The SAGE Dictionary of Statistics

Duncan Cramer and Dennis Howitt

The SAGE Dictionary of Statistics

The SAGE Dictionary of Statistics

a practical resource for students
in the social sciences

Duncan Cramer and Dennis Howitt

SAGE Publications
London ● Thousand Oaks ● New Delhi

First published 2004

SAGE Publications Ltd
1 Oliver's Yard
55 City Road
London EC1Y 1SP

SAGE Publications Inc.
2455 Teller Road
Thousand Oaks, California 91320

SAGE Publications India Pvt Ltd
B-42, Panchsheel Enclave
Post Box 4109
New Delhi 110 017

British Library Cataloguing in Publication data

A catalogue record for this book is available from the British Library

ISBN 0 7619 4137 1
ISBN 0 7619 4138 X (pbk)

Library of Congress Control Number: 2003115348

Typeset by C&M Digitals (P) Ltd.
Printed in Great Britain by The Cromwell Press Ltd, Trowbridge, Wiltshire

Contents

To our mothers – it is not their fault that lexicography took its toll.

Preface

Writing a dictionary of statistics is not many people's idea of fun. And it wasn't ours. Can we say that we have changed our minds about this at all? No. Nevertheless, now the reading and writing is over and those heavy books have gone back to the library, we are glad that we wrote it. Otherwise we would have had to buy it. The dictionary provides a valuable resource for students – and anyone else with too little time on their hands to stack their shelves with scores of specialist statistics textbooks.

Writing a dictionary of statistics is one thing – writing a practical dictionary of statistics is another. The entries had to be useful, not merely accurate. Accuracy is not that useful on its own. One aspect of the practicality of this dictionary is in facilitating the learning of statistical techniques and concepts. The dictionary is not intended to stand alone as a textbook – there are plenty of those. We hope that it will be more important than that. Perhaps only the computer is more useful. Learning statistics is a complex business. Inevitably, students at some stage need to supplement their textbook. A trip to the library or the statistics lecturer's office is daunting. Getting a statistics dictionary from the shelf is the lesser evil. And just look at the statistics textbook next to it – you probably outgrew its usefulness when you finished the first year at university.

Few readers, not even ourselves, will ever use all of the entries in this dictionary. That would be a bit like stamp collecting. Nevertheless, all of the important things are here in a compact and accessible form for when they are needed. No doubt there are omissions but even *The Collected Works of Shakespeare* leaves out *Pygmalion*! Let us know of any. And we are not so clever that we will not have made mistakes. Let us know if you spot any of these too – modern publishing methods sometimes allow corrections without a major reprint.

Many of the key terms used to describe statistical concepts are included as entries elsewhere. Where we thought it useful we have suggested other entries that are related to the entry that might be of interest by listing them at the end of the entry under 'See' or 'See also'. In the main body of the entry itself we have not drawn attention to the terms that are covered elsewhere because we thought this could be too distracting to many readers. If you are unfamiliar with a term we suggest you look it up.

Many of the terms described will be found in introductory textbooks on statistics. We suggest that if you want further information on a particular concept you look it up in a textbook that is ready to hand. There are a large number of introductory statistics texts that adequately discuss these terms and we would not want you to seek out a particular text that we have selected that is not readily available to you. For the less common terms we have recommended one or more sources for additional reading. The authors and year of publication for these sources are given at the end of the entry and full details of the sources are provided at the end of the book. As we have discussed some of these terms in texts that we have written, we have sometimes recommended our own texts!

The key features of the dictionary are:

- Compact and detailed descriptions of key concepts.
- Basic mathematical concepts explained.
- Details of procedures for hand calculations if possible.
- Difficulty level matched to the nature of the entry: very fundamental concepts are the most simply explained; more advanced statistics are given a slightly more sophisticated treatment.
- Practical advice to help guide users through some of the difficulties of the application of statistics.
- Exceptionally wide coverage and varied range of concepts, issues and procedures – wider than any single textbook by far.
- Coverage of relevant research methods.
- Compatible with standard statistical packages.
- Extensive cross-referencing.
- Useful additional reading.

