PROFESSIONAL DEVELOPMENT

Handling Surgical Data – Choosing a Statistical Test

By
Egyptian Group for Surgical Science and Research
Said Rateb, EGSSR Moderator
Nabil Dowidar, EGSSR Secretary General
Mohamed Farid
Ahmed Hussein
Ahmed Hazem

INTRODUCTION

Considerations in choosing a formal statistical test

In the previous issue of the Egyptian Journal of Surgery we covered the first stage of data analysis, the description and exploration of the data set. In the present issue, we turn to the second stage, namely, the making of formal statistical tests. It may be helpful to suggest that choosing a statistical test depends upon three general considerations:

1. the research question;
2. the nature of the data;
3. the design of the research.

The research question

Consider the following two situations:

a) Suppose that, in order to determine whether drug A or drug B is better at controlling postoperative pain. You have two groups of patients, the first group is given drug A while the other group is given drug B. You ask the two groups to rate their pain at the end of the their first postoperative day on a visual analogue scale (VAS). You want to know whether the pain control is better in one group than in the other. Translating your question into statistical terms, you want to know whether the average VAS is higher for one group than for the other: your research question is one of comparing averages.

b) Now suppose that you want to know whether there is a tendency for long surgical incisions to cause more postoperative pain. In statistical terms, you are asking whether there is a relation between the measured variables of pain (VAS) and the length (cm) of the surgical incision: your research question is one of association.

The answers to questions of comparison (situation a) and association (situation b) are provided by quite different statistical tests.

The nature of the data

Another crucial consideration when choosing a statistical test is the nature of the data set (*see previous issue of the EJS*).

Three kinds of data are known:

1. interval data, which are measurements on an independent scale with units;
2. ordinal data, consisting of ranks, of assignments to ordered categories, or of sequencing information;.
3. nominal or categorial data, which are records of qualitative category membership.

Different statistics are appropriate for descriptions of the three types of data; and the corresponding formal tests are also different.

The design of the research

Number of samples (groups)

If the VAS of a group of patients who were given drug A for the control of their postoperative pain is compared with that of a group of patients who were given drug B, we shall obtain two samples of scores; but if, we test a third group with drug C, there will be three samples of data. Again, the choice of a formal statistical test will be different in the two cases.

Independent versus related samples

If we compare the VAS of patients given drug A with those given drug B. The experiment will yield two independent samples of data. However, if we compare the VAS of patients who were given drug A at the end of the first postoperative day to their VAS at the end of the second postoperative day. The experiment will yield two related samples of data for that group of patients. The scores in two related samples are likely to be substantially correlated; whereas that is unlikely to be the case with scores in independent samples. Accordingly, different statistical tests are appropriate for use with independent and related samples of data.

CHOOSING APPROPRIATE TESTS FOR COMPARISON

To enter the table, one must consider:

1. Whether the samples of the data are related or independent (i.e. whether the design is between subjects or within subjects);
2. How many samples of data there are;
3. The type of data (interval, ordinal or nominal) that will be used in the test.

Type of data	Between subjects (independent samples)	Within subjects (related samples)
	TWO SAMPLES	TWO SAMPLES
Interval	Independent samples t-test	Paired samples t-test
Ordinal	Wilcoxon-Mann-Whitney Test	Wilcoxon signed ranks test. Sign test
Nominal	Chi-square test	McNemar test
	THREE OR MORE SAMPLES	THREE OR MORE SAMPLES
Interval	One-way ANOVA	Repeated measures ANOVA
Ordinal	Kruskal-Wallis k-sample test	Friedman test
Nominal	Chi-square test	Cochran's Q test (dichotomous nominal data only)

Parametric and nonparametric test

The t-test is an example of a parametric test, that is, it is assumed that the data are samples from a population with a normal distribution. Other tests, known as nonparametric tests such as Wilcoxon-Mann-Whitney test, do not make specific assumptions about population distributions. For that reason, they are often referred to as distribution-free tests.

Some experts strongly recommend the use of nonparametric tests, others emphasise the robustness of the parametric t-tests to violations of their assumptions and the loss of power incurred by the use of equivalent nonparametric tests. It is suggested that, provided the data show no obvious contra-indications such as the presence of outliers, marked skewness or great

disparity of variances, a t-test can safely be used; otherwise a nonparametric equivalent (see *tests used for ordinal data in the previous table)* should be considered.

On the other hand, if the data comprise measurements at the ordinal level or nominal data, a nonparametric test is the only possibility.

CHOOSING APPROPRIATE TESTS FOR ASSOCIATION

Consider the following research question: do longer surgical incisions cause more postoperative pain? Your question is whether there is a statistical association between the two variables; incision length and pain. To answer the question, you would need the lengths of a substantial sample of incisions and the VAS of the pain experienced by the patients. Since you would have interval data, an appropriate statistic would be a **Pearson correlation coefficient**.

Now suppose we have asked two surgeons to rank twenty analgesics in order of preference. We should have a data set consisting of (20) pairs of ranks. Do the surgeons agree? Again, our question is one of statistical association, but a rank correlation would be an appropriate statistic. Since you would have ordinal data, an appropriate statistic would be a **Spearman's rank correlation** or **Kendall's tau statistics**.

In the case of nominal data, an appropriate statistic would be **phi coefficient** or **Cramer's V**.

The followng table shows some of the correlation coefficients and other measures of strength of association that are available for the different kinds of data.

Measures of association between two variables

	Statistic test
Interval Data	Pearson correlation (r)
Ordinal Data	Spearman's rho, Kendall's tau-a, tau-b, tau-c
Nominal Data	Phi coefficient Cramer's V