

ENDGAMES

STATISTICAL QUESTION

Parametric v non-parametric statistical tests

Philip Sedgwick *senior lecturer in medical statistics*

Centre for Medical and Healthcare Education, St George's, University of London, Tooting, London, UK

Researchers investigated five year mortality in patients with chronic heart failure by comparing those with impaired left ventricular function (n=359) with those with preserved function (n=163).¹ A prospective cohort study design was used, with patients enrolled if they had had stable symptomatic chronic heart failure for at least three months.

Characteristics of patients measured at recruitment included age and heart rate. Patients with preserved function were a similar age to those with impaired function (62.5 (standard deviation 10.7) v 62.3 (9.10) years; independent samples *t* test: P=0.80). Those with preserved function had a lower median heart rate (69 (interquartile range 63-82) v 76 (66-89) beats/min; Mann-Whitney U test P<0.001). Five year mortality was significantly greater in patients with impaired left ventricular systolic function (41.5% v 25.2%; P<0.001).

Which of the following statements, if any, are true?

- The independent samples *t* test is a parametric test.
- The use of the independent samples *t* test assumed that age was normally distributed for the patient groups in the population.
- The Mann-Whitney test is a non-parametric test.
- The use of the Mann-Whitney U test assumed the variance of heart rate was equal between patient groups in the population.

Answers

Statements *a*, *b*, and *c* are true, whereas *d* is false.

Two types of statistical methods may be used when analysing data—parametric or non-parametric tests. Non-parametric methods are also referred to as distribution-free methods or methods of rank order. Parametric methods make assumptions about the distribution of the data, whereas non-parametric methods make none.

The independent samples *t* test, also known as the student's *t* test, is a parametric method (*a* is true). Described in a previous question,² the independent samples *t* test compares the means of a variable measured on a continuous scale between two independent groups. The patients with preserved and impaired function were compared in mean age. The null hypothesis states

that the mean age of the patient groups with preserved and impaired function was equal in the population. The population was all those patients with chronic heart failure that met the criteria for entry to the study. Parametric methods make the assumption that the variable being analysed has a particular distribution in the population, typically a Normal distribution. The Normal distribution, described in a previous question,³ is a theoretical distribution described by its mean and standard deviation. The use of the independent samples *t* test assumed the distribution of age in the population was Normal for each patient group (*b* is true). A further assumption also made when using the independent samples *t* test was that the variance of age was the same for each patient group in the population.

Before using the independent samples *t* test, the assumptions described above needed to be verified with the sample data used to estimate the properties of the variable in the population. The distributional assumption of normality could have been confirmed by inspection of the histogram for age in each patient group. Equality of variances for age between patient groups would have been verified by a statistical test, such as Levene's test, which is provided routinely by statistical software.

If the distribution of the sample data is skewed then a transformation—for example, a logarithmic transform—may make the data suitable for analysis using parametric methods. Data transformations will be discussed in a future question. If variances are not equal between patient groups, then statistical software will typically make an adjustment in the application of the independent samples *t* test. The assumptions required for the independent samples *t* test are particularly important when sample sizes are small, with small usually thought to be fewer than 30 in each group; if the assumptions cannot be verified then non-parametric methods should be used.

The assumptions of distributional normality and equal variances between groups could not be made for heart rate so the independent samples *t* test could not be used. The Mann-Whitney U test—the non-parametric equivalent of the independent samples *t* test (*c* is true)—was used instead. Non-parametric methods make no assumptions about the distribution of data in the population or equality of variances between groups (*d* is false). When the Mann-Whitney U test

was used, the null hypothesis stated that the distribution of heart rate was similar for the two groups in the population—that is, the median heart rate for the two groups in the population was equal. The Mann-Whitney U test is based on ranking the values of heart rate regardless of group. Under the null hypothesis, if the distribution of heart rate was similar in each group in the population then the average rank for values of heart rate would be expected to be equal for the two groups in the sample. The Wilcoxon rank sum test is sometimes used instead of the Mann-Whitney U test; the two tests are equivalent and give the same P value, so the same conclusion would be made with respect to statistical hypothesis testing.

The application of parametric methods can generate strong views, and it is sometimes suggested that parametric methods should be used only to analyse data measured on a continuous scale. Variables measured on an ordinal scale, such as a

depression rating scale, that have a large spread in potential values are often analysed using parametric methods. Variables measured on an ordinal scale with a limited range of values should be analysed using non-parametric methods. However, there are no rules regarding the range of values needed before non-parametric methods are used.

Competing interests: None declared.

- 1 MacCarthy PA, Kearney MT, Nolan J, Lee AJ, Prescott RJ, Shah AM, et al. Prognosis in heart failure with preserved left ventricular systolic function: prospective cohort study. *BMJ* 2003;327:78-9.
- 2 Sedgwick P. Independent samples t test. *BMJ* 2010;340:c2673.
- 3 Sedgwick P. The Normal distribution. *BMJ* 2010;341:c6085.

Cite this as: *BMJ* 2012;344:e1753

© BMJ Publishing Group Ltd 2012