

First-year Statistics for Psychology Students  
Through Worked Examples

3. Analysis of Variance

by

**Charles McCreery, D.Phil**

**Formerly Lecturer in Experimental Psychology  
Magdalen College  
Oxford**

**Copyright © Charles McCreery, 2007**

**OXFORD FORUM**

**Psychological Paper No. 2007-3**



# Contents

## **1. Introduction**

- 1.1 Concerning the Appendices
- 1.2 Recommended reading:
- 1.3 Comments

## **2. The Question**

## **3. The Answer**

- 3.1 What is meant by randomization?
- 3.2 The plot
- 3.3 The ANOVA analysis
  - 3.3.1 Hypothesis and assumptions
  - 3.3.2 Computation of the ANOVA
  - 3.3.3 The results
- 3.4 Design precautions which should have been taken
- 3.5 Two potential benefits from a repeated measures design

**Appendix I – steps in a two-way ANOVA**

**Appendix II – the theoretical distinction between populations and samples**

**Appendix III – repeated measures and interaction effects**

**Appendix IV – parametric versus non-parametric tests**

## 1. Introduction

This tutorial is the third in a series based on handouts I developed while teaching statistics to first-year psychology students at Magdalen College, Oxford. The main body of the tutorial is a detailed worked answer to a question from the Oxford University Psychology Prelims examination. The rationale for basing teaching of statistics on worked examples is given in the Introduction to the first tutorial.<sup>1</sup>

### 1.1 Concerning the Appendices

In this worked example of a Two-way ANOVA the arithmetical workings are set out in the most explicit way possible for heuristic purposes; i.e. the object is to show as clearly as possible what is going on. In practice, various short-cut methods of calculation are possible, particularly in an examination context. For example, first-year Oxford University psychology students are allowed the use of a hand calculator in the Psychology Prelims exam, and the Department of Statistics recommends the method of calculation summarized in Appendix 1.

The remaining three appendices deal with more general theoretical points. The first of these, Appendix II, summarises in schematic form the distinction between populations and samples. This distinction is important to an understanding of the theoretical basis of the ANOVA, as well as to the understanding of other tests in statistics.

Appendix III summarises some points concerning repeated measures and interaction effects. These are topics which are not directly involved in the particular worked example in this tutorial, although the concept of repeated measures is alluded to. However, the Appendix may be found useful by people learning about other forms of the ANOVA.

ANOVA is an example of a *parametric* statistical test, and the requirements of such a test are alluded to in the present example. Appendix IV summarises the differences between parametric and non-parametric tests, and lists a number of parametric tests and their non-parametric equivalents.

---

<sup>1</sup> McCreery, C. (2007). 'First-year statistics for psychology students through worked examples: I – The Chi-square test, a test of association between categorical variables'. *Psychological paper No. 2007, 1*. Oxford: Oxford Forum. Online pdf at: <http://www.celiagreen.com/charlesmccreery.html>

## 1.2 Recommended reading:

The present tutorial will probably be found more useful if some preliminary reading around the topic has already been done in advance. I suggest that it does not matter if not everything in the reading is clear on the first go through. After the preliminary reading you can work through the example in the present tutorial and then return to the books, when you may have a better feel for the topic.

I particularly suggest that it is important not to get bogged down in trying to understand the details of areas of the topic which may not be relevant to one's syllabus. Some of the books mentioned below go into the subject in great depth, and with a degree of detail which is probably unnecessary for the purposes of most examinations.

For the purposes of understanding what is going on in a two-way ANOVA I particularly recommend following the worked example given on pp. 355-357 of Murray Spiegel's book listed below. The arithmetical workings are laid out very explicitly, with a minimum of abstract notation. In my view going through an explicit worked example of this kind can give a better understanding of the underlying theoretical basis of ANOVA than a verbal exposition.

Hoel, Paul G. (1976). *Elementary Statistics* (4<sup>th</sup> edition). New York: Wiley. Chapter 11.

*Covers one-way and two-way ANOVAs in a relatively user-friendly manner.*

Howell, David C. (1997). *Statistical Methods for Psychology* (4<sup>th</sup> edition). London: Duxbury Press. Chapters 11-14.

