Bootstrap Methods for Judgment Post Stratification

Mozhgan Alirezaei Dizicheh[1] · Nasrollah Iranpanah[1] · Ehsan Zamanzade[1]

Abstract

It has been shown in the literature that judgment post stratification (JPS) sampling design often leads to more efficient statistical inference than what is possible to obtain in simple random sampling (SRS) design of comparable size. Since the JPS is a cost-efficient sampling design, a *large enough* sample size may not be available to use normal theory of the estimators. In this paper, we describe two bootstrap methods for JPS sampling scheme, one of which has been already used in the literature without studying its consistency and the other is new. We also show that both bootstrap approaches are consistent. We then investigate the use of the bootstrap methods for constructing confidence intervals for the population mean and compare them with the confidence interval of the population mean obtained via normal approximation (NA) method using Monte Carlo simulation. It is found that for the asymmetric distributions, one of the bootstrap methods we describe in the paper often leads to a closer coverage probability (CP) to the nominal level than NA method. Finally, a real dataset is analysed for illustration.

Keywords Confidence interval · Coverage probability · Judgment post stratified sample · Resampling methods · Simple random sample

Mathematics Subject Classification 62G09 · 62F40 · 62D05

1 Introduction

Judgment post stratification (JPS) sampling scheme was introduced by MacEachern et al. (2004) as an alternative to simple random sampling (SRS) scheme for situa-

---

✉ Ehsan Zamanzade
e.zamanzade@sci.ui.ac.ir; ehsanzamanzadeh@yahoo.com

1 Department of Statistics, Faculty of Mathematics and Statistics, University of Isfahan, Isfahan 81746-73441, Iran

tions where the variable of interest is expensive or difficult to measure, but judgment ranking is cheap and can be done relatively easily. These situations happen in many practical research including educational studies (Wang et al. 2016), sport sciences (He et al. 2018; Qian et al. 2019) and medical studies (Mahdizadeh and Zamanzade 2019; Zamanzade and Mahdizadeh 2020). A JPS sample is constructed using a simple random sample and its supplementary judgment ranks. The JPS sampling procedure is shortly described as follows.

To draw a JPS sample of size n using set size m from an infinite population, a SRS sample of size n is collected and measured. Then for each measured unit in the simple random sample, $m - 1$ supplemental units are identified from the same population to create a set of size m. Finally, the judgment rank of measured unit in the simple random sample among $m - 1$ supplemental units is noted. Thus, a JPS sample contains a SRS sample plus its preparatory judgment ranks and can be represented as $(Y_1, R_1), \ldots, (Y_n, R_n)$, where R_i is the judgment rank of Y_i among $m - 1$ supplemental units in the set of size m. Throughout of this paper, the term of *judgment rank* is used to point out that the rank of R_i is obtained using an inexpensive method such as visual inspection, covariate or an expert's personal judgment which does not require referring to exact values of supplemental units in the set and thus it may not to be accurate and contains error. If the rank is obtained without error (perfect ranking), then the conditional distribution of Y_i given its rank $R_i = r$ is the same as the rth order statistic of a sample of size m. Otherwise, the conditional distribution follows the rth judgment order statistic of a sample of size m. A ranking process is called consistent if

$$F(t) = \frac{1}{m} \sum_{r=1}^{m} F_{[r]}(t),$$

where $F(t)$ and $F_{[r]}(t)$ are cumulative distribution functions (CDFs) of the parent distribution and rth judgment order statistic of a sample of size m, evaluated at the point t, respectively (see Presnell and Bohn 1999).

Let n_r be the number of JPS units with judgment rank r, then one can simply show that under a consistent ranking process, the vector $\mathbf{n} = (n_1, \ldots, n_m)$ follows a multinomial distribution with mass parameter n and probability vector $\left(\frac{1}{m}, \ldots, \frac{1}{m}\right)$.

Conditioning on the vector of the ranks $\mathbf{R} = (R_1, \ldots, R_n)$, a JPS sample can be regarded as an unbalanced ranked set sample (RSS). To draw a RSS sample, one first determines a set size m and a vector of post strata sample sizes $\mathbf{n} = (n_1, \ldots, n_m)$, so that $n = \sum_{r=1}^{m} n_r$ is the total sample size. For $r = 1, \ldots, m$, n_r samples of size m are identified from the population of interest, ranked in an increasing magnitude, and the unit with rank r is selected for actual measurement.

