frac{10(10+1)}{2} - 128 = 27$$

Because 27 is smaller, $U = 27$.

Two tables for determining the critical values of U can be found in Appendix B (Tables B-5 and B-6). If we wished to test for significance at the $p < .05$ level, we would use Table B-5. The value for U when $N_1 = 10$ and $N_2 = 10$ is 23. To be significant, our value must be equal to or *smaller than* this critical value. Because 27 is not, it is not statistically significant at this level.

Note that the Mann–Whitney U test is different from the other tests because in order to be significant the value must be smaller rather than larger than the value in the table. To find a table for values of N_1 smaller than 7, you will have to use a more advanced text than this one. For values of N_2 larger than 20, U must be converted to a z score using this formula:

$$z = \frac{U - \frac{N_1N_2}{2}}{\sqrt{\frac{(N_1)(N_2)(N_1+N_2+1)}{12}}}$$

The z score can then be compared with the critical values listed in Table B-7 in Appendix B.

Reporting in Text. $U\ (N_1 = 10, N_2 = 10) = 27, p < .05$

WILCOXON MATCHED-PAIRS SIGNED-RANKS TEST

The **Wilcoxon matched-pairs signed-ranks test** is used to determine the probability that an observed difference (D) between two conditions for the same or matched participants occurred by chance. It differs from the t-test for correlated measures in that it can be used with ordinal data and the underlying distributions need not be normal.

$$\left.\begin{array}{l} T = \Sigma R_+ \\ T = |\Sigma R_-| \end{array}\right\} \text{whichever is smaller}$$

where

R_+ is a rank having a positive difference
R_- is a rank having a negative difference

Example

Pair	Condition 1	Condition 2	Difference (D)	Rank of D ignoring sign	Rank having a positive D	Rank having a negative D
1	54	50	4	3	3	
2	47	32	15	9	9	
3	39	33	6	4	4	
4	42	45	−3	2.5		−2.5
5	51	38	13	7	7	
6	46	39	7	5	5	
7	42	44	−2	1		−1
8	54	46	8	6	6	
9	42	39	3	2.5	2.5	
10	47	33	14	8	8	
					$\Sigma R_+ = 45.5$	$\|\Sigma R_-\| = 3.5$

The smaller of 45.5 and 3.5 is 3.5; thus:

$$T = |\Sigma R_-| = 3.5$$

To test for statistical significance, look at Table B-8 in Appendix B. To reach significance, T must be equal to or smaller than the number listed. In the example there are 10 pairs of scores, so $n = 10$, and assuming that we did not predict the direction of the difference between conditions, a two-tailed test is appropriate. We see, then, that 3.5 is smaller than 5 but not 3, so $p < .02$.

Reporting in Text. $T\ (n = 10) = 3.50, p < .02$

ANALYSIS OF VARIANCE

Analysis of variance (ANOVA) can be used for interval or ratio data when the underlying distributions are approximately normal. ANOVA tests are available for either within-subject (repeated measures) or between-subjects (separate groups) designs and for designs with multiple independent variables. In this appendix, however, we limit our consideration to a between-subjects design with one independent variable. In the following example the independent variable has three levels. However, the formulas given can also be used for designs having more than three groups.

Although the calculations for ANOVA appear to be complicated, the rationale behind the test is really quite simple. Suppose that you conduct an experiment in which you collect data from three groups. The experimental question is whether the three samples come from the same population and differ only by chance variation or whether the samples come from different populations and differ owing to the independent variable, as well as to chance variation. ANOVA allows you to partition the variance found in the distribution containing all the scores you sampled. Part of the variance in this distribution is due to differences between groups, including variance due to the independent variable. A second part is due to chance variation among participants within groups.

