Early Human Development 123 (2018) 48–49

Contents lists available at ScienceDirect

Early Human Development

journal homepage: www.elsevier.com/locate/earlhumdev

Best practice guidelines

WASP (Write a Scientific Paper): Parametric vs. non-parametric tests

ARTICLE INFO

Keywords:
Research
Teaching
Educational Measurement/methods*
Statistics as Topic

ABSTRACT

Data may be normally distributed in which case standard and relatively familiar tests may be used for inferential testing. Or data may be non-normally distributed, such as ordinal data, in which case, non-parametric tests are usually advocated. This paper will review current thinking with regard to these tests, and the appropriateness of their application/s.

1. Introduction

The statistical papers in this series of Write a Scientific Paper (WASP) guidelines in *Early Human Development* have alluded to non-parametric tests. This paper will delve more deeply into current thinking with regard to these tests, and the appropriateness of their application/s.

2. Differences between parametric and non-parametric tests

Parametric data has an underlying normal distribution such that the variable in question, when plotted, demonstrates a predictable and symmetrical bell-shaped graph, a so-called Gaussian distribution. The principal advantage of data of this type is that since the shape of its distribution is known, inferences may be drawn about values that lie within any part of the distribution curve. These distributions are defined by parameters, namely the mean and standard deviation and it is possible to predict critical values such as those that encompass 68.27%, 95.45% and 99.73% of the values of the variable; 1, 2 and 3 standard deviations respectively [1,3].

Nonparametric tests are distribution-free and make less, if any, assumptions about the distribution of the data's pattern. They instead use the rank order of individual measurements in the dataset rather than the measurements themselves [1]. The properties of parametric and non-parametric datasets are outlined in Table 1.

Non-parametric tests have conventionally been advocated for smaller datasets and/or non-normally distributed data. However, there has been a progressive relaxation of these restrictions. A list of some parametric tests (which are usually more familiar than their non-parametric test) and their non-parametric equivalents are shown in Table 2 [3].

3. Parametric after all?

Compellingly, Geoff Norman, a recognised leader in medical education research, has reviewed this topic, analysing arguments which date back to the 1930s. He has provided evidence (with real and simulated datasets) that parametric tests can in fact be used on ordinal data (such as Likert scales) [6], and are robust (likelier to avoid type 1 and type 2 errors), even when the data in question is very skewed and non-normally distributed [7,8]. Furthermore, parametric tests also perform well when the spread (dispersion, as measured by standard deviation and variance) of each group is different, while non-parametric tests assume that the dispersion of the two groups is similar. Indeed, parametric testing software often have an options dialog which allows the user to specify that the two groups being tested have unknown (and therefore potentially unequal dispersions (variance). Moreover, parametric tests usually have a higher statistical power than nonparametric tests and are thus likelier to avoid a false negative (type 2 error) result.

The overall utility of parametric tests has also been confirmed in separate studies dealing specifically with particular tests [8–10]. It has thus been averred that criticism against the usage of parametric tests may result from excessive reviewer zeal, a tradition that encourages "criticism for its own sake" with "inappropriate statistical dogmatism" ([2]). Nonetheless, a dose of caution would be advocated, especially since the publishability of a paper is likely to be threatened by the use of what could be considered as inappropriate significance tests.

4. When to go nonparametric

It may be appropriate to use a nonparametric technique if the data is so skewed that the median is a far better representation of the data than the mean, as would be the case with severe outliers. This dictum applies even with a large sample size. Another reason to go nonparametric would be a very small sample size. Indeed, sometimes, the sample size available may be so small that it may be impossible to ascertain whether the variable in question is normally distributed or not, in which case a nonparametric test is more appropriate. Minimal requirements for some common parametric tests are displayed in Table 3.

Thus, with small datasets that are non-normal, the inherently smaller statistical power of a small dataset coupled with the necessity for utilising a nonparametric test (which has less statistical power than its parametric equivalent) inevitably increases the chances of a type 2 error.

It is worth noting that while Microsoft Excel does not natively perform these tests at the time of writing, spreadsheets are available

https://doi.org/10.1016/j.earlhumdev.2018.04.014

Table 1
Properties of parametric and non-parametric datasets.

	Parametric	Non-parametric
Distribution	Assumed to be Gaussian (normal)	Any
Variance	Assumed to be homogeneous	Any
Typical data	Discrete or continuous interval	Ordinal or nominal
Data set relationships	Independent	Any
Central measurement	Mean	Median

Table 2
Some commonly used parametric tests and their non-parametric equivalents.