Although JPS and RSS are in close connection with each other, two obvious differences between them can be found from their definitions. The first one is related to how judgment ranks are connected to the sample units. Note that in JPS (RSS) setting, the ranks are assigned to each sample unit after (before) its measurement, and therefore they are loosely (strongly) attached to the sample units and can (cannot) be ignored. This means that a JPS sample has more variability than its RSS counterpart, yet more

flexibility to be used in practice since the standard SRS techniques can still be used for a JPS sample when the ranks are ignored. The second difference between JPS and RSS is related to the vector of post strata sample sizes $\mathbf{n} = (n_1, \ldots, n_m)$. While $\mathbf{n}$ is a fixed and pre-specified vector in RSS, it is a random vector in JPS and follows a multinomial distribution.

The JPS has been the subject of many studies since its introduction, including estimation of the population mean (Wang et al. 2008; Frey and Feeman 2012; Dastbaravarde et al. 2016; Frey 2016), estimation of the population variance (Frey and Feeman 2013; Zamanzade and Vock 2015; Zamanzade 2016), estimation of the cumulative distribution function (CDF; Frey and Ozturk 2011; Wang et al. 2012; Duembgen and Zamanzade 2020), estimation of the population proportion (Zamanzade and Wang 2017), estimation of the population quantile (Ozturk 2014), two-sample problems (Ozturk 2015; Dastbaravarde and Zamanzade 2020), finite mixture model analysis (Omidvar et al. 2018), finite sample size corrections (Ozturk 2016) and perfect ranking test (Zamanzade and Vock 2018).

It is shown in the most of the above literature that JPS sampling design provides more efficient statistical inference than what is possible in SRS design of comparable size provided that the sample size is not too small and quality of ranking is fairly good (Dastbaravarde et al. 2016). Most of the JPS literature focus on drawing statistical inference of the population characteristics based on normal approximation (NA) method. However, since the JPS is applicable for situations in which obtaining exact values of sample units is hard/expensive, a *large enough* sample size may not be available to obtain a valid asymptotic inference due to time/cost considerations. So, there is a need of existing an alternative method to draw statistical inference for the population characteristics based on a JPS sample.

Bootstrap approach proposed by Efron (1979), is a computer-based method that repeats a simple operation so many times. This method enjoys from the property that it makes few distributional assumptions and provides solutions to many standard statistical problems. Specifically, the bootstrap approach can be used to estimate CDF and constructing a confidence interval for the population characteristics without making any distributional assumptions.

This paper describes two bootstrap approaches for a JPS sample, one of which has been already used by Ozturk (2016) without studying its consistency and the other is new. In Sect. 2, two bootstrap methods based on JPS sample scheme are introduced and their asymptotic consistencies are established in Sect. 3. In Sect. 4, the bootstrap techniques are used to construct confidence intervals for the population mean based on a JPS sample. The bootstrap confidence intervals we developed in this section are then compared with the confidence interval based on NA using Monte Carlo simulation. In Sect. 5, a real dataset is used to show the applicability and efficiency of the introduced methods in practice. Some concluding remarks and directions for future research are provided in Sect. 6.

2 Bootstrap Methods for a Judgment Post Stratified Sample

Let $(Y_1, R_1), \ldots, (Y_n, R_n)$, be a JPS sample of size n from a population with CDF F, where R_i is the judgment rank of Y_i among $m - 1$ supplemental units in the set of size m. Let n_r be the number of the JPS sample units with judgment rank r, then it is easy to show that under a consistent ranking process, the vector $(n_1, \ldots, n_m)$ follows a multinomial distribution with mass parameter n and probability vector $\left(\frac{1}{m}, \ldots, \frac{1}{m}\right)$. As it is mentioned earlier, the conditional distribution of Y_i given its judgment rank $R_i = r$ is given by $F_{[r]}(t) = P(Y_i \leq t | R_i = r) = P(Y_{[r]} \leq t)$, where $Y_{[r]}$ is the rth judgment order statistic of a sample of size m. Here, the square bracket [.] is used to indicate imperfect rankings. If the rankings are perfect, then the square brackets are replaced with the round ones and $F_{(r)}(t) = B(F(t), r, m + 1 - r)$, where $B(t, \alpha, \beta)$ is the CDF of the beta distribution with parameters α and β, evaluated at the point t. The empirical estimator of $F_{[r]}(t)$ based on a JPS sample is given by