Rank having a positive D	Rank having a negative D
3	
9	
4	
	−2.5
7	
5	
	−1
6	
2.5	
8	
$\Sigma R_+ = 45.5$	$\lvert \Sigma R_- \rvert = 3.5$

in Appendix B. To reach
he number listed. In the
assuming that we did not
itions, a two-tailed test is
but not 3, so $p < .02$.

val or ratio data when the
1. ANOVA tests are avail-
r between-subjects (sepa-
independent variables. In
on to a between-subjects
owing example the inde-
rmulas given can also be

ar to be complicated, the
pose that you conduct an
groups. The experimental
the same population and
nples come from different
nt variable, as well as to
the variance found in the
Part of the variance in this
including variance due to
chance variation among

The final number calculated when doing ANOVA is called an *F value.* It is a ratio of the variance between groups to the variance within groups. If the groups sampled come from the same population and the independent variable has no effect, we would expect the ratio to be close to 1. That is, the between-group variance should be about the same size as the within-group variance. However, if the independent variable has an effect and the groups come from different populations, we would expect the between-group variance to be larger than the within-group variance. The F value would then be greater than 1. As the value of F gets larger, we would become increasingly confident that the differences among groups were due to the effects of the independent variable rather than to chance variation.

In the following example, we first calculate a quantity called the *total sum of squares* (SS_{TOT}) followed by a *sum of squares between groups* (SS_{bg}) and *within groups* (SS_{wg}). We then divide SS_{bg} and SS_{wg} by their appropriate degrees of freedom to get the mean squares between groups (MS_{bg}) and within groups (MS_{wg}). MS_{bg} is then divided by MS_{wg} to find the value of F.

You should be able to follow the example, but if you get into trouble, the following definitions might help:

T = the total sum of all scores for all groups
T_j = the total sum of scores in group j
N = the number of scores in all groups
n_j = the number of scores in group j
$\sum_{j=1}^{k}$ means to sum for all groups from 1 to k

k = the number of groups

Example

Group 1		*Group 2*		*Group 3*	
X_1	X_1^2	X_2	X_2^2	X_3	X_3^2
3	9	9	81	10	100
5	25	6	36	8	64
4	16	5	25	11	121
3	9	8	64	10	100
1	1	7	49	9	81
2	4	7	49	10	100
5	25	6	36	11	121
2	4	4	16	12	144
3	9	8	64	10	100
1	1	7	49	9	81
$T_1 = 29$	103	$T_2 = 67$	469	$T_3 = 100$	1012
$n_1 = 10$		$n_2 = 10$		$n_3 = 10$	

$N = 10 + 10 + 10 = 30$
$T = 29 + 67 + 100 = 196$
$k = 3$

$$SS_{TOT} = \Sigma X^2 - \frac{T^2}{N} = (103 + 469 + 1012) - \frac{(196)^2}{30}$$

$$= 1584 - \frac{38416}{30} = 1584 - 1281 = 303$$

$$SS_{bg} = \sum_{j=1}^{k} \frac{T_j^2}{n_j} - \frac{T^2}{N} = \frac{29^2}{10} + \frac{67^2}{10} + \frac{100^2}{10} - \frac{(196)^2}{30}$$

$$= \frac{841}{10} + \frac{4489}{10} + \frac{10000}{10} - 1281$$

$$= 84.1 + 448.9 + 1000 - 1281 = 1533 - 1281 = 252$$

$$SS_{wg} = SS_{TOT} - SS_{bg} = 303 - 252 = 51$$

$$df_{bg} = k - 1 = 3 - 1 = 2$$

$$df_{wg} = N - k = 30 - 3 = 27$$

$$MS_{bg} = \frac{SS_{bg}}{df_{bg}} = \frac{252}{2} = 126$$

$$MS_{wg} = \frac{SS_{wg}}{df_{wg}} = \frac{51}{27} = 1.89$$

$$F = \frac{MS_{bg}}{MS_{wg}} = \frac{126}{1.89} = 66.7$$

We can now compare this number with the critical values for F listed in Table B-9 in Appendix B. With 2 *df* in the numerator and 27 *df* in the denominator, F must equal or exceed 3.38 to be significant at $p < .05$ and equal or exceed 5.57 to be significant at $p < .01$. Because 66.7 far exceeds these critical values, the difference between the groups is highly significant. Note that the test could reach statistical significance owing to a difference between any two groups. To determine which means are statistically different from one another, further tests would have to be conducted. These tests are beyond the scope of this book. They can be found in the recommended texts at the end of Chapter 12.

