A Note on Teaching Large–Sample Binomial Confidence Intervals

Thomas J. Santner[1]
Department of Statistics
Ohio State University
Columbus, OH, 43210

Abstract

This paper addresses the question of which large–sample binomial confidence interval method to teach in elementary courses. Recently, Goodall (1995), Simon (1996), and the references therein, have discussed the merits of several large–sample systems of binomial intervals. Three of the systems considered these articles are the z-interval, a "t-based interval," and a continuity-corrected interval, the latter denoted by c-interval herein. The 95% nominal z limits are

$$\hat{p} \pm 1.96 \times S.E.(\hat{p}) \tag{0.1}$$

where $S.E.(\hat{p}) = \sqrt{\hat{p} \times (1-\hat{p})/n}$ is an estimate of the standard error of $\hat{p}$. The t-intervals widen the (0.1) limits by using a larger critical point; they are $\hat{p} \pm t \times S.E.(\hat{p})$, where t is the two-sided upper-0.05 critical point of the t-distribution with $n-1$ degrees of freedom. The c-intervals are

$$\hat{p} \pm 1.96 \times \{S.E.(\hat{p}) + 1/2n\}\,; \tag{0.2}$$

they also widen z-intervals but do so by increasing $S.E.(\hat{p})$ with a "continuity correction."

Simon (1996) compares the achieved coverage of these systems for sample sizes 5 to 40. He concludes "the t-based interval achieves better coverage than the z-based interval," and furthermore that the c-intervals are an attractive alternative to t-intervals.

This article suggests that a third alternative to z-intervals, called q-intervals herein, should be strongly preferred in elementary courses to either t- or c-intervals. First, q-intervals are more easily motivated than z-intervals because they require only a straightforward application of the Central Limit Theorem (*without* the need to estimate the variance of $\hat{p}$ and to justify that this perturbation does not affect the normal limiting distribution). Second, the q-intervals do not involve ad-hoc continuity corrections. Third, q-intervals have substantially superior achieved coverage than either the t or the c systems.

For teachers who are interested in learning more advanced material on this subject, there are several systems of "exact" intervals that have been proposed for the binomial problem. Here *exact*

means that the systems are constructed using the binomial distribution rather than the approximate normal distribution. None of the exact methods have achieved confidence level equal to the nominal level for all true p but all are conservative in that they attain *at least* their nominal level; they differ in their degree of conservatism. The earliest system is due to Clopper and Pearson (1934), with counter proposals by Sterne/Crow, Blyth/Still, and Duffy/Santner (Sterne, 1954 as modified by Crow, 1956; Blyth and Still, 1983; Duffy and Santner, 1987). Section 2.1 of Santner and Duffy (1989) describes the proposals in a unified manner and compares them.

The comparisons presented in this section are not exhaustive but were selected to facilitate comparison with those made in the previous articles in this series. Ghosh (1979) presented a detailed study comparing several small-sample properties of z- and q-intervals using additional criteria to those presented here (but not considering t- or c-intervals). Recently, Agresti and Coull (1997) have compared z-intervals, q-intervals, Clopper-Pearson "exact" intervals, Blyth/Still intervals, and a large-sample proposal of their own by using a weighted average coverage criteria. The conclusions of both these articles concerning z- and q-intervals are consistent with given here.