Confidence intervals for a binomial proportion: comparison of methods and software evaluation

Ana M. Pires[1]

[1] Department of Mathematics and Applied Mathematics Centre, IST, Technical University of Lisbon, Portugal.

Keywords. Arcsine transformation, Binomial distribution, Bootstrap confidence intervals, Continuity correction, Normal approximation.

There are a great number of practical situations where it is necessary to obtain a confidence interval for a population proportion, p (e. g. medical statistics, acceptance sampling by attributes, marketing research, survey sampling). It is well known that there are several methods for doing this. However most of the methods rely on asymptotic approximations and the validity of the approximations is not always stated, except for the fact that the usual approximation (considered in almost every introductory course) is poor when the true p is close to zero or to one. Comparisons between the methods are usually based on single cases (it was not possible to find a textbook listing the most common methods and where a general comparison is made, not even in Fleiss, 1981, dedicated exclusively to rates and proportions). Some papers addressing this problem have appeared recently (Vollset, 1993, Newcombe, 1998) but they have missed important aspects and apparently have had no impact on statistical software or on textbooks.

For this study twelve noniterative methods were selected and numerically compared in terms of coverage probability and expected length, at 25000 parameter space points. The twelve methods are: the usually called exact method, based on the inversion of the binomial test, known as Clopper-Pearson interval (I); a Bayesian interval (II); the method based on the normal approximation with the true proportion on the variance, known as score or Wilson interval, with (III) and without (IV) continuity correction; the usual normal approximation, with (V) and without (VI) continuity correction; four types of bootstrap intervals for which it is not necessary to use Monte Carlo (VII) to (X); method based on the arcsine (variance stabilising) transformation, with two types of continuity correction (XI) and (XII). The explicit limits of the confidence intervals for the twelve methods considered are given in Table 1 (X denotes the number of successes in a random sample of size n). Only five of these methods were included in the comparative studies of Vollset (1993) or Newcombe (1998). Moreover, the explicit expressions in terms of Beta percentiles, which allow a very easy determination of the exact interval, were not considered by those authors.

From the results (not presented due to lack of space) a clear classification of the methods as emerged. A first group of conservative methods, even for p close to zero or to one, contains methods (I), (IV), (XI) and (XII). A second group of methods, attaining in many situations a coverage probability smaller than the specified, but without large deviations, is formed by methods (II) and (III). Finally, all the others, which can have, depending on p and n, much smaller coverage probability than specified.

The methods available for the computation of confidence intervals for a binomial proportion in three major statistical packages (SAS, S-Plus and SPSS) were analysed. The main conclusion is that they need an urgent (but

Table 1. Explicit limits of the confidence intervals for the twelve methods.