*There are nearly 200 pages on ANOVA in these four chapters, covering all aspects of the subject, including repeated measures designs.*

Hays, William L. (1994). *Statistics* (5<sup>th</sup> edition). Orlando, Florida: Harcourt Brace. Chapters 10-12.

*Rather more than 200 pages in this case. Similar in comprehensiveness to Howell.*

Spiegel, Murray R. (1992). *Schaum's Outline of Theory and Problems of Statistics* (2<sup>nd</sup> edition). New York: McGraw-Hill. Chapter 16.

*The theoretical introduction to this chapter is compressed, but the worked examples in the remainder may be found helpful.*

### **1.3 Comments**

If anyone has any comment to make on this tutorial, or notices any typographical or other error which should be corrected, I should be pleased to hear from them, via:

[reply@oxfordforumpublisher.com](mailto:reply@oxfordforumpublisher.com)

## 2. The Question<sup>2</sup>

(i) What is meant by *randomisation*?

An undesirable effect of some antihistamines is drowsiness, which is a consequence of the effect of the drugs on the central nervous system. These data come from an experiment of Hedges, Hills, Maclay, Newman-Taylor and Turner (1971) to compare the effect on the central nervous system of a placebo and two antihistamines. This was done by measuring the *flicker frequency*<sup>3</sup> some time after drug administration in four volunteers who have taken the three treatments. The data presented here are scaled measures based on the flicker frequency.

Subject Number	Meclastine	Promethazine	Placebo
1	112	112	131
2	48	37	61
3	106	93	112
4	51	46	70

(ii) Plot these data in a meaningful way and comment.

(iii) Carry out an appropriate analysis to examine whether there is a difference between the effects of the different drugs, stating clearly your hypotheses, conclusions and any assumptions made.

(iv) State two precautions which should have been taken in running this experiment.

(v) Give two benefits which would have resulted if more than one measurement for each drug for each subject had been obtained.

(3;4;7;3;3)

**[The numbers in the bottom right hand corner of the question are as they appear in the Oxford Prelims exam paper and indicate to the candidate how many marks each part of the question is potentially worth if answered correctly.]**

---

<sup>2</sup> The question is taken from the Prelims Statistics paper for first-year psychology students at Oxford University, Hilary Term, 1999.

<sup>3</sup> The *flicker frequency*, often referred to as the *critical flicker frequency* or CFF, is that frequency at which a flickering light appears to a particular observer to be continuous rather than discontinuous. It is usually estimated by taking a series of readings from both below and above; that is to say, the experimenter starts with a frequency below the fusion point and gradually increases it until the flicker disappears; he/she then starts again from a frequency above the fusion point and lowers the frequency until the flicker appears again. The CFF reading for a given subject at a given time would then be the mean of several observations taken in both directions, to allow for 'noise' (i.e. random moment-to-moment small fluctuations in the CFF).

### 3. The answer

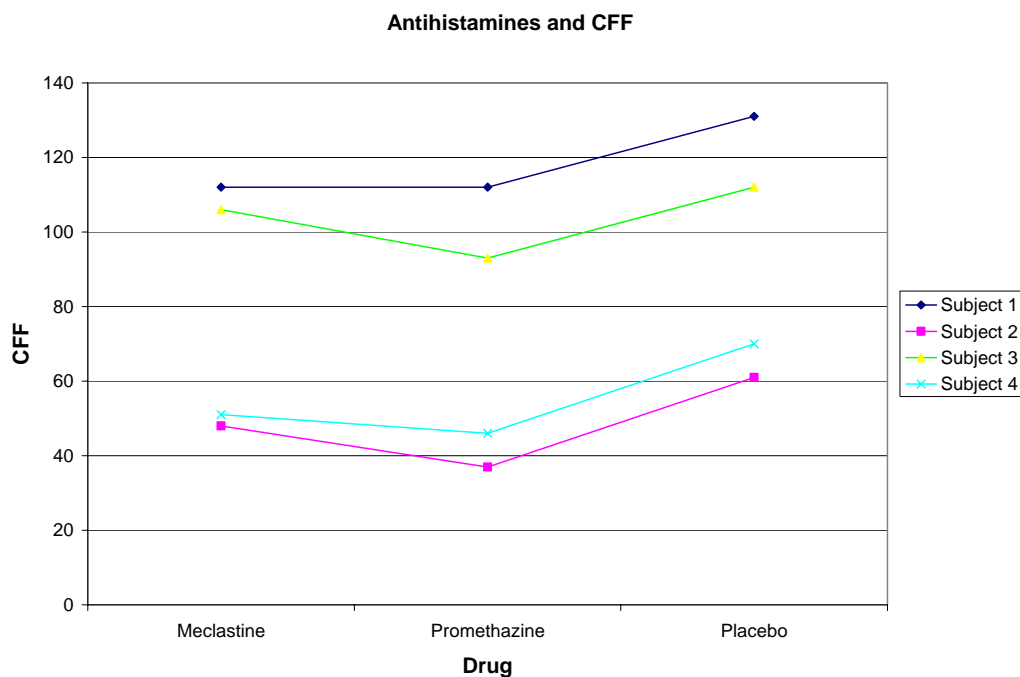
#### 3.1 What is meant by randomization?