One good thing, we guess, is that since this statistics dictionary would be hard to distinguish from a two-author encyclopaedia of statistics, we will not need to write one ourselves.

Duncan Cramer
Dennis Howitt

Some Common Statistical Notation

Roman letter symbols or abbreviations:

a	constant
df	degrees of freedom
F	F test
log n	natural or Napierian logarithm
M	arithmetic mean
MS	mean square
n or N	number of cases in a sample
p	probability
r	Pearson's correlation coefficient
R	multiple correlation
SD	standard deviation
SS	sum of squares
t	t test

Greek letter symbols:

α (lower case alpha) Cronbach's alpha reliability, significance level or alpha error
β (lower case beta) regression coefficient, beta error
γ (lower case gamma)
σ (lower case delta)
η (lower case eta)
κ (lower case kappa)
λ (lower case lambda)
ρ (lower case rho)
τ (lower case tau)
φ (lower case phi)
χ (lower case chi)

Some common mathematical symbols:

$\sum$	sum of
∞	infinity
$=$	equal to
$<$	less than
$\leq$	less than or equal to
$>$	greater than
$\geq$	greater than or equal to
$\sqrt{}$	square root

A

a posteriori tests: see ***post hoc* tests**

a priori comparisons or **tests:** where there are three or more means that may be compared (e.g. analysis of variance with three groups), one strategy is to plan the analysis in advance of collecting the data (or examining them). So, in this context, a priori means before the data analysis. (Obviously this would only apply if the researcher was not the data collector, otherwise it is in advance of collecting the data.) This is important because the process of deciding what groups are to be compared should be on the basis of the hypotheses underlying the planning of the research. By definition, this implies that the researcher is generally disinterested in general or trivial aspects of the data which are not the researcher's primary focus. As a consequence, just a few of the possible comparisons are needed to be made as these contain the crucial information relative to the researcher's interests. Table A.1 involves a simple ANOVA design in which there are four conditions – two are drug treatments and there are two control conditions. There are two control conditions because in one case the placebo tablet is for drug A and in the other case the placebo tablet is for drug B.

An appropriate a priori comparison strategy in this case would be:

- Mean_a against Mean_b
- Mean_a against Mean_c
- Mean_b against Mean_d

Table A.1 *A simple ANOVA design*

Drug A	Drug B	Placebo control A	Placebo control B
Mean_a =	Mean_b =	Mean_c =	Mean_d =

Notice that this is fewer than the maximum number of comparisons that could be made (a total of six). This is because the researcher has ignored issues which perhaps are of little practical concern in terms of evaluating the effectiveness of the different drugs. For example, comparing placebo control A with placebo control B answers questions about the relative effectiveness of the placebo conditions but has no bearing on which drug is the most effective overall.

The a priori approach needs to be compared with perhaps the more typical alternative research scenario – *post hoc* comparisons. The latter involves an unplanned analysis of the data following their collection. While this may be a perfectly adequate process, it is nevertheless far less clearly linked with the established priorities of the research than a priori comparisons. In *post hoc* testing, there tends to be an exhaustive examination of all of the possible pairs of means – so in the example in Table A.1 all four means would be compared with each other in pairs. This gives a total of six different comparisons.

In a priori testing, it is not necessary to carry out the overall ANOVA since this merely tests whether there are differences across the various means. In these circumstances, failure of some means to differ from the others may produce non-significant findings due to conditions which are of little or no interest to the researcher. In a priori testing, the number of comparisons to be made has been limited to a small number of key comparisons. It is generally accepted that if there are relatively few a priori comparisons to be made, no adjustment is needed for the number of comparisons made. One rule of thumb is that if the comparisons are fewer in total than the degrees of freedom for the main effect minus one, it is perfectly appropriate to compare means without adjustment for the number of comparisons.

