

# Calculating IRR

There are couple of ways of measuring IRR.

## Light's Kappa

If there are more than 2 raters, then the average of all possible two-raters kappa is known as Light's kappa (Conger 1980).

Light's kappa is calculated by taking the observed agreement between the two raters and subtracting the expected agreement, which is the agreement that would be expected by chance. The result is then adjusted for the number of categories that the raters are rating.

## Fliess Kappa

The Fleiss kappa is an inter-rater agreement measure that extends the Cohen's Kappa for evaluating the level of agreement between two or more raters, when the method of assessment is measured on a categorical scale. It expresses the degree to which the observed proportion of agreement among raters exceeds what would be expected if all raters made their ratings completely randomly. It is an extension of the Cohen's Kappa.

Fleiss kappa is calculated in a similar way as Light's kappa, but it takes into account the agreement between all of the raters. The result is then adjusted for the number of categories that the raters are rating and the number of raters.

The Fleiss Kappa can be specially used when participants are rated by different sets of raters. This means that the raters responsible for rating one subject are not assumed to be the same as those responsible for rating another (Fleiss et al., 2003).

Fleiss kappa is more accurate than Light's kappa when there are three or more raters. However, Light's kappa is simpler to calculate and interpret.

*## Light's Kappa works for instances with more than two raters.*

```
kappam.light(t1[, 2:6])
```

Light's Kappa for m Raters

Subjects = 21

Raters = 5

Kappa = 0.75

z = 0.311

p-value = 0.755

*## As does the Fliess's Kappa.*

```
kappam.fleiss(t1[, 2:6])
```

Fleiss' Kappa for m Raters

Subjects = 21

Raters = 5

Kappa = 0.747

z = 10.8

p-value = 0