Reporting in Text. $F\ (2, 27) = 66.70, p < .05$

KRUSKAL–WALLIS ONE-WAY ANOVA BY RANKS

If the assumptions of an interval or ratio scale or normal distributions cannot be met, a **Kruskal–Wallis one-way ANOVA** can be used to test for differences between two or more independent groups. Only an ordinal scale is necessary.

In the following example:

K = the number of groups
n_j = the number of scores per group
N = the total number of scores
R_j = the sum of ranks for group j
t = the number of ties for each score

Example

Group 1		*Group 2*		*Group 3*	
X_1	*Rank*	X_2	*Rank*	X_3	*Rank*
8	15	2	2.5	6	11
4	5.5	5	8.5	5	8.5
7	13	2	2.5	4	5.5
5	8.5	3	4	5	8.5
7	13	1	1	7	13
	$R_1 = 55.0$		$R_2 = 18.5$		$R_3 = 46.5$

$K = 3$
$n_j = 5$
$N = 15$

Rank all the scores to get the ranks for each group.

Score	*Rank*	*Average for ties*	*t*
1	1	1	
2	2	2.5	2
2	3		
3	4	4	
4	5	5.5	2
4	6		
5	7	8.5	4
5	8		
5	9		
5	10		
6	11	11	
7	12	13	3
7	13		
7	14		
8	15	15	

Now place the ranks from this table next to the individual scores for each group in the previous table, and sum them to get R_1, R_2, and R_3.

$$H = \frac{12}{N(N+1)} \sum_{j=1}^{k} \frac{R_j^2}{n_j} - 3(N+1)$$

$$= \frac{12}{15(15+1)} \left[\frac{(55)^2}{5} + \frac{(18.5)^2}{5} + \frac{(46.5)^2}{5} \right] - 3(15+1)$$

$$= \frac{12}{15(16)} \left[\frac{3025}{5} + \frac{342.25}{5} + \frac{2162.25}{5} \right] - 3(16)$$

$$= \frac{12}{240} \left[\frac{5529.5}{5} \right] - 48$$

$$= .05(1105.9) - 48 = 55.295 - 48 = 7.295$$

The correction for ties is to divide H by $1 - \frac{\Sigma(t^3 - t)}{N^3 - N}$.

$$1 - \frac{(2^3 - 2) + (2^3 - 2) + (4^3 - 4) + (3^3 - 3)}{15^3 - 15}$$

$$1 - \frac{(8 - 2) + (8 - 2) + (64 - 4) + (27 - 3)}{3375 - 15}$$

$$1 - \frac{96}{3360} = 1 - .029 = .971$$

$$H = \frac{7.295}{.971} = 7.51$$

According to Table B-10 in Appendix B, for group sizes of 5, 5, and 5, the probability of having an H as large as 7.51 is less than .049. Thus, the difference between groups is statistically significant at the $p < .05$ level. Because this value is smaller than the 7.98 required for the $p < .01$ level, the difference is not significant at that level.

If the groups contain more than five participants, H is distributed like the chi-square. To determine the critical value in this case, refer to Table B-3 with $k - 1$ degrees of freedom.

Reporting in Text. $H\,(5, 5, 5) = 7.51, p < .05$

Conclusion

This appendix should allow you to compute some very basic statistical operations. However, if you go much beyond a basic course in experimentation, you will need to do at least three additional things. First, you will need to learn to use more complex tests for designs having multiple independent variables and mixtures of within-subject and between-subjects variables. Second, you will need to learn to use packaged computer programs to save time and effort. Third, and probably most important, you must go beyond a cookbook approach to statistics. As a researcher you should understand why you do what you do.

An understanding of the concepts underlying statistical operations not only allows you to choose the most powerful way to analyze your data but also allows you to design research so that the data can be effectively analyzed. Statistical consultants tell horror stories about inexperienced researchers who dump volumes of data on their desks and ask, "How do I analyze this?" In some cases the data defy analysis.

The point is that design and statistical analysis are integrally linked. If you plan to design your research, you should also understand the concepts underlying the statistical operations that should be used to analyze the outcome.