Contrasts are examined in a priori testing. This is a system of weighting the means in order to obtain the appropriate mean difference when comparing two means. One mean is weighted (multiplied by) +1 and the other is weighted −1. The other means are weighted 0. The consequence of this is that the two key means are responsible for the mean difference. The other means (those not of interest) become zero and are always in the centre of the distribution and hence cannot influence the mean difference.

There is an elegance and efficiency in the a priori comparison strategy. However, it does require an advanced level of statistical and research sophistication. Consequently, the more exhaustive procedure of the *post hoc* test (multiple comparisons test) is more familiar in the research literature. See also: **analysis of variance; Bonferroni test; contrast; Dunn's test; Dunnett's C test; Dunnett's T3 test; Dunnett's test; Dunn–Sidak multiple comparison test; omnibus test; *post hoc* tests**

abscissa: this is the horizontal or *x* axis in a graph. See ***x* axis**

absolute deviation: this is the difference between one numerical value and another numerical value. Negative values are ignored as we are simply measuring the distance between the two numbers. Most

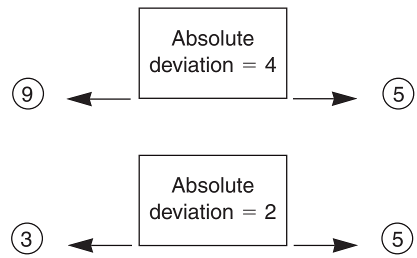

Figure A.1 *Absolute deviations*

commonly, absolute deviation in statistics is the difference between a score and the mean (or sometimes median) of the set of scores. Thus, the absolute deviation of a score of 9 from the mean of 5 is 4. The absolute deviation of a score of 3 from the mean of 5 is 2 (Figure A.1). One advantage of the absolute deviation over deviation is that the former totals (and averages) for a set of scores to values other than 0.0 and so gives some indication of the variability of the scores. See also: **mean deviation; mean, arithmetic**

acquiescence or **yea-saying response set** or **style:** this is the tendency to agree or to say 'yes' to a series of questions. This tendency is the opposite of disagreeing or saying 'no' to a set of questions, sometimes called a nay-saying response set. If agreeing or saying 'yes' to a series of questions results in a high score on the variable that those questions are measuring, such as being anxious, then a high score on the questions may indicate either greater anxiety or a tendency to agree. To control or to counteract this tendency, half of the questions may be worded in the opposite or reverse way so that if a person has a tendency to agree the tendency will cancel itself out when the two sets of items are combined.

adding: see **negative values**

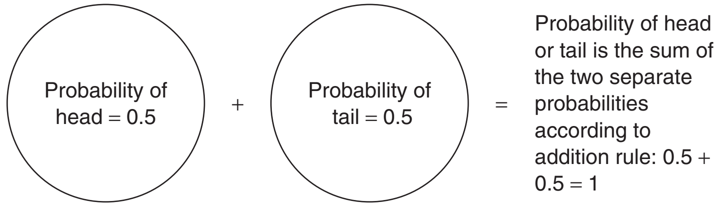

Figure A.2 *Demonstrating the addition rule for the simple case of either heads or tails when tossing a coin*

addition rule: a simple principle of probability theory is that the probability of either of two different outcomes occurring is the sum of the separate probabilities for those two different events (Figure A.2). So, the probability of a die landing 3 is 1 divided by 6 (i.e. 0.167) and the probability of a die landing 5 is 1 divided by 6 (i.e. 0.167 again). The probability of getting either a 3 or a 5 when tossing a die is the sum of the two separate probabilities (i.e. 0.167 + 0.167 = 0.333). Of course, the probability of getting any of the numbers from 1 to 6 spots is 1.0 (i.e. the sum of six probabilities of 0.167).

adjusted means, analysis of covariance: see **analysis of covariance**

agglomeration schedule: a table that shows which variables or clusters of variables are paired together at different stages of a cluster analysis. See **cluster analysis**

Cramer (2003)

algebra: in algebra numbers are represented as letters and other symbols when giving equations or formulae. Algebra therefore is the basis of statistical equations. So a typical example is the formula for the mean:

$$m = \frac{\sum X}{N}$$

In this m stands for the numerical value of the mean, X is the numerical value of a score, N is the number of scores and $\sum$ is the symbol indicating in this case that all of the scores under consideration should be added together.