*Random sampling*: a sample is said to be taken at random from a population when each member of the population has an equal chance of being chosen. One result is that each member would be selected approximately the same number of times as any other member given a large number of repetitions of the experiment.

The concept of *randomisation* in the design of experiments is a distinct but related concept. It refers to the process of allocating subjects to two or more experimental conditions, or 'treatment groups', by some chance process such as the use of random number tables. Assuming that we wish to have an equal number of subjects in each group, the chances of a given subject being allocated to the each possible condition are thus equal.

The purpose of random allocation of this kind is to produce treatment groups that are as nearly similar as possible prior to the experimental procedure whose effect it is desired to measure. Thus any difference between the groups subsequent to the treatment may be attributed to the experimental treatment and not to some systematic, prior difference between them.

#### 3.2 The plot



### **Comment on the plot:**

The most notable effect appears to be large individual differences in baseline critical flicker frequency or CFF. It also appears that both antihistamines have the effect of depressing the CFF, since the highest mean value in all four subjects is for the placebo condition. Promethazine would appear to depress it more than Meclastine, though this difference may not be statistically significant, given the small number of subjects.

[The data are consistent with the idea that the CFF is a measure of sensory efficiency and hence cortical arousal.]

### **3.3 The ANOVA analysis**

We will perform a two-way ANOVA on the data to determine whether there is any significant difference between the effects of the different drugs. There will be two factors in this analysis: 'Drug', which has three levels<sup>4</sup>, Meclastine, Promethazine and Placebo; and 'Subject', which has four. The inter-subject variability referred to above suggests that there may be a significant main effect<sup>5</sup> for Subject. Whether the intra-subject variability is sufficient to give us a main effect for Drug is less clear.

#### **3.3.1 Hypothesis and assumptions**

The hypothesis is that the two antihistamines depress cortical arousal, and hence visual efficiency as measured by the visual system's power of temporal resolution. We predict that the CFF will therefore be significantly lower under the influence of the antihistamines than under the placebo condition.

Our assumptions are: (1) that the observations under each condition are drawn from a normal distribution, even though there are too few under each condition to constitute a normal sample; and (2) that the variances of each of these underlying normal distributions are approximately equal.

[These are general assumptions underlying all applications of the ANOVA test. Note that the question explicitly mentions that the data are from 'scaled measures'; i.e. they are at least interval data, so a parametric test such as ANOVA is appropriate. For an explanation of the distinctions between interval and other types of measurements, and the implications for what type of test to

---

<sup>4</sup> 'Level' in analysis of variance is simply a technical term for the different possible values of a variable or 'factor'.

<sup>5</sup> For a definition of 'main effect' and a discussion of the difference between this and an 'interaction effect' see Appendix III.



use on a given type of data, please refer to my tutorial on the chi-square test, Appendix B.<sup>6</sup>

Appendix IV to the present tutorial summarises the differences between parametric tests and non-parametric tests in general.]