Method	Lower limit	Upper limit
I [a]	0 if $X=0$, $(\alpha/2)^{1/n}$ if $X=n$ $B_{X,n-X+1;\alpha/2}$ if $0<X<n$	$1-(\alpha/2)^{1/n}$ if $X=0$, 1 if $X=n$ $B_{X+1,n-X;1-\alpha/2}$ if $0<X<n$
II [a]	0 if $X=0$, $\alpha^{1/(n+1)}$ if $X=n$ $B_{X+1,n-X+1;\alpha/2}$ if $0<X<n$	$1-\alpha^{1/(n+1)}$ if $X=0$, 1 if $X=n$ $B_{X+1,n-X+1;1-\alpha/2}$ if $0<X<n$
III [b]	$\frac{2X+c^2-c\sqrt{c^2+4X(1-X/n)}}{2(n+c^2)}$	$\frac{2X+c^2+c\sqrt{c^2+4X(1-X/n)}}{2(n+c^2)}$
IV [b]	0 if $X=0$, otherwise $\frac{2X+c^2-1-c\sqrt{c^2-2-1/n+4X(1-(X-1)/n)}}{2(n+c^2)}$	1 if $X=n$, otherwise $\frac{2X+c^2+1+c\sqrt{c^2+2-1/n+4X(1-(X+1)/n)}}{2(n+c^2)}$
V [b]	$\max\left\{\frac{X}{n}-c\sqrt{\frac{X}{n^2}\left(1-\frac{X}{n}\right)};0\right\}$	$\min\left\{\frac{X}{n}+c\sqrt{\frac{X}{n^2}\left(1-\frac{X}{n}\right)};1\right\}$
VI [b]	$\max\left\{\frac{X}{n}-c\sqrt{\frac{X}{n^2}\left(1-\frac{X}{n}\right)}-\frac{1}{2n};0\right\}$	$\min\left\{\frac{X}{n}+c\sqrt{\frac{X}{n^2}\left(1-\frac{X}{n}\right)}+\frac{1}{2n};1\right\}$
VII [c]	$\frac{\mathrm{Bin}_{n,X/n;\alpha/2}}{n}$	$\frac{\mathrm{Bin}_{n,X/n;1-\alpha/2}}{n}$
VIII [c]	$\max\left\{\frac{\mathrm{Bin}_{n,X/n;\alpha/2}}{n}-\frac{1}{2n};0\right\}$	$\min\left\{\frac{\mathrm{Bin}_{n,X/n;1-\alpha/2}}{n}+\frac{1}{2n};1\right\}$
IX [c][d]	$\frac{\mathrm{Bin}_{n,X/n;\alpha''}}{n}$	$\frac{\mathrm{Bin}_{n,X/n;\alpha''}}{n}$
X [c][d]	$\max\left\{\frac{\mathrm{Bin}_{n,X/n;\alpha'}}{n}-\frac{1}{2n};0\right\}$	$\min\left\{\frac{\mathrm{Bin}_{n,X/n;\alpha''}}{n}+\frac{1}{2n};1\right\}$
XI [b]	0 if $X=0$, otherwise $\sin^2\left(\arcsin\sqrt{\frac{X-0.5}{n}}-\frac{c}{2\sqrt{n}}\right)$	1 if $X=n$, otherwise $\sin^2\left(\arcsin\sqrt{\frac{X+0.5}{n}}+\frac{c}{2\sqrt{n}}\right)$
XII [b]	0 if $X=0$, otherwise $\sin^2\left(\arcsin\sqrt{\frac{3/8+X-0.5}{n+3/4}}-\frac{c}{2\sqrt{n+1/2}}\right)$	1 if $X=n$, otherwise $\sin^2\left(\arcsin\sqrt{\frac{3/8+X+0.5}{n+3/4}}+\frac{c}{2\sqrt{n+1/2}}\right)$

[a] $B_{\theta_1,\theta_2;\gamma}$ is the γ percentile of the Beta(θ_1,θ_2) distribution. [b] $c=z_{1-\alpha/2}$, where z_γ is the γ percentile of the $\mathcal{N}(0,1)$ distribution. [c] $\mathrm{Bin}_{n,\theta;\gamma}$ is the γ percentile of the Bin(n,θ) distribution. [d] α' and α'' are bias corrected percentiles.

very easy) revision. In SAS the only method used is method (V), which is one of the worst. In S-Plus the `binom.test` command should also give the exact interval (line I of Table 1); the `prop.test` command, which gives the intervals by methods (III) and (IV), should have the information given in the previous paragraph (it was also noticed that, for method (IV), when $X=0$ and $X=n$, the intervals given are wrong). SPSS provides the exact and the asymptotic tests for a binomial proportion but no confidence intervals.

References

Fleiss, J. L. (1981). *Statistical Methods for Rates and Proportions (2nd ed.)*. New York: Wiley.

Newcombe, R. G. (1998). Two-sided confidence intervals for the single proportion: comparison of seven methods. *Statistics in Medicine*, **17**, 857-872.

Vollset, S. E. (1993). Confidence intervals for a binomial proportion. *Statistics in Medicine*, **12**, 809-827.