Bayesian Interval Estimation of the Difference Between Two Proportions for the Small Sample Case

Suk Hoon Lee
Chungnam National University, Department of Statistics
220 Kung-dong, Yusung-gu
Taejon, 305-764, KOREA
shlee@stat.cnu.ac.kr

Nae Hyun Park
Chungnam National University, Department of Statistics
220 Kung-dong, Yusung-gu
Taejon, 305-764, KOREA
nhpark@stat.cnu.ac.kr

Seon Woo Kim
Chungnam National University, Department of Statistics
220 Kung-dong, Yusung-gu
Taejon, 305-764, KOREA
ksw6571@hanmail.net

1. Problem Domain

Suppose a clinical trial is designed to test a new schedule of radiotherapy for treatment of brain tumors. It is known that standard dosing and schedule induce at least 60% response rate and the improved treatment is expected to have an increased response rate by 25%. One would with to construct a 95% confidence interval for the amount of increased response rate due to the new treatment. Consider a pair of independent binomial random variables X and Y with corresponding parameters p_1 and p_2 and sample sizes m and n. We put $q = p_1 - p_2$.

2. Standard Methods

The standard non-Bayesian method of constructing a 95% confidence interval for q is to use the normal approximation but this method is known to badly behave when m, n and the observations x, y are small. In particular, if $x = y = 0$ this interval fails to cover the true q unless $p_1 = p_2$ since the interval length is zero. Another standard one to overcome these difficulties uses a Bayesian approach. One of the latest results is using the non-informative prior distribution of (p_1, p_2) to get the joint posterior distribution of q and j where $j = p_1 + p_2$. The authors try to get the bounds by integrating the joint distribution using the numerical analysis and show the superiority to the conventional normal approximation approach.

3. Marginal Posterior of q

In this study we propose a more refined Bayesian method using the marginal posterior distribution of q derived from the joint distribution of (p_1, q) not that of (q, j). Here we use a nice lemma which produces the marginal distribution of q. Under some circumstances of small m, n we compare the actual coverage percentages of two confidence intervals to the confidence level and show that in most of the cases considered our proposed intervals have the actual coverage percentage closer to the specified level then those from the previous Bayesian method.

REFERENCE

[1] Yongdai Kim and Daewoo Choi, "On the Interval Estimation of the Difference Between Independent Proportions with Rare Events." The Korean Communications in Statistics, vol.7, No.2, 481-487, 2000

[2] R. G. Newcombe, "Interval Estimation for the Difference Between Independent Proportion : comparison of eleven methods" Statistics in Medicine, vol.17, 873-890, 1998

[3] Suk Hoon Lee, Nae Hyun Park, Chong Jin Park, "Bayesian Inference of Multinomial Probabilities when Some Frequency Counts are Merged", 7-th Conference of the International federation of Classification Societies, 2000

RESUME

Suk Hoon Lee and Nae Hyun Park are professors at Department of Statistics, Chungnam National University and Seon Woo Kim is a Ph. D. Student at the Same Department.