



Statistical Corner

Wilfred Hing-Sang WONG 黃慶生

Department of Paediatrics & Adolescent Medicine, The University of Hong Kong, Hong Kong

Question 1:

How to test inter-rater agreement?

Answer :

Many clinical researches monitor parameters scored by different scorers. It is important to look for reliability between different scorers. Kappa can be used to measure inter-rater agreement. Kappa statistics is meaningful when the scale of interest is nominal (e.g. Yes/No). If the agreement categories are ordered, e.g. mild, moderate and severe, the kappa does not account for the degree of disagreement and treat all disagreements equally. In this case weighted kappa is preferred, since it assigns different weights to subjects for whom the raters differ by categories, so that different levels of agreements are accounted for. If the scale of interest is ordinal (e.g. poor, moderate and good asthma control), weighted kappa is preferred as Kappa takes into account the agreement between raters but also quantify the relative difference between categories in ordinal scale.

Question 2:

How to interpret Kappa?

Answer:

The result of Kappa is less than or equal to 1. A value of 1 which is means that perfect agreement between the rater. Below is the suggestion of interpretation of Kappa.¹

Kappa result	Possible interpretation
Less than 0.2	Poor
0.2 to 0.4	Fair
0.41 to 0.6	Moderate
0.61 to 0.8	Good
0.81 to 1.00	Very good

Question 3:

How to use SPSS to compute Kappa?

Answer:

Select Analyze Descriptive statistics Crosstabs. In the dialog box, click Statistics and then select Kappa

Question 4:

The calculation of Kappa can also be done on the web?

Answer:

<http://www.dmi.columbia.edu/homepages/chuangj/kappa/>

Reference

1. Altman DG. Practical Statistics for Medical Research. Chapman & Hall. First edition 1991; page:404.