Statistical Science
2005, Vol. 20, No. 4, 375–379
DOI 10.1214/088342305000000395

Comment: Fuzzy and Randomized Confidence Intervals and P-Values

Lawrence D. Brown, T. Tony Cai and Anirban DasGupta

Professor Geyer and Professor Meeden have given us an intriguing article with much material for thought and exploration, and they deserve our congratulations. Although the idea of randomized procedures has long existed, this paper has revitalized the discussion on randomized confidence intervals and randomized P-values.

Interval estimation of a binomial proportion is a very basic but very important problem with an extensive literature. Brown, Cai and DasGupta (2001) revisited this problem and showed that the performance of the standard Wald interval, which is used extensively in textbooks and in practice, is far more erratic and inadequate than is appreciated. Several natural alternative confidence intervals for p were recommended in Brown, Cai and DasGupta (2001). See also Agresti and Coull (1998). These intervals are all what the authors call crisp intervals.

The coverage probability of these crisp confidence intervals contains significant oscillation, which is intrinsic in all crisp intervals due to the lattice structure of the binomial distributions. In the present paper, Geyer and Meeden introduce the notion of fuzzy confidence intervals with the goal to eliminate oscillation and to have the exact coverage probability. The confidence intervals are obtained by inverting families of randomized tests. In addition, the authors introduce the notion of fuzzy P-values. The introduction of the critical function ϕ as a function of three variables x, α and θ provides a unified description of fuzzy decision, fuzzy confidence interval and fuzzy P-values.

Our discussion here will focus on four issues: (1) What is new in this paper?; (2) exact versus approximate coverage; (3) expected length; (4) generalization of abstract randomized confidence intervals to simultaneous inference.

Lawrence D. Brown and T. Tony Cai are Professor and Associate Professor, Department of Statistics, The Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA. Anirban DasGupta is Professor, Department of Statistics, Purdue University, West Lafayette, Indiana 47907, USA (e-mail: dasgupta@stat.purdue.edu).

1. WHAT IS NEW IN THIS PAPER?

As the authors observe, the notion of a randomized confidence interval has a long history. Such intervals are a natural consequence of the formulation of randomized tests in the Neyman–Pearson lemma, and appear in Lehmann (1959, page 81; 2nd ed., 1986, page 93), Blyth and Hutchinson (1960) and Pratt (1961). It is thus important to try to clarify which portions of the current paper are new, which represent a valuable new focus on a classical concept and which are an informative survey of key elements of that concept.

The earlier authors mentioned above, and others, realized that there are several ways to represent randomized confidence intervals. Most preferred versions in which the statistician produces a particular interval. In view of the discussion in the present Section 1.4, it appears this is what the authors would call a realized randomized interval, but some preferred what the present paper would refer to as an abstract randomized interval. For example, Lehmann (1959) created realized randomized intervals by introducing an auxiliary independent uniform random variable. Pratt (1961) created such intervals for the binomial problem by the equivalent device of constructing nonrandomized intervals on the basis of observation of $X + U$, where X is binomial and U is an independent uniform$(0, 1)$ random variable. However, the discussion in Cohen and Strawderman (1973), Brown and Cohen (1995) and Brown, Casella and Hwang (1995) is in terms of abstract randomized intervals.

From a formal mathematical perspective there seems to be nothing about the definition of abstract randomized intervals here that is different from the treatment in these earlier papers. Thus, while the descriptive language is different, the formal structure here for "fuzzy intervals" is the same as that for abstract randomized intervals. We find one feature in this new descriptive language to be very appealing: the pictorial representation in the figure near (1.2). This representation allows the user to think of abstract randomized intervals as a minor extension of ordinary ones, and helps the statistician in some circumstances to avoid the need for more precise but cumbersome statements like "the probability is 30% that I have 95% confidence in the