Making sense of statistics for family practitioners: The "Chi-square test" – getting it right

Ogunbanjo GA, Durrheim DN

The chi-square test is the most commonly used statistical test for investigating the differences between proportions arising from research. It helps to determine where two or more series of proportions are significantly different from one another or whether a single series of proportions differs from a theoretically expected distribution. In addition, it allows us to test whether any observed relationship could simply be due to "chance". Proportions would obviously differ from sample to sample selected simply due to the play of chance. The chi-square test thus assists us in making this judgment in an explicit way, and is a measure of the difference between the proportions observed and those that would have been expected if the null hypothesis of no difference between groups had been true.

Once this test statistic is calculated, and we shall consider how this is done in an example below, one uses a statistical table to look up the two-tailed probability (P) of finding a chi-squared value of this magnitude due to chance alone, taking into account the "degrees of freedom". The latter traditionally appears on the left-hand side of such tables and indicates the number of groups being compared, where the degrees of freedom usually are one less than the number of groups being compared. Thus where the proportions from two groups are being compared, for example in a classic case-control study, there is one degree of freedom. For the family practitioner, the question may arise – "How does one interpret chi-square test results provided in scientific publications?" Practically, a P value in the chi-square table which is less than 0,05, implies that the observed difference between groups would be expected to occur by chance less than 5 times in 100 samples. As the P value increases above 0,05, the likelihood that the observed differences between groups occurred simply due to chance increases and this is traditionally recorded as "not significant" (NS).

There are different types of chi-square that we think the family practitioner should be aware of, namely:

- Pearson's chi-square: This is the most commonly used chi-square test for testing the hypothesis of no association between column and row data in a table.
- Chi-square goodness-of-fit test: This is a different use of Pearsonian chi-square, to test whether an observed distribution conforms to a specific statistical distribution.
- Mantel-Haenszel chi-square: This is a pooled summary statistic from a series of 2x2 tables.

At this stage, you may be wondering about the relevance of the chi-square test to your clinical practice. We will demonstrate its importance by using the following example.

Assume that you have 150 male patients in your practice between the ages of 30 and 39 years and, 60 smoke dagga (marijuana). You have in the past year diagnosed 50 male patients in this age group with myocardial ischaemia, of which 40 were dagga smokers. You are concerned that myocardial ischaemia may be associated with dagga smoking.

Step 1: Construct a 2x2 table

	Non dagga smoker	Dagga smoker	Total
Myocardial ischaemia	10	40	50
No myocardial ischaemia	80	20	100
Total	90	60	150

Step 2: Construct a 2x2 table if dagga smoking were not associated with myocardial ischaemia (to determine the expected value for the cells you multiply the total at the bottom of column with the total of row of that cell).

EXPECTED	Non dagga smoker	Dagga smoker	Total
Myocardial ischaemia	90*50/150 = 30	60*50/150 = 20	50
No myocardial ischaemia	90*100/150 = 60	60*100/150 = 40	100
Total	90	60	150

Step 3: Compare the numbers in the expected table with the actual ones to see if there is a difference. If there were no real difference between dagga smokers and non-dagga smokers, then we would expect only small differences between cells, if any. The chi-square test is now used to formally calculate whether there is a significant difference by subtracting the expected value from the observed value for each cell, squaring this difference and then dividing this square by the expected value of the cell, before adding the values for each cell.

In this example: Chi-square = $(10-30)^2/30 + (40-20)^2/20 + (80-60)^2/60 + (20-40)^2/40 = 50$.

When one looks up this value in the chi-square table, with one degree of freedom, we find that the $P < 0.0001$; i.e. the possibility of finding an association this extreme by chance alone is less than 1 chance in 10,000. This means that the association between dagga smoking and myocardial ischaemia in this clinical practice is "real" and not by "chance".

There are a few conditions that govern the valid use of the chi-square test. If these are not satisfied then it is necessary to use the Fischer's exact test, which we will discuss in a future article in this series. To use the chi-square test, the total sample size should be greater than 40, or if between 20 and 40 no cell's expected value should be smaller than 5. The next time you read about chi-square, bear in mind that, a large chi-square indicates that the relationship is "real" and "statistically significant" (not due to chance), and a small chi-square indicates that the relationship is probably due to "chance".

Prof. Gboyega A Ogunbanjo MBBS, MFGP (SA), M Fam Med (Medunsa), FACRRM
Department of Family Medicine & Primary Health Care, Medunsa Pretoria South Africa.

Prof. David N Durrheim MBChB, DTM&H, DCH, MPH&TM FACTM, FFTM
School of Public Health & Tropical Medicine, James Cook University, Townsville Australia