Journal of Modern Applied Statistical Methods

Volume 6 | Issue 2 Article 32

11-1-2007

Reply (to Ian R. White)

Kung-Jong Lui
San Diego State University, kjl@rohan.sdsu.edu

Journal of Modern Applied Statistical Methods
November, 2007, Vol. 6, No. 2, 668-669

1538 – 9472/07/$95.00

Reply

Kung-Jong Lui
San Diego State University

When evaluating the performance of an interval estimator, we generally use the coverage probability to measure the accuracy and the average length to measure the precision (Casella & Berger, 1990). An ideal interval estimator is the one which can consistently cover the underlying true parameter for all parameter values, while its average length is minimal so that one can almost pinpoint the underlying true parameter. In practice, however, such an ideal interval estimator does not exist. Note that an interval estimator, which has a high coverage probability but has a quite wide length, is of little practical value. For example, the interval estimate $(0, \infty)$ has the coverage probability of 100% covering a positive parameter, but is useless due to its length is too wide to be informative. Following the same arguments, we can easily see that the interval estimate [-1, 1] that also has the coverage probability of 100% for the difference between two proportions is also completely useless. Thus, the information on the coverage probability of an interval estimator alone is not sufficient to determine whether it can perform well or not. Given two interval estimators with the same coverage probability, the interval estimator with a shorter average length is obviously preferable to the other with a longer average length. This is because the former can allow us to draw a more precise inference. On the other hand, an interval estimator which has a short average length but has a low coverage probability is also of no practical value. These lead us to consider finding an interval estimator which has the shortest average length among all interval estimators with the coverage probability consistently larger than or equal to the desired confidence level.

Note that obtaining an interval estimate with an infinite length only suggests that the employed interval estimator based on the given data cannot provide us with an accurate estimate of the underlying parameter. This certainly does not imply that the interval estimator with an infinite length is valuable and useful. In fact, there are many problems and concerns by simply inverting the interval estimate for the risk difference to obtain an interval estimate for the number needed to treat (NNT). A systematic list of these concerns and references as well as a simple logic solution to alleviate these concerns can be found elsewhere (Lui, 2004).

It is incorrect and misleading to state that "When the parameter is transformed to a different scale, confidence interval retains their coverage properties, but not their mean length. Thus, mean length on different scales could have been considered". First, this statement about the confidence interval is generally not true unless the transformation is, for example, continuous and monotonic. The mean interval length, just like the standard error, has a unit scale. This certainly does not deter its use once when the parameter of primary interest is selected. The average length for all interval estimators will have the same unit scale as that for the parameter of interest. Thus, there will be no concern that we may compare the average length of different interval estimators at different unit scales. Note also that the relative precision is not invariant with respect to the reciprocal transformation and hence a relatively more precise interval estimate for the risk difference does not necessarily lead to produce a relatively more precise interval estimate for the NNT.

Because the sampling distribution of a statistic on which we are based to derive an interval estimator is not necessarily symmetric, we can obtain an interval estimator with the coverage probability larger than the other one, but the former also has the average length less than the latter. For example, as shown elsewhere (Lui, 2006), we can easily find the situations in which interval estimator (4) using $\tanh^{-1}(x)$ transformation has the largest coverage probability and the shortest average length among interval estimators considered in the paper. It is senseless to put a penalty on an interval estimator when its coverage probability can be even higher than the desired confidence level without sacrificing its precision. Based on the coverage probability exclusively, we indiscriminately select which interval estimator is the best can be subject to the above concern.

It is certainly desirable that test results between using hypothesis testing and various interval estimators can always be consistent with each other. If readers wish to have this property, test-based confidence intervals will be the choice. However, for given an adequately large sample size, the chance to obtain an inconsistent conclusion between hypothesis testing (in which we generally account for the null conditions when calculating the estimated variance of the test statistic) and interval estimators (in which we calculate the estimated variance of statistic without having the null conditions) should be generally small.

References

Casella, G. & Berger, R. L. (1990). *Statistical Inference*. Belmont, CA: Duxbury.

Lui, K.-J. (2004). A simple logical solution to eliminate the limitations of using the number needed to treat. *Evaluation & the Health Professions*, 27, 206-214.

Lui, K.-J. (2006). Interval estimation of risk difference in simple compliance randomized trials. *Journal of Modern Applied Statistical Methods*, 5, 395-407.