One difficulty in statistics is that there is a degree of inconsistency in the use of the symbols for different things. So generally speaking, if a formula is used it is important to indicate what you mean by the letters in a separate key.

algorithm: this is a set of steps which describe the process of doing a particular calculation or solving a problem. It is a common term to use to describe the steps in a computer program to do a particular calculation. See also: **heuristic**

alpha error: see **Type I** or **alpha error**

alpha (α) reliability, Cronbach's: one of a number of measures of the internal consistency of items on questionnaires, tests and other instruments. It is used when all the items on the measure (or some of the items) are intended to measure the same concept (such as personality traits such as neuroticism). When a measure is internally consistent, all of the individual questions or items making up that measure should correlate well with the others. One traditional way of checking this is split-half reliability in which the items making up the measure are split into two sets (odd-numbered items versus

Table A.2 *Preferences for four foodstuffs plus a total for number of preferences*

	Q1: bread	Q2: cheese	Q3: butter	Q4: ham	Total
Person 1	0	0	0	0	0
Person 2	1	1	1	0	3
Person 3	1	0	1	1	3
Person 4	1	1	1	1	4
Person 5	0	0	0	1	1
Person 6	0	1	0	0	1

Table A.3 *The data from Table A.2 with Q1 and Q2 added, and Q3 and Q4 added*

	Half A: bread + cheese items	Half B: butter + ham items	Total
Person 1	0	0	0
Person 2	2	1	3
Person 3	1	2	3
Person 4	2	2	4
Person 5	0	1	1
Person 6	1	0	1

even-numbered items, the first half of the items compared with the second half). The two separate sets are then summated to give two separate measures of what would appear to be the same concept. For example, the following four items serve to illustrate a short scale intended to measure liking for different foodstuffs:

1 I like bread *Agree Disagree*
2 I like cheese *Agree Disagree*
3 I like butter *Agree Disagree*
4 I like ham *Agree Disagree*

Responses to these four items are given in Table A.2 for six individuals. One split half of the test might be made up of items 1 and 2, and the other split half is made up of items 3 and 4. These sums are given in Table A.3. If the items measure the same thing, then the two split halves should correlate fairly well together. This turns out to be the case since the correlation of the two split halves with each other is 0.5 (although it is not significant with such a small sample size). Another name for this correlation is the split-half reliability.

Since there are many ways of splitting the items on a measure, there are numerous split halves for most measuring instruments. One could calculate the odd–even reliability for the same data by summing items 1 and 3 and summing items 2 and 4. These two forms of reliability can give different values. This is inevitable as they are based on different combinations of items.

Conceptually alpha is simply the average of all of the possible split-half reliabilities that could be calculated for any set of data. With a measure consisting of four items, these are items 1 and 2 versus items 3 and 4, items 2 and 3 versus items 1 and 4, and items 1 and 3 versus items 2 and 4. Alpha has a big advantage over split-half reliability. It is not dependent on arbitrary selections of items since it incorporates all possible selections of items.

In practice, the calculation is based on the repeated-measures analysis of variance. The data in Table A.2 could be entered into a repeated-measures one-way analysis of variance. The ANOVA summary table is to be found in Table A.4. We then calculate coefficient alpha from the following formula:

$$\text{alpha} = \frac{\text{mean square between people} - \text{mean square residual}}{\text{mean square between people}}$$

$$= \frac{0.600 - 0.200}{0.600} = \frac{0.400}{0.600} = 0.67$$

Of course, SPSS and similar packages simply give the alpha value. See **internal consistency; reliability**

