Binomial Confidence Intervals for Proportions

Introduction

Binomial Confidence Intervals for Proportions (BinoCIProps) (http://www.genecalcs.weebly.com/binociprops.html) is designed to calculate 90, 95, or 99% confidence intervals for a sample proportion. The calculations are based on the binomial distribution and are, therefore, considered 'exact' (as opposed to being based on an approximation). BinoCIProps can be used to calculate confidence intervals (CIs) for any scenario where a binomial proportion is calculated (e.g. any scenario with *n* trials where the outcome can be classed as either a success or a failure and where the probability of success is the same for each trial). In clinical molecular genetics, prime examples relate to the calculation of sensitivity and specificity for a qualitative diagnostic test (e.g. the proportion of variants detected by a test in a validation exercise). The validation of genetic tests, the issues of sensitivity and specificity, and the importance of correctly calculating and stating CIs is discussed in an article by Mattocks *et al*. (2010). The calculations undertaken by BinoCIProps are in accordance with those recommended in that article for qualitative tests.

The CI calculations undertaken by BinoCIProps differ according to whether the proportion is complete or incomplete. In the following formulae, *pL* is the lower CI value, *pU* is the upper CI value, *n* is the number of trials, *x* is the number of success, and *α* is the alpha level (e.g. 1-0.95 = 0.05 for 95% confidence).

Complete Proportions

When the proportion is complete (i.e. *x* = 0 or *n*), a one-sided binomial calculation is undertaken.

When *x* = 0: $pL = 0$ and $pU = 1 - (\sqrt[n]{\propto})$ (or $1 - (\propto^{1/n})$.
When *x* = *n*: $pL = \sqrt[n]{\propto}$ (or $\propto^{1/n}$) and $pU = 100$.

These formulae are the correct calculations of which the so called 'rule of 3' is an approximation, and it is these calculations that BinoCIProps undertakes for a complete proportion. For an excellent explanation of the approximate 'rule of 3' and the correct one-sided binomial calculation, please refer to Hanley & Lippman-Hand (1983).

Incomplete Proportions

When the proportion is incomplete (i.e. 0 < *x* < *n*), a two-sided binomial calculation based on the method described by Clopper & Pearson (1934) is undertaken. These calculations are based on the following two equations, which BinoCIProps solves for the values of *pL* and *pU* using formulae based on the *F* distribution:

$$\sum_{k=0}^{x-1} \binom{n}{k} pL^k (1 - pL)^{n-k} = 1 - (\frac{\alpha}{2})$$

The lower CI value is determined by solving the above equation for the value of *pL*.

$$\sum_{k=0}^{x} \binom{n}{k} pU^k (1 - pU)^{n-k} = \frac{\alpha}{2}$$

The upper CI value is determined by solving the above equation for the value of *pU*.

References

Clopper & Pearson (1934). *Biometrika*, 26(4): 404-413.
Hanley & Lippman-Hand (1983). *JAMA*, 249(13): 1743-1745. PMID: 6827763.
Mattocks *et al*. (2010). *Eur J Hum Genet*, 18(12): 1276-1288. PMID: 20664632.