$$F_{[r], n_r}(t) = \left(\frac{1}{n_r} \sum_{i=1}^{n_r} \mathbb{I}(Y_i \leq t) \mathbb{I}(R_i = r)\right) \mathbb{I}(n_r > 0), \tag{1}$$

where $\mathbb{I}(.)$ is the indicator function. The empirical estimator of CDF based on a JPS sample is then given by

$$F_{n, jps}(t) = \frac{1}{d_m} \sum_{r=1}^{m} F_{[r], n_r}(t), \tag{2}$$

where $d_m = \sum_{r=1}^{m} \mathbb{I}(n_r > 0)$.

The standard mean estimator in the JPS setting can be defined in a similar fashion as follows

$$\hat{\mu}_{jps} = \frac{1}{d_m} \sum_{r=1}^{m} \bar{Y}_{[r]}, \tag{3}$$

where

$$\bar{Y}_{[r]} = \left(\frac{1}{n_r} \sum_{i=1}^{n} Y_i \mathbb{I}(R_i = r)\right) \mathbb{I}(n_r > 0), \tag{4}$$

is the mean of sample units with judgment rank r. Dastbaravarde et al. (2016) examined the finite sample size and asymptotic properties of $\hat{\mu}_{jps}$. They showed that this estimator is unbiased regardless of issue of rankings. Dastbaravarde et al. (2016) also proved that if the sample size n goes to infinity then $\sqrt{n}(\hat{\mu}_{jps} - \mu)$ converges in distribution to a mean zero normal distribution with variance $\sigma_{jps}^2 = \frac{1}{m} \sum_{r=1}^{m} \sigma_{[r]}^2$, where $\sigma_{[r]}^2$ is the variance of the rth judgment order statistic in a set of size m. One can simply show that $\sigma_{jps}^2 \leq \sigma^2$, and therefore $\hat{\mu}_{jps}$ is at least as asymptotically efficient as the standard mean estimator in SRS.

In what follows, we develop two algorithms to draw a bootstrap sample from a JPS sample. The first algorithm is called simple bootstrap JPS (SBJPS) which is based on drawing simple random samples of JPS units with replacement from the original JPS sample. In the second algorithm, the JPS sample units are partitioned into different strata based on their judgment ranks and then a simple random sample with replacement is drawn from each stratum. The second algorithm is called bootstrap JPS by stratum (BJPSS). These two algorithms are delineated as follows.

2.1 SBJPS: Simple Bootstrap JPS

The simple bootstrap JPS algorithm is based on the fact that the pairs of (Y_i, R_i), $i = 1, \ldots, n$ in the JPS sample, are identically and independently distributed. Thus, the original bootstrap algorithm can be applied on the pairs (Y_i, R_i). This technique was firstly used by Ozturk (2016), and here, we describe it in detail and establish its asymptotic properties.

Algorithm 1: SBJPS

1. Assign probability $\frac{1}{n}$ to each pair of $(Y_1, R_1), \ldots, (Y_n, R_n)$.
2. Randomly draw n pairs with replacement from $(Y_1, R_1), \ldots, (Y_n, R_n)$ to obtain $(Y_1^*, R_1^*), \ldots, (Y_n^*, R_n^*)$.
3. Define the bootstrap empirical distribution function as

$$F_{n,jps}^*(t) = \frac{1}{d_m^*} \sum_{r=1}^{m} F_{[r],n_r}^*(t),$$

where

$$F_{[r],n_r}^*(t) = \left(\frac{1}{n_r^*} \sum_{j=1}^{n_r^*} \mathbb{I}\left(Y_i^* \le t\right) \mathbb{I}\left(R_i^* = r\right) \right) \mathbb{I}\left(n_r^* > 0\right),$$