### 3.3.2 Computation of the ANOVA

Subject Number	Meclastine	Promethazine	Placebo	Row Mean
1	112	112	131	<b>118.33</b>
2	48	37	61	<b>48.67</b>
3	106	93	112	<b>103.67</b>
4	51	46	70	<b>55.67</b>
<b>Column Mean</b>	<b>79.25</b>	<b>72</b>	<b>93.5</b>	

*Calculation of the total sum of squares of the data:*

$$\begin{aligned}
 \text{Grand mean} &= (112 + 48 + 106 + 51 \\
 &\quad + 112 + 37 + 93 + 46 \\
 &\quad + 131 + 61 + 112 + 70)/12 \\
 &= 979/12 \\
 &= 81.58.
 \end{aligned}$$

Observation	Deviation of observation from grand mean	Deviation squared
112	30.42	925.38
48	-33.58	1,127.62
106	24.42	596.34
51	-30.58	935.14
112	30.42	925.38
37	-44.58	1,987.38
93	11.42	130.42
46	-35.58	1,265.94
131	49.42	2,442.34
61	-20.58	423.54
112	30.42	925.38
70	-11.58	134.10
<b>Grand mean =81.58</b>	<b>Total sum of squares (TSS):</b>	<b>11,818.92</b>

<sup>6</sup> McCreery, C. (2007). 'First-year statistics for psychology students through worked examples: I – The Chi-square test, a test of association between categorical variables'. *Psychological paper No. 2007, 1*. Oxford: Oxford Forum. Online pdf at: <http://www.celiagreen.com/charlesmccreery.html>

$$\begin{aligned}
\text{Total number of observations: } n &= (\text{no. of rows}) * (\text{no. of columns}) \\
&= r * c \\
&= 4 * 3 \\
&= 12.
\end{aligned}$$

[The calculation of TSS is laid out in the explicit form above for heuristic purposes. In practice you are likely to be able to use a short-cut method for calculating the TSS. For example, in the Oxford University Psychology Prelims examination the student is allowed the use of a hand calculator, and may therefore calculate the TSS by entering all 12 observations into the calculator, working out their standard deviation (SD) and squaring the SD to get the variance. Having got the variance, you can derive the TSS by multiplying the variance by (n-1), where n is the total number of observations. See Appendix I for a summary of the steps using this method.

N.B. There are two separate operational definitions of variance (and hence of SD, its square root). If we are calculating the variance of a *population*, we divide TSS by N, the total number in the population; but if we are dealing with only a *sample* of a population, as in the present example, we divide by (n-1), where n is the total number in the sample. This is an adjustment designed to allow for *sampling error* in our data, or random variation from sample to sample.

If using a calculator, therefore, it is advisable to be aware which definition is being used by it, as some calculators are able to calculate both versions.

For more on the distinction between populations and samples, see Appendix II.]

*Row calculations:*

Row means	Deviation of row mean from grand mean	Deviation squared
118.33	36.75	1350.81
48.67	-32.91	1083.29
103.67	22.09	487.82
55.67	-25.91	671.50
<b>Grand mean =81.58</b>	<b>Sum of squares:</b>	<b>3593.42</b>

$$\begin{aligned}
\text{Row Sum of Squares (RSS)} &= \text{Sum of squares of row means} * \text{No. of columns} \\
&= 3592.92 * 3 \\
&= 10,778.76.
\end{aligned}$$

[Note that the sum of squares of the row means is the variation in just one column of figures, albeit a representative one. To get an estimate of the overall

variation due to the 'row factor', Subject, we have to extrapolate this variation across the whole two-dimensional array of columns. Hence the need to multiply by the number of columns.

As with the computation of the TSS, the calculation of the RSS shown above is laid out in a manner designed to show the underlying rationale of ANOVA, whereas a short-cut method may be used in an examination in which a calculator is permitted. For details of this method, see Appendix I.]

*Column calculations:*

Column means	Deviation of column mean from grand mean	Deviation squared
79.25	-2.33	5.44
72.00	-9.58	91.84
93.50	11.92	142.01
<b>Grand mean =81.58</b>	<b>Sum of squares:</b>	<b>239.29</b>

[Note that the sum of squares of the column means is much less than that of the row means, reflecting the fact that, simply from visual inspection of the data, there seems to be more variability from one subject to another than within-subject variability due to the different effect of different drugs.]

$$\begin{aligned}
 \text{Column Sum of Squares (CSS)} &= \text{Sum of squares of column means} * \text{No. of rows} \\
 &= 239.29 * 4 \\
 &= 957.16
 \end{aligned}$$

[This time we are multiplying the sum of squares of the *column* means by the number of *rows*, for an analogous reason to that given for multiplying the sum of squares of the row means by the number of columns.