Cramer (1998)

alternative hypothesis: see **hypothesis; hypothesis testing**

AMOS: this is the name of one of the computer programs for carrying out structural

Table A.4 *Repeated-measures ANOVA summary table for data in Table A.2*

	Sums of squares	Degrees of freedom	Means square
Between treatments	0.000	3	0.000 (not needed)
Between people	3.000	5	0.600
Error (residual)	3.000	15	0.200

equation modelling. AMOS stands for *A*nalysis of *Mo*ment *S*tructures. Information about AMOS can be found at the following website:
http://www.smallwaters.com/amos/index.html

See **structural equation modelling**

analysis of covariance (ANCOVA): analysis of covariance is abbreviated as ANCOVA (*an*alysis of *cova*riance). It is a form of analysis of variance (ANOVA). In the simplest case it is used to determine whether the means of the dependent variable for two or more groups of an independent variable or factor differ significantly when the influence of another variable that is correlated with the dependent variable is controlled. For example, if we wanted to determine whether physical fitness differed according to marital status and we had found that physical fitness was correlated with age, we could carry out an analysis of covariance. Physical fitness is the dependent variable. Marital status is the independent variable or factor. It may consist of the four groups of (1) the never married, (2) the married, (3) the separated and divorced, and (4) the widowed. The variable that is controlled is called the covariate, which in this case is age. There may be more than one covariate. For example, we may also wish to control for socio-economic status if we found it was related to physical fitness. The means may be those of one factor or of the interaction of that factor with other factors. For example, we may be interested in the interaction between marital status and gender.

There is no point in carrying out an analysis of covariance unless the dependent variable is correlated with the covariate. There are two main uses or advantages of analysis of covariance. One is to reduce the amount of unexplained or error variance in the dependent variable, which may make it more likely that the means of the factor differ significantly. The main statistic in the analysis of variance or covariance is the F ratio which is the variance of a factor (or its interaction) divided by the error or unexplained variance. Because the covariate is correlated with the dependent variable, some of the variance of the dependent variable will be shared with the covariate. If this shared variance is part of the error variance, then the error variance will be smaller when this shared variance is removed or controlled and the F ratio will be larger and so more likely to be statistically significant.

The other main use of analysis of covariance is where the random assignment of cases to treatments in a true experiment has not resulted in the groups having similar means on variables which are known to be correlated with the dependent variable. Suppose, for example, we were interested in the effect of two different programmes on physical fitness, say swimming and walking. We randomly assigned participants to the two treatments in order to ensure that participants in the two treatments were similar. It would be particularly important that the participants in the two groups would be similar in physical fitness before the treatments. If they differed substantially, then those who were fitter may have less room to become more fit because they were already fit. If we found that they differed considerably initially and we found that fitness before the intervention was related to fitness after the intervention, we could control for this initial difference with analysis of covariance. What analysis of covariance does is to make the initial means on fitness exactly the same for the different treatments. In doing this it is necessary to make an adjustment to the means after the intervention. In other words, the adjusted means will differ from the unadjusted ones. The more the initial means differ, the greater the adjustment will be.

Analysis of covariance assumes that the relationship between the dependent variable and the covariate is the same in the different groups. If this relationship varies between the groups it is not appropriate to use analysis of covariance. This assumption is known as homogeneity of regression. Analysis of covariance, like analysis of variance, also assumes that the variances within the groups are similar or homogeneous. This assumption is called homogeneity of variance. See also: **analysis of variance; Bryant–Paulson simultaneous test procedure; covariate; multivariate analysis of covariance**