$d_m^* = \sum_{r=1}^{m} \mathbb{I}\left(n_r^* > 0\right)$ and n_r^* is the number of bootstrap sample units with judgment rank r (for $r = 1, \ldots, m$).
4. The bootstrap estimate of the parameter $\theta = g(F)$, for an arbitrary function of $g(.)$, is then obtained as $\hat{\theta}^* = g\left(F_{n,jps}^*\right)$.
5. Repeat steps 1–4 for B times to obtain bootstrap sample $\left(\hat{\theta}_1^*, \ldots, \hat{\theta}_B^*\right)$.

2.2 BJPSS : Bootstrap JPS by Stratum

This bootstrap algorithm is based on the idea that artificial post-strata can be constructed based on the judgment ranks in a JPS sample. Thus, independent bootstrap samples can be first drawn from different post-strata and then combined to obtain the final bootstrap sample. Let $(Y_1, R_1), \ldots, (Y_n, R_n)$, be a JPS sample of size n from a population with CDF F. Let $\mathcal{Y}_r = \{(Y_1, R_1), \ldots, (Y_n, R_n); Y_i | R_i = r; i = 1, \ldots, n\}$

be the sample units with judgment rank r (for $r = 1, \ldots, m$). Then for a fixed r, the sample units in $\mathcal{Y}_r$ are independent and identical random variables. Thus, the following bootstrap algorithm is proposed.

Algorithm 2: BJPSS

1. If $n_r > 0$, then assign probability $\frac{1}{n_r}$ to each element of $\mathcal{Y}_r$ and randomly draw n_r pairs of (Y_i^*, R_i^*) with replacement from $\mathcal{Y}_r$ to construct $\mathcal{Y}_r^*$.
2. Repeat the step (1) for $r = 1, \ldots, m$.
3. The bootstrap sample is then obtained as $\cup_{r=1}^{m} \mathcal{Y}_r^*$.
4. The bootstrap empirical distribution function and estimate of the parameter $\theta = g(F)$ are then defined in a similar fashion as in algorithm 1.
5. Repeat steps 1–4 for B times to obtain bootstrap sample $\left(\hat{\theta}_1^*, \ldots, \hat{\theta}_B^* \right)$.

Finally, bootstrap statistical inference can be made based on the bootstrap sample $\left(\hat{\theta}_1^*, \ldots, \hat{\theta}_B^* \right)$ obtained from each of the algorithms defined above. For example, bootstrap variance estimate of $\hat{\theta}$ is obtained from $\widehat{Var}_B \left(\hat{\theta} \right) = \frac{1}{B-1} \sum_{b=1}^{B} \left(\hat{\theta}_b^* - \bar{\hat{\theta}}^* \right)$, where $\bar{\hat{\theta}}^* = \frac{1}{B} \sum_{b=1}^{B} \hat{\theta}_b^*$.

3 Asymptotic Results

In this section, we establish the consistency of the bootstrap algorithms we described in Sect. 2 using Mallows metric which was firstly used by Bickel and Freedman (1981) in the context of the bootstrap technique. The definition of Mallows metric is given below.

Definition 1 (*Mallows metric*) Let Γ_2 be the set of CDFs having finite second moments. Let X and Y be random variables with CDFs $G, H \in \Gamma_2$, respectively, and define $\rho_2(G, H) = \inf_{\tau_{X,Y}} \mathbb{E}^{1/2}(|X - Y|^2)$ where $\tau_{X,Y}$ is the set of all possible joint distributions of the pair (X, Y) whose marginal CDFs are G and H, respectively.

Definition 2 (*Concept of consistency for bootstrap estimators*) Let $T_n = T(X_1, X_2, \ldots, X_n)$ be a statistic based on a random sample $X_1, \ldots, X_n$ and T_n^* be its corresponding bootstrap replicate. Then the bootstrap procedure is strongly consistent under ρ_2 for T if $\rho_2(H_n, H_n^*) \xrightarrow{a.s.} 0$, where H_n is the sampling distribution of T_n and H_n^* is the sampling distribution of T_n^* and ρ_2 is a metric on the space of CDFs.