	Parametric test	Non-parametric test
Correlation	Pearson	Spearman/Kendall and others
Two groups, independent measures	*t*-test	Mann-Whitney *U* test (also known as Wilcoxon rank-sum test)
More than two groups, independent measures	One-way analysis of variance (ANOVA)	Kruskal-Wallis test
Repeated measures	Paired *t*-test	Wilcoxon signed-rank test

Table 3
Minimal requirements for common parametric tests.

Test	Requirement
1-sample *t*-test (paired *t*-test)	> 20
2-sample *t*-test (unpaired *t*-test)	Each group should be > 15
One-way ANOVA	If 2–9 groups, each group should be > 15. If 10–12 groups, each group should be > 20.
Pearson correlation	> 30 data points.

online that permit these tests to be done within Excel. Alternatively, an Excel statistics add-in that contains non-parametric tests, or a dedicated statistics software package, can both be used.

5. Normality testing

Should one wish to test for normality before making the choice above, there are several choices of methods that one can use. Whilst the Kolmogorov-Smirnov test is the obvious choice of many, this test is known not to perform well with small datasets – which is where we are typically faced with this decision. The Shapiro-Wilks and Shapiro-Francia tests are better indicated for smaller tests, whilst not perfect. In view of the discussion above, the best approach we have found is using quantile-quantile plots (commonly known as q-q plots) and applying one's discretion. In short, if the plotted data produces a curve, cutting across the diagonal line twice, then it is worth considering a non-parametric test or a transformation. Should the plotted data produce either a perfectly straight line which is nicely adherent to the diagonal line, or a sinusoidal trace that crosses the diagonal line three times (heavy tailed distribution), then a parametric test could be considered.

6. Conclusion

Use parametric tests where possible!

Acknowledgments

The inspiration for this series of papers arises from the international Write a Scientific Paper course (WASP – http://www.ithams.com/wasp) [4,5].

Conflict of interest statement

There are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome.

References

[1] D.G. Altman, J.M. Bland, Parametric v non-parametric methods for data analysis, BMJ 338 (2009 Apr 2) a3167, , https://doi.org/10.1136/bmj.a3167.
[2] P. Bacchetti, Peer review of statistics in medical research: the other problem, BMJ 324 (7348) (2002 May 25) 1271–1273.
[3] M. Fine-Golden, V. Grech, Statistics, in: T. Lissauer, W. Carrol (Eds.), The Science of Paediatrics MRCPCH Mastercourse, 1st edition, Elsevier, London, 2016, pp. 723–738.
[4] V. Grech, WASP - Write a Scientific Paper course: why and how, J. Vis. Commun. Med. 40 (3) (2017 Jul) 130–134, https://doi.org/10.1080/17453054.2017.1366825 (Epub 2017 Sep 19).
[5] V. Grech, S. Cuschieri, Write a Scientific Paper (WASP) - a career-critical skill, Early Hum. Dev. (2018 Jan 17), https://doi.org/10.1016/j.earlhumdev.2018.01.001 ([Epub ahead of print] PMID: 29361328, pii: S0378-3782(18)30005-7).
[6] R. Likert, A technique for the measurement of attitudes, Arch. Psychol. 22 (140) (1932) 55.
[7] G. Norman, Likert scales, levels of measurement and the “laws” of statistics, Adv. Health Sci. Educ. Theory Pract. 15 (5) (2010) 625–632.
[8] G.M. Sullivan, A.R. Artino Jr., Analyzing and interpreting data from Likert-type scales, J. Grad. Med. Educ. 5 (4) (2013) 541–542.
[9] L.M. Sullivan, R.B. D'Agostino Sr., Robustness and power of analysis of covariance applied to ordinal scaled data as arising in randomized controlled trials, Stat. Med. 22 (8) (2003 Apr 30) 1317–1334.
[10] A.J. Vickers, Parametric versus non-parametric statistics in the analysis of randomized trials with non-normally distributed data, BMC Med. Res. Methodol. 5 (2005 Nov 3) 35.

Victor Grech[a,*], Neville Calleja[b],
[a] *Academic Department of Paediatrics, Mater Dei Hospital, Malta*
[b] *Department for Policy in Health - Health Information and Research, Malta*
E-mail addresses: victor.e.grech@gov.mt (V. Grech), neville.calleja@gov.mt (N. Calleja)

* Corresponding author at: Department of Paediatrics, Mater Dei Hospital, Malta.