Cramer (2003)

analysis of variance (ANOVA): analysis of variance is abbreviated as ANOVA (*an*alysis *o*f *va*riance). There are several kinds of analyses of variance. The simplest kind is a one-way analysis of variance. The term 'one-way' means that there is only one factor or independent variable. 'Two-way' indicates that there are two factors, 'three-way' three factors, and so on. An analysis of variance with two or more factors may be called a factorial analysis of variance. On its own, analysis of variance is often used to refer to an analysis where the scores for a group are unrelated to or come from different cases than those of another group. A repeated-measures analysis of variance is one where the scores of one group are related to or are matched or come from the same cases. The same measure is given to the same or a very similar group of cases on more than one occasion and so is repeated. An analysis of variance where some of the scores are from the same or matched cases and others are from different cases is known as a mixed analysis of variance. Analysis of covariance (ANCOVA) is where one or more variables which are correlated with the dependent variable are removed. Multivariate analysis of variance (MANOVA) and covariance (MANCOVA) is where more than one dependent variable is analysed at the same time. Analysis of variance is not normally used to analyse one factor with only two groups but such an analysis of variance gives the same significance level as an unrelated *t* test with equal variances or the same number of cases in each group. A repeated-measures analysis of variance with only two groups produces the same significance level as a related *t* test. The square root of the *F* ratio is the *t* ratio.

Analysis of variance has a number of advantages. First, it shows whether the means of three or more groups differ in some way although it does not tell us in which way those means differ. To determine that, it is necessary to compare two means (or combination of means) at a time. Second, it provides a more sensitive test of a factor where there is more than one factor because the error term may be reduced. Third, it indicates whether there is a significant interaction between two or more factors. Fourth, in analysis of covariance it offers a more sensitive test of a factor by reducing the error term. And fifth, in multivariate analysis of variance it enables two or more dependent variables to be examined at the same time when their effects may not be significant when analysed separately.

The essential statistic of analysis of variance is the *F* ratio, which was named by Snedecor in honour of Sir Ronald Fisher who developed the test. It is the variance or mean square of an effect divided by the variance or mean square of the error or remaining variance:

$$F \text{ ratio} = \frac{\text{effect variance}}{\text{error variance}}$$

An effect refers to a factor or an interaction between two or more factors. The larger the *F* ratio, the more likely it is to be statistically significant. An *F* ratio will be larger, the bigger are the differences between the means of the groups making up a factor or interaction in relation to the differences within the groups.

The *F* ratio has two sets of degrees of freedom, one for the effect variance and the other for the error variance. The mean square is a shorthand term for the mean squared deviations. The degrees of freedom for a factor are the number of groups in that factor minus one. If we see that the degrees of freedom for a factor is two, then we know that the factor has three groups.

Traditionally, the results of an analysis of variance were presented in the form of a table. Nowadays research papers are likely to contain a large number of analyses and there is no longer sufficient space to show such a table for each analysis. The results for the analysis of an effect may simply be described as follows: 'The effect was found to be statistically significant, $F_{2,12} = 4.72$, $p = 0.031$.' The first subscript (2) for F refers to the degrees of freedom for the effect and the second subscript (12) to those for the error. The value (4.72) is the F ratio. The statistical significance or the probability of this value being statistically significant with those degrees of freedom is 0.031. This may be written as $p < 0.05$. This value may be looked up in the appropriate table which will be found in most statistics texts such as the sources suggested below. The statistical significance of this value is usually provided by statistical software which carries out analysis of variance. Values that the F ratio has to be or exceed to be significant at the 0.05 level are given in Table A.5 for a selection of degrees of freedom. It is important to remember to include the relevant means for each condition in the report as otherwise the statistics are somewhat meaningless. Omitting to include the relevant means or a table of means is a common error among novices.

Table A.5 *Critical values of F*

***df* for error variance**	***df* for effect variance**					
	1	**2**	**3**	**4**	**5**	**∞**
8	5.32	4.46	4.07	3.84	3.69	2.93
12	4.75	3.89	3.49	3.26	3.11	2.30
20	4.35	3.49	3.10	2.87	2.71	1.84
30	4.17	3.32	2.92	2.69	2.53	1.62
40	4.08	3.23	2.84	2.61	2.45	1.51
60	4.00	3.15	2.76	2.53	2.37	1.39
120	3.92	3.07	2.68	2.45	2.29	1.25
∞	3.84	3.00	2.60	2.37	2.21	1.00

If a factor consists of only two groups and the F ratio is significant we know that the means of those two groups differ significantly. If we had good grounds for predicting which of those two means would be bigger, we should divide the significance level of the F ratio by 2 as we are predicting the direction of the difference. In this situation an F ratio with a significance level of 0.10 or less will be significant at the 0.05 level or lower ($0.10/2 = 0.05$).