In order to obtain asymptotic results for our bootstrap algorithms, the following lemmas are recalled from Bickel and Freedman (1981).

Lemma 1 *Let $G_n, G \in \Gamma_2$. Then $\rho_2(G_n, G) \xrightarrow{a.s.} 0$ as $n \to \infty$ is equivalent to*

$$G_n \to G \quad weakly \ and \quad \int t^2 dG_n(t) \xrightarrow{a.s.} \int t^2 dG(t).$$

Lemma 2 *Suppose the U_j are independent, likewise for V_j. Assume that their corresponding CDFs are in Γ_2 and $\mathbb{E}(U_j) = \mathbb{E}(V_j)$. Then*

$$\rho_2 \left(\sum_j U_j, \sum_j V_j \right)^2 \leq \sum_j \rho_2(U_j, V_j)^2.$$

Lemma 3 *Let the CDFs of U, V are in Γ_2. Then*

$$\rho_2(U, V)^2 = \rho_2[U - \mathbb{E}(U), V - \mathbb{E}(V)]^2 + |\mathbb{E}(U) - \mathbb{E}(V)|^2.$$

The first theorem in this section shows that the empirical estimator of $F_{[r]}(t)$ based on a JPS sample is consistent.

Theorem 3.1 *Let $(Y_1, R_1), \ldots, (Y_n, R_n)$, be a JPS sample of size n from a population with CDF $F \in \Gamma_2$ which is obtained under a consistent ranking process. Let $F_{[r],n_r}(t)$ be the empirical estimator of $F_{[r]}(t)$ as given in Eq. (1). Then $\rho_2(F_{[r],n_r}, F_{[r]}) \xrightarrow{a.s.} 0$ as the sample size n goes to infinity.*

Proof Let $\mathcal{Y}_r = \{(Y_1, R_1), \ldots, (Y_n, R_n) ; Y_i | R_i = r; i = 1, \ldots, n\}$ be the sample units with judgment rank r (for $r = 1, \ldots, m$). Then for a fixed value of r and under a consistent ranking assumption, the sample units in $\mathcal{Y}_r$ are independent and identically distributed with CDF $F_{[r]}$. A straightforward application of Glivenko–Cantelli theorem indicates that

$$\sup_t \left| F_{[r],n_r}(t) - F_{[r]}(t) \right| \xrightarrow{a.s.} 0 \quad as \; n_r \to \infty \; ; \; \forall r.$$

If $F \in \Gamma_2$, then $F_{[r]} \in \Gamma_2$. Thus, it follows from strong law of large numbers (SLLN) that:

$$\int t^2 dF_{[r],n_r}(t) = \frac{1}{n_r} \sum_{j=1}^{n_r} X_{[r]j}^2 \xrightarrow{a.s.} \int t^2 dF_{[r]}(t).$$

Therefore by Lemma 1, $\rho_2 \left(F_{[r],n_r}, F_{[r]} \right) \xrightarrow{a.s.} 0$ as $n_r \to \infty$ for $r = 1, \ldots, m$. □

Theorem 3.2 *Let $(Y_1, R_1), \ldots, (Y_n, R_n)$, be a JPS sample of size n from a population with CDF $F \in \Gamma_2$ which is obtained under a consistent ranking process. Let $F_{n,jps}(t)$ be the empirical estimator of $F(t)$ based on the JPS sample as given in Eq. (2). Then $\rho_2(F_{n,jps}, F) \xrightarrow{a.s.} 0$ as $n \to \infty$.*

Proof Note that under a consistent ranking process assumption, we can write

$$F(t) = \frac{1}{m} \sum_{r=1}^{m} F_{[r]}(t).$$

Therefore, we have

$$\sup_t |F_n(t) - F(t)| = \sup_t \left| \frac{1}{d_m} \sum_{r=1}^{m} F_{[r],n_r}(t) - \frac{1}{m} \sum_{r=1}^{m} F_{[r]}(t) \right|$$

$$\leq \sum_{r=1}^{m} \sup_t \left| \frac{1}{d_m} F_{[r],n_r}(t) - \frac{1}{m} F_{[r]}(t) \right|.$$