As with the RSS, an alternative method of calculating the CSS is given in Appendix I.]

Within groups/error sum of squares (ESS): -

$$\begin{aligned}
 \text{ESS} &= \text{TSS} - \text{RSS} - \text{CSS} \\
 &= 11,818.92 - 10,778.76 - 957.16 \\
 &= 83.
 \end{aligned}$$

[In an experimental design like this there are only two sources of variation in the CFF readings which are of interest to us: variation associated with different drugs (i.e. between-groups or 'horizontal' variation), and variation associated

with different subjects (i.e. within-groups or 'vertical' variation). So if we know the value of the total variation and the value of both these two sorts of variation, then 'error variation' is simply what is left over when we have subtracted both sorts from the total.

Note that the term 'error' here does not imply any sort of mistake. In another context the error variation might be of interest to us. Some of it might be due to a time of day effect, for example, if the subjects were not all run at the same time of day.

The error term in this case is extremely small relative to the main effects for Subject and Drug. The reason for this can be seen from the plot: the pattern of variation due to drug is very similar in each subject, so that by the time one has taken out the variation in the scores due to drug and that due to subject the effect of any other factor appears to be minimal.]

*Computation of mean squares:*

$$\begin{aligned}\text{Row Mean Square (RMS)} &= \text{RSS}/(r-1) \\ &= 10,778.76/3 \\ &= 3592.92\end{aligned}$$

$$\begin{aligned}\text{Column Mean Square (CMS)} &= \text{CSS}/(c-1) \\ &= 957.16/2 \\ &= 478.58\end{aligned}$$

$$\begin{aligned}\text{Error Mean Square (EMS)} &= \text{ESS}/(r-1)(c-1) \\ &= 83/6 \\ &= 13.83\end{aligned}$$

*Computation of degrees of freedom*

**Rows:**

$$\begin{aligned}\text{Degrees of freedom (d.f.)} &= (r-1) \\ &= (4-1) \\ &= 3\end{aligned}$$

**Columns:**

$$\begin{aligned}\text{D.f.} &= (c-1) \\ &= (3-1) \\ &= 2\end{aligned}$$

**Error**

$$\begin{aligned}
\text{D.f.} &= (r-1)(c-1) \\
&= 3 * 2 \\
&= 6
\end{aligned}$$

[Note that there are always two separate sets of degrees of freedom to be taken into account when looking up the probability of an F value in a two-way ANOVA, one for the numerator in the F ratio (e.g. the RMS if one is looking up the Row main effect) and one for the denominator (always the EMS).]

**3.3.3 Results and conclusions**

<b>ANOVA TABLE</b>				
<b>Source of variation</b>	<b>Degrees of freedom</b>	<b>Sum of Squares</b>	<b>Mean Square</b>	<b>F</b>
<b>Rows</b>	3	10,778.76	3592.92	<i>RMS/EMS=259.79</i>
<b>Columns</b>	2	957.16	478.58	<i>CMS/EMS=34.60</i>
<b>Error</b>	6	83.00	13.83	
<b>Total:</b>	<b>11</b>	<b>11,818.92</b>		

The critical value for F at the 0.05 level with 2 and 6 degrees of freedom is 5.14 (see, for example, Table IX in Hoel, 1976, pp. 338-341). Our F value for the Column Mean Square/Error Mean Square variance ratio (also called the Main Effect for Drug) far exceeds this value; we can therefore reject the null hypothesis that the antihistamines do not significantly depress visual performance.

The Main Effect for Subject is even more significant, as suggested by the plot.

[It can be seen simply from visual inspection of the data that there is far greater variability from subject to subject than there is from one condition to another. This fact is reflected in the far larger value for the row sum of squares than for the columns sum of squares, which is the one we are interested in. In another context we might of course be interested in the within-samples variability, e.g. in a study of individual differences in response to antihistamines, or individual differences in base-line CFF.]

