

## PRESENTATION OF MATHEMATICAL AND STATISTICAL DATA

### Numbers and units

Numbers below ten are written in words; numbers over ten (inclusive) are written in figures, except where part of a sequence:

*There were nine learners in the last pilot programme BUT Out of the sample of 26 adults interviewed, only 9 were literate.*

Use a space as a thousands separator – never a comma:

*1 073; 153 655*

Compound numbers below ninety-nine are hyphenated when spelt out, for instance, at the start of a sentence:

*twenty-one BUT one hundred and one*

Fractions are hyphenated when spelt out:

*two-thirds*

Avoid elision:

*156–159 NOT 156–9; 2003–2012 NOT 2003–12*

Use the decimal point (never a decimal comma):

*25.9%*

Insert a space between numerals and units:

*1.23 mm; 2.45 kg ha<sup>-1</sup> BUT 17%; 23 °C; 23½° S*

### Presenting variables

All single-letter mathematical variables must be in italics:

*v = lbh* (where *v* = volume, *l* = length, *b* = breadth, *h* = height)

The italicised single-letter variable rule also applies to subscripts. Thus:  $x_1, x_2, x_3 \dots x_n$ , where 1, 2 and 3 are roman but the subscript *n* is in italics (it is a variable), and *x* is a variable and therefore in italics throughout.

This style convention has implications for reporting of statistics:

*r*<sup>2</sup> (for correlation coefficient); Student's *t*-test; *F* = 34.234 AND *p* < 0.001 (note the spaces either side of the < sign)

For two-letter variables use roman:

SD for standard deviation; CI for confidence interval; FL for fork length

**Note:** The reason for making two- or three-letter variables roman and not italic is that, for example, SD, the abbreviation for standard deviation, means something different to *SD* which means, in mathematics, the variable *S* multiplied by the variable *D*.

### Use of upper and lower case

Many statistical packages are derived from programs written for main frame computers and output was usually all in capitals. Statistical packages do not always distinguish between capital and lower case, italics and roman. Do not simply copy and paste the output from the statistical package directly into your text or tables. Capital letters for a variable may mean something different to the same letter in lower case; for example, *N* is not the same as *n* and *R* means something different to *r*.

### Values less than 1

Use a leading zero when relevant, even when presenting probabilities or statistical values that cannot exceed 1. Do not present probabilities as *p* = .34 or .045. This should always be *p* = 0.34 or 0.045 for the sake of consistency, even if the statistical package outputs the values without a leading zero.

**Note:** Our convention of a leading zero is different to the APA style, which advocates removing the leading zero from any statistics that have a maximum below 1.

## Probabilities less than 0.001

Use sensible probability reporting. If a probability associated with a statistical value is very, very low, e.g.  $p = 0.000005$ , it is customary to represent this as  $p < 0.001$  (to no more than three decimal places). Any probability below 0.1% is insignificantly small.

**Note:** It is incorrect to present a low probability as  $p = 0.000$  and shows that the stats were probably cut-and-pasted from the printout of the statistics package. Replace this representation with  $p < 0.001$ .

## Names of statistical tests

The names of some statistical procedures are derived from the names of two individuals and therefore need to be linked by an en-dash instead of a hyphen. Examples are Mann–Whitney test (or sometimes Mann–Whitney  $U$ -test, with an en-dash AND a hyphen); Kolmogorov–Smirnov test for comparing two samples.

## Other symbols

**Minus sign:** Be careful to use the minus sign, when required, not the hyphen and not the en-dash.

**Primes (')** are used to denote derivatives of mathematical variables and for minutes of angles, especially in denoting latitude or longitude.

*a and a'; 12°14' N; 108°33'54" E*

NISC (Pty) Ltd  
18 August 2014

**TABLE 1: Some common statistics and their presentation**

Statistic	Abbreviation or symbol	Notes
degree of freedom	df	Roman, lower case
<i>F</i> -statistic ( <i>F</i> -value)	<i>F</i>	Most commonly used when reporting the results of ANOVAs. Only the <i>F</i> is italicised. Indicate the associated degrees of freedom in brackets afterwards in the form <i>F</i> (1, 124). Note that the numbers are not subscripted
mean	<i>M</i> or $\bar{x}$	Social scientists tend to use capital <i>M</i> , $\bar{x}$ is used by natural scientists, or better still write out 'mean' <i>M</i> is sometimes used (incorrectly) as an abbreviation for median (a case for writing it out). Median has a variety of different abbreviations (e.g. Mdn, MED, etc.)
sample size (subsample)	<i>n</i>	Three schools were sampled and results combined (School A, <i>n</i> = 17; School B, <i>n</i> = 15; School C, <i>n</i> = 12) [beware of incorrect use of <i>n</i> and <i>N</i> ]
sample size (full sample or population)	<i>N</i>	Three schools were sampled and results combined ( <i>N</i> = 44)
probability ( <i>p</i> -value)	<i>p</i>	Give exact value, such as <i>p</i> = 0.35 or if a very small associated probability, <i>p</i> < 0.005 (or similar) Note the use of lower-case <i>p</i> and a spaced < sign
correlation coefficient <i>r</i> -value	<i>r</i>	
coefficient of determination <i>R</i> <sup>2</sup> -value	<i>R</i> <sup>2</sup>	Note: Superscript is roman
standard deviation	SD	Roman, upper case
standard error	SE	Roman, upper case
<i>t</i> -test or Student's <i>t</i> -test ( <i>t</i> -value)	<i>t</i>	Note lower-case <i>t</i> in <i>t</i> -test. Never start a sentence with 't-test' because you cannot use a capital 'T', which means something else
chi-square	$\chi^2$	Note: Chi-square (not chi-squared). Use the lower-case Greek letter chi