Besides

$$\sup_t \left| \frac{1}{d_m} F_{[r],n_r}(t) - \frac{1}{m} F_{[r]}(t) \right|$$

$$\leq \left| \frac{1}{d_m} - \frac{1}{m} \right| F_{[r],n_r}(t) + \frac{1}{m} \sup_t \left| F_{[r],n_r}(t) - F_{[r]}(t) \right| = o(1).$$

Therefore $\sup_t \left| F_n(t) - F(t) \right| = o(1)$. If $F \in \Gamma_2$ then $F_{[r]} \in \Gamma_2$. Thus, by SLLN we have:

$$\int t^2 dF_n(t) \xrightarrow{\text{a.s.}} \sum_{r=1}^{m} \frac{1}{m} \int t^2 dF_{[r]}(t) = \int t^2 d\left(\frac{1}{m} \sum_{r=1}^{m} F_{[r]}(t) \right) = \int t^2 dF(t).$$

Then by Lemma 1, $\rho_2(F_n, F) \xrightarrow{\text{a.s.}} 0$ as $n \to \infty$. $\qquad\qquad\square$

Theorem 3.3 *Let* $\left(Y_1^*, R_1^*\right), \ldots, \left(Y_n^*, R_n^*\right)$, *be a SBJPS sample of size n based on a JPS sample of the same size from a population with CDF* $F \in \Gamma_2$ *which is obtained under a consistent ranking process. Then* $\rho_2(F_{[r],n_r^*}^*, F_{[r],n_r}) \xrightarrow{\text{a.s.}} 0$ *as* $n_r \to \infty$ *for* $r = 1, \ldots, m$, *where* $F_{[r],n_r}^*$ *is the empirical estimator of* $F_{[r],n}$ *based on the SBJPS sample.*

Proof Let $\mathcal{Y}_r^* = \left\{ \left(Y_1^*, R_1^*\right), \ldots, \left(Y_n^*, R_n^*\right); Y_i^* | R_i^* = r; i = 1, \ldots, n \right\}$ be the bootstrap units with judgment rank r (for $r = 1, \ldots, m$). Note that the units in $\mathcal{Y}_r^*$ are independent and identically distributed with CDF $F_{[r]}$. Thus it follows from Glivenko–Cantelli theorem that

$$\sup_t \left| F_{[r],n_r^*}^*(t) - F_{[r],n_r}(t) \right| \xrightarrow{\text{a.s.}} 0 \quad \text{as } n_r \to \infty; \quad \forall r.$$

If $F \in \Gamma_2$, then $F_{[r]} \in \Gamma_2$. By SLLN, we have

$$\int t^2 dF_{[r],n_r^*}^*(t) \xrightarrow{\text{a.s.}} \int t^2 dF_{[r],n_r}(t).$$

Therefore by Lemma 1, $\rho_2\left(F_{[r],n_r^*}^*, F_{[r],n_r} \right) \xrightarrow{\text{a.s.}} 0$ as $n_r \to \infty$ for $r = 1, 2, \ldots, m$.
$\square$

Theorem 3.4 *Let $\left(Y_1^*, R_1^*\right), \ldots, \left(Y_n^*, R_n^*\right)$, be a SBJPS sample of size n based on a JPS sample of the same size from a population with CDF $F \in \Gamma_2$ which is obtained under a consistent ranking process. Then $\rho_2(F_{n,jps}^*, F_{n,jps}) \xrightarrow{a.s.} 0$ as $n \to \infty$, where $F_{n,jps}^*$ is the empirical estimator of F based on the SBJPS sample.*

Proof

$$
\sup_t \left| F_{n,jps}^*(t) - F_{n,jps}(t) \right| = \sup_t \left| \sum_{r=1}^m \frac{1}{d_m^*} F_{[r],n_r^*}^*(t) - \frac{1}{m} \sum_{r=1}^m F_{[r],n_r}(t) \right|
$$

$$
\leq \sum_{r=1}^m \sup_t \left| \frac{1}{d_m^*} F_{[r],n_r^*}^*(t) - \frac{1}{m} F_{[r],n_r}(t) \right| = o(1).
$$

If $F \in \Gamma_2$ then $F_{[r]} \in \Gamma_2$. Thus, by SLLN we have

$$
\int t^2 dF_n^*(t) \xrightarrow{a.s.} \sum_{r=1}^m \frac{1}{m} \int t^2 dF_{[r],n_r}(t) = \int t^2 dF_n(t)
$$

and this completes the proof. $\qquad\square$

Now, we are ready to present the main results. To do so, we only present the results for SBJPS algorithm and the results for BJPSS can be obtained in a similar fashion.