### **3.4 Design precautions which should have been taken:**

(1) The order in which the drugs/placebo were administered to each subject should have been randomised. This would prevent the possibility of the *latent effect*<sup>7</sup> of one of the drugs potentiating the effect of another.

(2) Sufficient time should have been allowed to elapse between the administration of each drug to avoid any possibility of a *carry-over effect*<sup>8</sup> (i.e. the effect of the first drug treatment not having worn off before the second was given).

### **3.5 Two potential benefits from a repeated measures design:**

(1) Administering each drug more than once to each subject would have increased the reliability of the readings.

(2) Repeated administration of each drug to each subject would have meant that it was possible to randomly vary the position of each drug for each subject in successive experiments. E.g. if Meclastine was the first drug to be administered to Subject 1 in the first experiment, then it could have been administered second or third in a second experiment. This randomisation of the order of administration between experiments and within subjects would have provided an even stronger control for any latent effect than simply randomising the order of administration between subjects in one experiment as recommended in section 3.4, (1).

---

<sup>7</sup> See Appendix III

<sup>8</sup> See Appendix III

## Appendix I

### ‘Short-cut’ Method for Computing a Two-Way ANOVA<sup>9</sup>

#### Definitions:

*Number of rows* =  $r$

*Number of columns* =  $c$

1. Calculate the **row means** and **column means**.
2. (a) Calculate the **total variance** of all the observations in the data using a calculator,  
(b) Calculate the **total sum of squares (TSS)**, by multiplying the total variance by  $(n - 1)$ , where  $n$  is the total number of observations.

[*Caveat:* As mentioned in the text above, there are two separate operational definitions of variance (and hence of SD, its square root). Make sure your calculator is using the version for samples, which uses  $(n-1)$ , rather than the version for populations, which uses  $N$ .]

3. Row calculations:
  - (i) Compute the **variance of the row means**; i.e. enter your various row means from 1. into the calculator and find their variance,
  - (ii) Compute the **row mean square (RMS)** by multiplying the variance of the row means by  $c$  (the number of *columns*),
  - (iii) Compute **row sum of squares (RSS)**, by multiplying RMS by  $(r - 1)$ .

[The same caveat applies here as in the case of the short-cut computation of TSS given above: be sure you know which definition of variance your calculator is using.]

---

<sup>9</sup> This Appendix summarises the method recommended by the handbook accompanying the Oxford University Psychology Prelims Statistics course: *Definitions and Formulae with Statistical Tables for Elementary Statistics and Quantitative Methods Courses*, Department of Statistics, University of Oxford, 1999.

4. Column calculations:

- (i) Compute the **variance of the column means**; i.e. enter your various column means from 1. into the calculator and find their variance,
- (ii) Compute **column mean square (CMS)** by multiplying the variance of the column means by  $r$  (the number of *rows*).
- (iii) Compute the **column sum of squares (CSS)** by multiplying CMS by  $(c - 1)$ .

[Observe the usual caveat with regard to the definition of variance.]

5. Compute the **error sum of squares (ESS)**:

i.e. take both row sum of squares (RSS) and column sum of squares (CSS) from total sum of squares (TSS). In other words, take results 3 (iii) and 4 (iii) from result 2. The error sum of squares is whatever is left.

6. Compute the **error mean square (EMS)** from (5).

i.e. divide the result of (5), the error sum of squares, by the product of  $(r - 1)$  and  $(c - 1)$ .

7. You now have all the values you need to complete your ANOVA table (see worked example, above, for a model), work out your F ratios and look up their corresponding p-values.

*Note that there are always two separate sets of degrees of freedom to be taken into account when looking up the probability of an F value in a two-way ANOVA, one for the numerator in the F ratio (e.g. the RMS if one is looking up the Row main effect) and one for the denominator (always the EMS).*



## Appendix II

### THE THEORETICAL DISTINCTION BETWEEN SAMPLES AND POPULATIONS

It's importance: '*Statistical methods may described as methods for drawing conclusions about populations by means of samples.*' Hoel, p.2