When a factor consists of more than two groups, the F ratio does not tell us which of those means differ from each other. For example, if we have three means, we have three possible comparisons: (1) mean 1 and mean 2; (2) mean 1 and mean 3; and (3) mean 2 and mean 3. If we have four means, we have six possible comparisons: (1) mean 1 and mean 2; (2) mean 1 and mean 3; (3) mean 1 and mean 4; (4) mean 2 and mean 3; (5) mean 2 and mean 4; and (6) mean 3 and mean 4. In this situation we need to compare two means at a time to determine if they differ significantly. If we had strong grounds for predicting which means should differ, we could use a one-tailed t test. If the scores were unrelated, we would use the unrelated t test. If the scores were related, we would use the related t test. This kind of test or comparison is called a planned comparison or a priori test because the comparison and the test have been planned before the data have been collected.

If we had not predicted or expected the F ratio to be statistically significant, we should use a *post hoc* or an a posteriori test to determine which means differ. There are a number of such tests but no clear consensus about which tests are the most appropriate to use. One option is to reduce the two-tailed 0.05 significance level by dividing it by the number of comparisons to obtain the familywise or experimentwise level. For example, the familywise significance level for three comparisons is 0.0167 ($0.05/3 = 0.0167$). This may be referred to as a Bonferroni adjustment or test. The Scheffé test is suitable for unrelated means which are based on unequal numbers of cases. It is a very conservative test in that means are less likely to differ significantly than with some other tests. Fisher's protected LSD (Least Significant Difference) test is used for unrelated means in an analysis of variance where the means have been adjusted for one or more covariates.

A factorial analysis of variance consisting of two or more factors may be a more sensitive test of a factor than a one-way analysis of variance because the error term in a factorial analysis of variance may be smaller than a one-way analysis of variance. This is because some of the error or unexplained variance in a one-way analysis of variance may be due to one or more of the factors and their interactions in a factorial analysis of variance.

There are several ways of calculating the variance in an analysis of variance which can be done with dummy variables in multiple regression. These methods give the same results in a one-way analysis of variance or a factorial analysis of variance where the number of cases in each group is equal or proportionate. In a two-way factorial analysis where the number of cases in each group is unequal and disproportionate, the results are the same for the interaction but may not be the same for the factors. There is no clear consensus on which method should be used in this situation but it depends on what the aim of the analysis is.

One advantage of a factorial analysis of variance is that it determines whether the interaction between two or more factors is significant. An interaction is where the difference in the means of one factor depends on the conditions in one or more other factors. It is more easily described when the means of the groups making up the interaction are plotted in a graph as shown in Figure A.3.

The figure represents the mean number of errors made by participants who had been deprived of either 4 or 12 hours of sleep and who had been given either alcohol or no alcohol. The vertical axis of the graph reflects the dependent variable, which is the number of errors made. The horizontal axis depicts one of the independent variables, which is sleep deprivation, while the two types of lines in the graph show the other independent variable, which is alcohol. There may be a significant interaction where these lines are not parallel as in this case. The difference in the mean number of errors between the 4 hours' and the 12 hours' sleep deprivation conditions was greater for those given alcohol than those not given alcohol. Another way of describing this interaction is to say the difference in the mean number of errors between the alcohol and the no alcohol group is greater for those deprived of 12 hours of sleep than for those deprived of 4 hours of sleep.

