Bootstrap Estimation for the Swap-rate by Preliminary Selection in University Entrance Examination

Hirohito Sakurai
Graduate School of Engineering, Hokkaido University
Kita 13, Nishi 8, Kita-ku, Sapporo 060–8628, JAPAN
sakurai@main.eng.hokudai.ac.jp

Hisayuki Tsukuma
Graduate School of Science and Technology, Chiba University
1-33 Yayoi-cho, Inage-ku, Chiba 263-8522, JAPAN
97um0204@g.math.s.chiba-u.ac.jp

Masaaki Taguri
Faculty of Science, Chiba University
1-33 Yayoi-cho, Inage-ku, Chiba 263-8522, JAPAN
taguri@math.s.chiba-u.ac.jp

1. Introduction

In the current Japanese university entrance examination system, the selection of candidates is based on a composite score of two kinds of tests. One is the National Center Test for University Admission (First-stage Examination; FE), which is jointly administered throughout Japan, and the other is administered by each university (Second-stage Examination; SE). In some universities, FE is used for preliminary selection of candidates. However, if the preliminary selection is not carried out, some candidates may pass the entrance examination. Our interest is to decrease such probability as small as possible (see Yano *et al.* (1990)). Similar idea is presented in Kikuchi and Mayekawa (1995), where it is called by "swap-rate."

In this paper we construct three kinds of bootstrap confidence intervals for the swap-rate, and the accuracy of each confidence interval is numerically evaluated from the viewpoint of the coverage probability and the length of the confidence interval.

2. Formulation of the Problem

Let (X, Y) be a random variable, corresponding to FE and SE, following a bivariate normal distribution with parameters, $\mu_X = E[X]$, $\mu_Y = E[Y]$, $\sigma_X^2 = V[X]$, $\sigma_Y^2 = V[Y]$ and $\rho = \mathrm{cov}(X, Y)/(\sigma_X \sigma_Y)$. Suppose that we are to select N_A students out of N candidates based on the result of an entrance examination, and our interest lies on the inference of $p = \Pr[X <$

$\gamma \cap X+Y \geq \delta]$, where γ is the minimum passing mark for X, and δ is that for the composite score $X+Y$. Yano *et al.* (1990) derived the estimator $\hat{p}$ of p as $\hat{p} = L(-\hat{u}_0,\ \hat{v}_0;\ \hat{\rho}_0)$, where $L(h,k;\rho) = \int_h^\infty \int_k^\infty \phi_0(u,v;\rho)\,du\,dv$, $\phi_0(u,v;\rho) = \exp[-(u^2-2\rho uv+v^2)/\{2(1-\rho^2)\}]/(2\pi\sqrt{1-\rho^2})$, $\hat{u}_0 = (\hat{\gamma}-\hat{\mu}_X)/\hat{\sigma}_X$, $\hat{v}_0 = \{\hat{\delta}-(\hat{\mu}_X+\hat{\mu}_Y)\}/(\hat{\sigma}_X^2+2\hat{\rho}\hat{\sigma}_X\hat{\sigma}_Y+\hat{\sigma}_Y^2)^{1/2}$, $\hat{\rho}_0 = -\{(\hat{\sigma}_X/\hat{\sigma}_Y)+\hat{\rho}\}/\{(\hat{\sigma}_X/\hat{\sigma}_Y)^2+2\hat{\rho}(\hat{\sigma}_X/\hat{\sigma}_Y)+1\}^{1/2}$, $\hat{\delta} = \hat{\mu}_X+\hat{\mu}_Y+z_q(\hat{\sigma}_X^2+2\hat{\rho}\hat{\sigma}_X\hat{\sigma}_Y+\hat{\sigma}_Y^2)^{1/2}$, $q = N_A/N$, and z_q is the upper $100\,q$ percentile of $N(0,1)$, respectively. $\hat{\mu}_X$, $\hat{\mu}_Y$, $\hat{\sigma}_X$, $\hat{\sigma}_Y$, $\hat{\rho}$ and $\hat{\gamma}$ are some appropriate estimators of μ_X, μ_Y, σ_X, σ_Y, ρ and γ.

Since an analytic derivation of the sampling distribution of $\hat{p}$ is difficult, we numerically approximate it by the bootstrap. For the swap-rate, we consider the following three kinds of bootstrap confidence intervals with nominal coverage probability $1-\alpha$ based on B times resampling: $\hat{I}_1^* = [2\hat{p}-\hat{p}^*_{([B\cdot(1-\alpha/2)])},\ 2\hat{p}-\hat{p}^*_{([B\cdot\alpha/2]+1)}]$, $\hat{I}_2^* = [\hat{p}^*_{([B\cdot\alpha/2]+1)},\ \hat{p}^*_{([B\cdot(1-\alpha/2)])}]$, and $\hat{I}_3^* = [\hat{p}^*_{([B\cdot\beta_L]+1)},\ \hat{p}^*_{([B\cdot\beta_U])}]$, where $[k]$ is the integer part of k, $\hat{p}^*_{(b)}$ is the b-th order statistic of $\hat{p}_1^*,\ldots,\hat{p}_B^*$, $\beta_L = \Phi(2c+z_{\alpha/2})$, $\beta_U = \Phi(2c+z_{1-\alpha/2})$, $\Phi(\cdot)$ is the cumulative distribution function of $N(0,1)$, $c = \Phi^{-1}(\sum_{b=1}^B I\{\hat{p}_b^* \leq \hat{p}\}/B)$, and $I\{\cdot\}$ denotes the indicator function, respectively. The second and third ones, viz., $\hat{I}_2^*$ and $\hat{I}_3^*$, are called percentile and BC confidence intervals.

3. Numerical Evaluation

We evaluate the accuracy of $\hat{I}_1^*$, $\hat{I}_2^*$ and $\hat{I}_3^*$ from the viewpoint of the coverage probability and the length of them for some typical parameter values. The evaluation for the former is made by the following steps: (i) construct the bootstrap confidence interval $\hat{I}_k^*$ $(k=1,2,3)$, and (ii) repeating step (i) an appropriate number of times R, compare $\#\{p \in \hat{I}_k^*\}/R$ with the nominal coverage probability $1-\alpha$.

The computational results will be presented at the meeting.

REFERENCES

Kikuchi, K. and Mayekawa, S. (1995). On the sampling distribution of swap-rate. *Behaviormetrika*, **22**, 185–204.

Yano, K., Ouchi, S. and Taguri, M. (1990). Investigation on the multiplier of preliminary selection in entrance examination to universities. *Jpn. J. Behav.*, **17**, 25–33 (in Japanese).

RESUME

Dans le système courant du concours d'admission de l'université japonaise, la selection des candidats est basée sur un total de points des deux genres de test. La contribution du chaque test peut se mesures par la proportion de d'échange, nomé par Kikuchi et Mayekawa (1995). Dans cet article, nous construirons trois sortes des intervalles de confidence du bootstrap pour la proportion de d'échange, et aussi l'exactitude pour chaque intervalle de confidence sera numériquement evalué au point de vue de la probabilité de couverture et la longueur de l'intervalle de confidence.