Theorem 3.5 *Let $\left(Y_1^*, R_1^*\right), \ldots, \left(Y_n^*, R_n^*\right)$, be a SBJPS sample of size n based on a JPS sample of the same size from a population with CDF $F \in \Gamma_2$ which is obtained under a consistent ranking process. Then $\rho_2(F_{n,jps}^*, F) \xrightarrow{a.s.} 0$ as $n \to \infty$, where $F_{n,jps}^*$ is the empirical estimator of F based on the SBJPS sample.*

Proof

$$
\sup_t \left| F_{n,jps}^*(t) - F(t) \right| \leq \sup_t \left| F_{n,jps}^*(t) - F_{n,jps}(t) \right| + \sup_t \left| F_{n,jps}(t) - F(t) \right| = o(1).
$$

If $F \in \Gamma_2$ then $F_{[r]} \in \Gamma_2$. So, SLLN indicates that

$$
\int t^2 dF_{n,jps}^*(t) \xrightarrow{a.s.} \sum_{r=1}^m \frac{1}{m} \int t^2 dF_{[r]}(t)
$$

$$
= \int t^2 d\left(\frac{1}{m} \sum_{r=1}^m F_{[r]}(t) \right) = \int t^2 dF(t),
$$

and this completes the proof. $\qquad\square$

Theorem 3.6 *Let $\left(Y_1^*, R_1^*\right), \ldots, \left(Y_n^*, R_n^*\right)$, be a SBJPS sample of size n based on a JPS sample of the same size from a population with CDF $F \in \Gamma_2$ which is obtained under a consistent ranking process. Define $T_n = \sqrt{n}(\hat{\mu}_{jps} - \mu)$ and $T_n^* = \sqrt{n}(\hat{\mu}_{jps}^* - \hat{\mu}_{jps})$, then $\rho_2(H_n^*, H_n) \xrightarrow{a.s.} 0$ as $n \to \infty$ where H_n and H_n^* are the sampling distributions of T_n and T_n^*, respectively.*

Proof

$$\rho_2\left(H_n^*, H_n\right) = \rho_2\left(\sqrt{n}\left(\hat{\mu}_{jps}^* - \hat{\mu}_{jps}\right), \sqrt{n}\left(\hat{\mu}_{jps} - \mu\right)\right)$$

$$= \rho_2\left(\sqrt{n}\sum_{r=1}^{m}\left(\frac{1}{d_m^*}\bar{X}_{[r]}^* - \frac{1}{d_m}\bar{X}_{[r]}\right), \sqrt{n}\sum_{r=1}^{m}\left(\frac{1}{d_m}\bar{X}_{[r]} - \frac{1}{m}\mu_{[r]}\right)\right)$$

$$= \rho_2\left(\sum_{r=1}^{m}\sqrt{\frac{n}{n_r}}\sqrt{n_r}\left(\frac{1}{d_m^*}\bar{X}_{[r]}^* - \frac{1}{d_m}\bar{X}_{[r]}\right),\right.$$

$$\left.\sum_{r=1}^{m}\sqrt{\frac{n}{n_r}}\sqrt{n_r}\left(\frac{1}{d_m}\bar{X}_{[r]} - \frac{1}{m}\mu_{[r]}\right)\right)$$

$$\xrightarrow{a.s.} \rho_2\left(\sum_{r=1}^{m}\sqrt{\frac{n_r}{m}}\left(\bar{X}_{[r]}^* - \bar{X}_{[r]}\right), \sum_{r=1}^{m}\sqrt{\frac{n_r}{m}}\left(\bar{X}_{[r]} - \mu_{[r]}\right)\right)$$