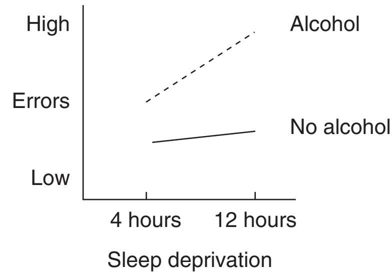

Figure A.3 *Errors as a function of alcohol and sleep deprivation*

The analysis of variance assumes that the variance within each of the groups is equal or homogeneous. There are several tests for determining this. Levene's test is one of these. If the variances are not equal, they may be made to be equal by transforming them arithmetically such as taking their square root or logarithm. See also: **Bartlett's test of sphericity; Cochran's *C* test; Duncan's new multiple range test; factor, in analysis of variance; *F* ratio; Hochberg *GT*2 test; mean square; repeated-measures analysis of variance; sum of squares; Type I hierarchical** or **sequential method; Type II classic experimental method**

Cramer (1998, 2003)

ANCOVA: see **analysis of covariance**

ANOVA: see **analysis of variance**

arithmetic mean: see **mean, arithmetic**

asymmetry: see **symmetry**

asymptotic: this describes a curve that approaches a straight line but never meets it. For example, the tails of the curve of a normal distribution approach the baseline but never touch it. They are said to be asymptotic.

attenuation, correcting correlations for: many variables in the social sciences are measured with some degree of error or unreliability. For example, intelligence is not expected to vary substantially from day to day. Yet scores on an intelligence test may vary suggesting that the test is unreliable. If the measures of two variables are known to be unreliable and those two measures are correlated, the correlation between these two measures will be attenuated or weaker than the correlation between those two variables if they had been measured without any error. The greater the unreliability of the measures, the lower the real relationship will be between those two variables. The correlation between two measures may be corrected for their unreliability if we know the reliability of one or both measures.

The following formula corrects the correlation between two measures when the reliability of those two measures is known:

$$R_c = \frac{\text{correlation between measure 1 and measure 2}}{\sqrt{\text{measure 1 reliability} \times \text{measure 2 reliability}}}$$

For example, if the correlation of the two measures is 0.40 and their reliability is 0.80 and 0.90 respectively, then the correlation corrected for attenuation is 0.47:

$$\frac{0.40}{\sqrt{0.80 \times 0.90}} = \frac{0.40}{\sqrt{0.72}} = \frac{0.40}{0.85} = 0.47$$

The corrected correlation is larger than the uncorrected one.

When the reliability of only one of the measures is known, the formula is

$$R_c = \frac{\text{correlation between measure 1 and measure 2}}{\sqrt{\text{measure 1 or measure 2 reliability}}}$$

For example, if we only knew the reliability of the first but not the second measure then the corrected correlation is 0.45:

$$\frac{0.40}{\sqrt{0.80}} = \frac{0.40}{0.89} = 0.45$$

Typically we are interested in the association or relationship between more than two variables and the unreliability of the measures of those variables is corrected by using structural equation modelling.

attrition: this is a closely related concept to drop-out rate, the process by which some participants or cases in research are lost over the duration of the study. For example, in a follow-up study not all participants in the earlier stages can be contacted for a number of reasons – they have changed address, they choose no longer to participate, etc.

The major problem with attrition is when particular kinds of cases or participants leave the study in disproportionate numbers to other types of participants. For example, if a study is based on the list of electors then it is likely that members of transient populations will leave and may not be contactable at their listed address more frequently than members of stable populations. So, for example, as people living in rented accommodation are more likely to move address quickly but, perhaps, have different attitudes and opinions to others, then their greater rate of attrition in follow-up studies will affect the research findings.

Perhaps a more problematic situation is an experiment (e.g. such as a study of the effect of a particular sort of therapy) in which drop-out from treatment may be affected by the nature of the treatment so, possibly, many more people leave the treatment group than the control group over time.

Attrition is an important factor in assessing the value of any research. It is not a matter which should be hidden in the report of the research. See also: **refusal rates**

average: this is a number representing the usual or typical value in a set of data. It is virtually synonymous with measures of central