How can I calculate a kappa statistic for variables with unequal score ranges?

Authored by

stats writer

June 30, 2024

The kappa statistic is a measure of agreement between two variables, often used in fields such as statistics, psychology, and medicine. It is used to determine the level of agreement between two raters or between two sets of measurements. However, calculating the kappa statistic for variables with unequal score ranges can be challenging. In order to accurately calculate the kappa statistic in these cases, it is important to first standardize the scores to a common range. This can be achieved by converting the scores to z-scores or by rescaling the scores to a common range. Once the scores have been standardized, the kappa statistic can be calculated using the standard formula. By properly addressing the issue of unequal score ranges, the kappa statistic can provide a reliable measure of agreement between variables.

How can I calculate a kappa statistic for variables with unequal score ranges? | SPSS FAQ

Suppose we would like to compare two raters using a kappa statistic but the raters have different range of scores. This situation most often presents itself where one of the raters did not use the same range of scores as the other rater.

Let us consider an example where two graduate students where asked to rate 12 movies based on a scale from 1-3. One rater used all of the three scores possible while rating the movies whereas the other student did not like any of the movies and therefore rated all of them as either a 1 or a 2. Thus, the range

of scores is the not the same for the two raters.

To obtain the kappa statistic in SPSS we are going to use the crosstabs command with the statistics
= kappa
option. By default,
SPSS will only compute the kappa statistics if the two variables have exactly the same categories, which is not the case in this
particular instance. We can get around this problem by adding a fake observation
and a weight variable shown below. The weight variable takes value of 1 for all
the real observations and value of 0.00001 (something very small) for the fake
observation that we have just added. The trick is then to weight the
observations using the
weight
command.

```
data list list
/rater1 rater2.
begin data.
```

```
1 1
1 1
1 1
1 1
2 2
2 2
2 2
2 2
3 2
3 2
3 2
3 2
end data.

save outfile = kappa.
data list list
/rater1 rater2.
begin data.
3 3
end data.

add files file = *
/file = kappa.
exe.
```

compute weight = 1.
if (rater1 =3 & rater2 =3) weight = .00001.
exe.

weight by weight.
crosstabs
/tables=rater1 by rater2
/statistics=kappa.

Symmetric Measures

	Value	Asymp. Std. Error(a)	Approx. T (b)	Approx. Sig.
Measure of Agreement	Kappa	.500	3.000	.003
N of Valid Cases		12		

a. Not assuming the null hypothesis.
b. Using the asymptotic standard error assuming the null hypothesis.