

## TEACHING NON-PARAMETRIC STATISTICS TO STUDENTS IN HEALTH SCIENCES

Michael J. Campbell

ScHARR, University of Sheffield, United Kingdom  
m.j.campbell@sheffield.ac.uk

*There are advantages and disadvantages to teaching non-parametric statistics. On the plus side is the fact that minimal assumptions have to be made, and many of the tests (such as the sign test) are intuitive. On the minus side are the difficulties in teaching statistics purely through significance tests and the difficulty of getting estimates and confidence intervals. My experience is that some disciplines, especially psychology, seem to teach almost exclusively non-parametric statistics, and I have been trying to steer them towards a more parametric approach, since, in my opinion, unless you can write a model, you don't really understand what to infer from a p-value.*

### INTRODUCTION

By non-parametric statistics I mean statistics based on ranks, and this involves continuous data. I do not consider the chi-squared test or the McNemar test to be non-parametric since they deal with binary data. In addition, these tests can be generalised to logistic regression and conditional logistic regression respectively, both of which are most definitely parametric. Another synonym for non-parametric tests is 'distribution free.' Again, I would restrict this to continuous data and not consider the Fisher's Exact Test, or its asymptotic equivalent, the chi-squared test to be in this class.

Statistics is taught at a number of levels to health science students at Sheffield University in the UK, principally at 1<sup>st</sup> year undergraduate, students doing an intercalated medical degree, and post graduate. Non parametric statistics is taught only to those on the intercalated BMedSci degree and a post graduate course for students doing a professional examination for Psychiatry.

In particular the emphasis to the medical undergraduates is on estimation and confidence intervals, following the philosophy espoused in Altman *et al.* (2000). Particularly in the medical literature there is still a prevailing misunderstanding that the objective of an experiment is to get a P-value, and abstracts of papers still contain only the P-value and no estimates of effects or confidence intervals. We feel it is important to correct this at the start and concentrate on estimation, but this is not well served by non-parametric statistics.

The advantages and disadvantages of teaching non-parametric statistics are given in Table 1.

Table 1: Advantages and disadvantages of teaching non-parametric statistics to early health science students

Advantages	Disadvantages
1. Intuitive 2. Relatively easy mathematics 3. Robust, especially in areas such as psychiatry where scales are skewed	1. Concentrates on P-values rather than estimation 2. Difficult to obtain confidence intervals 3. A 'bag of tricks' rather than a coherent methodology.

### *Non-Parametric Tests are Valid for Both Non-Normally Distributed Data and Normally Distributed Data, So Why Not Use Them All the Time?*

It would seem prudent to use non-parametric tests in all cases, which would save one the bother of testing for Normality. Parametric tests are preferred, however, for the following reasons:

1. As I have tried to emphasise in this book, we are rarely interested in a significance test alone; we would like to say something about the population from which the samples came, and this is best done with estimates of parameters and confidence intervals.
2. It is difficult to do flexible modelling with non-parametric tests, for example allowing for confounding factors using multiple regression (see Chapter 11).

*Do Non-Parametric Tests Compare Medians?*

It is a commonly held belief that a Mann-Whitney U test is in fact a test for differences in medians. However, two groups could have the same median and yet have a significant Mann-Whitney U test. Consider the following data for two groups, each with 100 observations. Group 1: 98 (0), 1, 2; Group 2: 51 (0), 1, 48 (2). The median in both cases is 0, but from the Mann-Whitney test  $P < 0.0001$ .

Only if we are prepared to make the additional assumption that the difference in the two groups is simply a shift in location (that is, the distribution of the data in one group is simply shifted by a fixed amount from the other) can we say that the test is a test of the difference in medians. However, if the groups have the same distribution, then a shift in location will move medians and means by the same amount and so the difference in medians is the same as the difference in means. Thus the Mann-Whitney U test is also a test for the difference in means.

*How is the Mann-Whitney U Test Related to the t Test?*

If one were to input the ranks of the data rather than the data themselves into a two sample *t* test program, the P value obtained would be very close to that produced by a Mann-Whitney U test.

*Reading and Reporting Rank Score Tests*

When reading a result from a non-parametric test, one is often faced with a bald P value. One should then ask what hypothesis is being tested? As shown above, one has to make further assumption before statements concerning parameters such as means can be made.

A Mann-Whitney U test such as that described earlier should be reported as ‘Mann-Whitney  $P = 0.075$ , difference in medians  $-5.5$ , 95% CI  $-10$  to  $1.0$ .’ Often, because of inappropriate software, confidence intervals for non-parametric tests are not reported, but there are now a number of packages which will calculate them. For the two sample test, a confidence interval is only interpretable if the distribution of the outcome is similar in the two groups; if the distributions are markedly different, then the two groups differ in more than just a shift in location. Differences in spread can be as important as differences in medians.

REFERENCES

- Carlin, J. B. and Doyle, L. W. (2003). Non-parametric methods for continuous or ordered data. *Journal of Paediatrics and Child Health*, 39(4), 309-311.
- Hart, A. (2001). Mann-Whitney test is not just a test of medians: Differences in spread can be important. *British Medical Journal*, 323, 391-3.
- Nooriafshar, M. (2001). Teaching non-parametric statistics to students from a non-mathematical background. *International Journal for Mathematical Teaching and Learning*, Sept. 17, <http://www.cimt.plymouth.ac.uk/journal/mnstats.pdf>.
- Swinscow, T. D. V. and Campbell, M. J. (2001). *Statistics at Square One* (10<sup>th</sup> edition). London: BMJ Publishing Group.