	<b>Samples</b>		<b>Populations</b>
<b>Nature of measures:</b>	Always empirical/ <i>a posteriori</i>		May be theoretical/ <i>a priori</i> (e.g. mean IQ score, or predictions from binomial); or, if unknown, may have to be represented by sample.
<b>Represented by:</b>	English alphabet		Mainly Greek alphabet
<b>Examples:</b>	$\bar{x}$	Mean	$\mu$
	$s^2$	Variance	$\sigma^2$
	$S$	Standard deviation	$\sigma$
	$\hat{p}$	Proportion	$p$

#### SAMPLING - Key concepts

##### Sampling error<sup>10</sup>

- or variability from sample to sample due to chance

##### Sampling distribution of a statistic

- 'the most basic concept underlying all statistical tests' (Howell, p.90)

---

<sup>10</sup> N.B. Does not imply any mistake.

- tells us 'what degree of sample-to-sample variability we can expect by chance as a function of sampling error' (Howell, p.90)
- or: 'the distribution of values obtained for that statistic over repeated sampling' (Ibid.)
- derived mathematically rather than empirically

### **Sampling distribution for the mean**

- 'distribution of means of an infinite number of random samples' (Howell, p.90)

### **Standard Error**

- the standard deviation of a sampling distribution of a statistic

## Appendix III

### REPEATED MEASURES AND INTERACTION EFFECTS

#### Repeated Measures

- **Definition:**

Any experiment in which the same variable is measured more than once for each subject is a repeated measures design; e.g. a paired t-test in which two measures, 'before' and 'after', are taken for each S.

- **Advantage of repeated measures designs:**

- they reduce overall variability by using fewer subjects
- they control for individual differences

- **Disadvantages:**

- the *carry-over effect* (e.g. effect of first drug treatment has not worn off before second is given and tested)
- the *latent effect* (second drug potentiates or activates the latent effect of the first)
- the *learning effect* (self-explanatory; most relevant to psychology experiments)

#### Interaction

An *interaction effect* should be thought of in contradistinction to a *main effect*.

- **Main Effects**

A *main effect* is what we would measure if we were to collapse all our data into the several levels of a single factor and compare the resulting means.

- **Interaction Effects**

An *interaction effect* can be thought of as a correlation or interaction between two or more factors. In other words, the influence of the two or more factors is not independent of each other - they interact.

- **Interaction Effects may be 'second-order' but they are not necessarily secondary in importance**

An interaction effect may be of greater interest than a main effect. In fact it may be the main object of investigation of an experiment.

*See Howell (1997) for discussions of the basic ideas of both interaction (pp.409-411) and repeated measures (pp. 450-452).*

**Reference:**

Howell, David C. (1997). *Statistical Methods for Psychology* (4<sup>th</sup> edition). London: Duxbury Press.

## Appendix IV

### PARAMETRIC VERSUS NON-PARAMETRIC TESTS

	<b>Parametric</b>	<b>Non-parametric</b>
<b>Differences:</b>	Assume normality in underlying population distribution	No assumptions about nature of underlying distribution
	Deal in means	Deal in medians
	Deal with intervals	Deal with ranks
<b>Advantages:</b>	More power	Fewer assumptions (so good for skewed distributions, for example)
		Less affected by outliers ( <i>re</i> mean and variance)
		Simpler to calculate
<b>Disadvantages:</b>	Mean and variance disproportionately affected by outliers	Less power (because use less information)
	Affected by skewness	

### CORRESPONDING TESTS

<b>Parametric</b>	<b>Non-parametric</b>
T-test, independent samples	Mann-Whitney ('Rank Sum') Median (less powerful than Mann-Whitney)
T-test, matched pairs	Wilcoxon for matched pairs Sign (less powerful than Wilcoxon)
ANOVA (one-way)	Kruskal-Wallis (generalization of Wilcoxon)
ANOVA (2-factor)	Friedman
Chi-square	Kolmogorov-Smirnov
Product-Moment Correlation	Spearman's Rank Correlation

Some other publications from **OXFORD FORUM** are shown below.  
For further information about each title please click on the image,  
which links to the relevant page on **Amazon UK**.

Alternatively they can be obtained from:

**Blackwells Bookshop**, Oxford (tel. 01865 792792)

or direct from the distributors:

**Book Systems Plus**, Linton (tel. 01223 894870)

