Product moments of nested bootstrap distributions

Yoko Ono
Department of Management Science, Science University of Tokyo
Kagurazaka, Shinjuku, Tokyo 162-8601, JAPAN
ono@ms.kagu.sut.ac.jp
Naoto Niki
Department of Management Science, Science University of Tokyo
Kagurazaka, Shinjuku, Tokyo 162-8601, JAPAN
niki@ms.kagu.sut.ac.jp

1. Introduction

Let $\mathbf{X}^{[0]} = (\mathbf{x}_1^{[0]}, \ldots, \mathbf{x}_n^{[0]})$ be a p variate random sample of size n drawn from an unknown population $\mathbf{F}^{[0]}$. It is assumed that $\mathbf{F}^{[0]}$ has the finite moments of requisite order. Let $T(\mathbf{X}^{[0]})$ be a statistic to be used as an estimate of a population parameter θ, providing that **(a)** T is smooth enough to be expanded in a Taylor series of requisite degree; **(b)** $T(\mathbf{X}^{[0]})$ is a function of the sample moments.

Bootstrapping is known as a versatile non-parametric method for estimating the sampling distribution $G^{[0]}$ of $T(\mathbf{X}^{[0]})$ from which the bias of $T(\mathbf{X}^{[0]})$, the confidence intervals for θ and other values appearing in statistical analysis are computed. The bootstrap estimate $G^{[1]}$ of $G^{[0]}$ is the sampling distribution of $T(\mathbf{X}^{[1]})$ where each sample $\mathbf{X}^{[1]}$ is drawn from the empirical distribution $\mathbf{F}^{[1]}$ composed proportionally to $\mathbf{X}^{[0]}$.

Discussions on properties of $G^{[1]}$ have been made by many workers, including Singh (1981) and Hall (1986), where it should be noted that those are based on the condition that the sample $\mathbf{X}^{[0]}$ is given. In other words, observations in $\mathbf{X}^{[1]}$ are not truly independent but conditionally independent each other.

In order to evaluate the influence of such kind of dependence, Ono and Niki (2001) have proposed a symbolic algorithm for obtaining the sampling moments of the moments of $G^{[1]}$. Let $\Lambda = (\lambda_1, \lambda_2, \ldots)$ denote a partition of a vector

$$\mu \in \mathbf{Z}_0^p = \{j \mid j \in \mathbf{Z},\, j \geq 0\}^p,$$

such that $\mu = \lambda_1 + \lambda_2 + \ldots$. $A_\Lambda(\mathbf{X}^{[i]})$ and $P_\Lambda(\mathbf{X}^{[i]})$ are vector version of the augmented symmetric polynomial and the power sum, respectively, associated with Λ. The main part of the algorithm can be summarized as

$$\mathrm{E}^{[1]} A_\Lambda(\mathbf{X}^{[1]}) = \int A_\Lambda(\mathbf{X}^{[1]})\, d\mathbf{F}^{[1]} = P_\Lambda(\mathbf{X}^{[0]}). \tag{1}$$

It is shown, in this article, that the algorithm due to Ono and Niki (2001) is applicable also to multi-level bootstrapping, with straightforward extension and modification, for making higher order discussion.

2. Algorithm

The second level (or doubly) bootstrapping makes $G^{[2]}$, the bootstrap estimate of $G^{[1]}$, in the same manner from the empirical distribution $\mathbf{F}^{[2]}$ defined for each $\mathbf{X}^{[1]}$. Approximation to $G^{[0]}$ by using $G^{[1]}$ is approximated with the relation between two known distributions $G^{[1]}$ and $G^{[2]}$. For obtaining higher level bootstrap distribution $G^{[i]}$, we apply (**??**) to the ith level bootstrapping version;

$$\mathrm{E}^{[i]} A_\Lambda(\mathbf{X}^{[i]}) = \int A_\Lambda(\mathbf{X}^{[i]})\, d\mathbf{F}^{[i]} = P_\Lambda(\mathbf{X}^{[i-1]}). \tag{2}$$

By using (**??**) and a symbolic algorithm `PtoA` designed by Nakagawa and Niki (1991) for changing bases from $P_\Lambda(\mathbf{X}^{[i]})$ to $A_\Lambda(\mathbf{X}^{[i]})$, we can wright moments of $G^{[i]}$ in terms of population moments. The sketch of the algorithm is as follows: **(a)** Let $w \leftarrow P_\Lambda(\mathbf{X}^{[i]})$ and transform it into $A_\Lambda(\mathbf{X}^{[i]})$ by applying `PtoA`. **(b)** Take expectation for each terms in w, by using (**??**), to give $\bar{w} \leftarrow \mathrm{E}^{[i]} w$ as a linear combination of $P_\Lambda(\mathbf{X}^{[i-1]})$. **(c)** $i \leftarrow i-1$; if $i = 0$ then return to $\bar{w}$ as output else go back to **(a)**.

3. Simple Example

Assume $\mathbf{F}^{[0]}$ has finite eighth moments and let $\mathbf{V}^{[0]} = T(\mathbf{X}^{[0]})$ be the sample variance. The expectations of the second level bootstrap moments are given as follows:

$$\mathrm{E}^{[0]}\mathrm{E}^{[1]}\mathrm{E}^{[2]}\Big(\sqrt{n}(\mathbf{V}^{[2]} - \mathbf{V}^{[1]})\Big) = -\frac{\kappa_2}{\sqrt{n}} + \frac{2\kappa_2}{n\sqrt{n}} - \frac{\kappa_2}{n^2\sqrt{n}} + O(n^{-\frac{7}{2}})$$

and

$$\mathrm{E}^{[0]}\mathrm{E}^{[1]}\mathrm{E}^{[2]}\Big(n\left(\mathbf{V}^{[2]} - \mathbf{V}^{[1]}\right)^2\Big) = 2\kappa_2^2 + \kappa_4 + \frac{1}{n}\Big(-13\kappa_2^2 - 12\kappa_4\Big) + \frac{1}{n^2}\Big(62\kappa_2^2 + 72\kappa_4\Big) + O(n^{-4}),$$

where κ_i $(i = 2, 4)$ is the ith population cumulant.

REFERENCES

Hall, P. (1986). On the bootstrap and confidence intervals, *Ann. Statist.*, **14**, 1431–1452.

Nakagawa, S. and Niki, N. (1991). Change of bases of multi-system symmetric polynomials and their applications in distribution theory of multi-variate statistics (in Japanese), *Bull. Comput. Statist. Jpn.*, **4**, 35–43.

Ono, Y. and Niki, N. (2001). Symbolic algorithms for obtaining moments and "moments of moments" of bootstrap distributions, *Jornal of Japan Society for Symbolic and Algebraic Computation*, vol. **8**, No. 2, 37–48.

Singh, K. (1981). On the asymptotic accuracy of Efron's bootstrap, *Ann. Statist.*, **9**, 1187–1195.

RESUME

Nous concentrons sur des calculs symboliques pour obtenir des moments des "*nested bootstrap moments.*"