$$\leq \frac{1}{\sqrt{m}}\sqrt{\sum_{r=1}^{m}\rho_2\left(\sqrt{n_r}\left(\bar{X}_{[r]}^* - \bar{X}_{[r]}\right), \sqrt{n_r}\left(\bar{X}_{[r]} - \mu_{[r]}\right)\right)}$$

$$\leq \frac{1}{\sqrt{m}}\sqrt{m}\sqrt{\sup_r\left[\rho_2\left(\sqrt{n_r}\left(\bar{X}_{[r]}^* - \bar{X}_{[r]}\right), \sqrt{n_r}\left(\bar{X}_{[r]} - \mu_{[r]}\right)\right)\right]} = o(1).$$

Because

$$\rho_2\left(\sqrt{n_r}\left(\bar{X}_{[r]}^* - \bar{X}_{[r]}\right), \sqrt{n_r}\left(\bar{X}_{[r]} - \mu_{[r]}\right)\right)^2$$

$$= \frac{1}{n_r}\rho_2\left(\sum_{j=1}^{n_r}\left(X_{[r]j}^* - \bar{X}_{[r]}\right), \sum_{j=1}^{n_r}\left(X_{[r]j} - \mu_{[r]}\right)\right)^2$$

$$\leq \frac{1}{n_r}\sum_{j=1}^{n_r}\rho_2\left(X_{[r]j}^* - \bar{X}_{[r]}, X_{[r]j} - \mu_{[r]}\right)^2$$

$$= \rho_2\left(X_{[r]1}^* - \bar{X}_{[r]}, X_{[r]1} - \mu_{[r]}\right)^2$$

$$= \rho_2\left(X_{[r]1}^*, X_{[r]1}\right)^2 - \left|\bar{X}_{[r]} - \mu_{[r]}\right|^2$$

$$= \rho_2\left(F_{[r],n_r^*}^*, F_{[r],n_r}\right)^2 - \left|\bar{X}_{[r]} - \mu_{[r]}\right|^2$$

$$= o(1).$$

$\square$

In the next section, we will use the above bootstrap methods for constructing confidence intervals of the population mean.

4 Monte Carlo Simulation

In this section, we compare the performance of bootstrap expanded percentile confidence intervals based on the standard mean estimator in JPS with confidence interval using NA technique. In the NA method, we used the fact that $\hat{\mu}_{jps}$ is an unbiased estimator of the population mean and $\sqrt{n}\left(\hat{\mu}_{jps} - \mu\right)$ converges in distribution to a mean zero normal distribution with variance σ_{jps}^2 as the sample size n tends to infinity. An unbiased estimator $\hat{\sigma}_{jps}^2$ for $\mathbb{V}(\hat{\mu}_{jps})$ is given by Ozturk (2016, Eq. 2.5 in Theorem 2). Based on this estimator, the $100\left(1 - \alpha\right)\%$ asymptotic confidence interval for the population mean due to Ozturk (2016), is given by

$$\left(\hat{\mu}_{jps} - t_{1-\alpha/2,n-1}\sqrt{\frac{\hat{\sigma}_{jps}^2}{n}}, \ \hat{\mu}_{jps} + t_{1-\alpha/2,n-1}\sqrt{\frac{\hat{\sigma}_{jps}^2}{n}}\right),$$

where $t_{p,v}$ is the pth quantile of the t-distribution with v degrees of freedom.

There are several methods for constructing a confidence interval using bootstrap approach. Here, we use bootstrap expanded percentile confidence interval since it is obtained using the similar adjustments to the NA confidence interval described above for the bootstrap setting and it often has better performance than percentile confidence interval (see for example, Hesterberg 2015).

The $100\left(1 - \alpha\right)\%$ expanded percentile confidence interval for the population mean using bootstrap approach can be constructed as follows. First, let $\alpha'/2 = \Phi\left(\sqrt{n/\left(n-1\right)}t_{\alpha/2,n-1}\right)$, where $\Phi\left(.\right)$ is the CDF of the standard normal distribution. Then, draw B bootstrap samples of size n from original JPS sample using each of the bootstrap methods introduced in Sect. 2. Next, obtain the estimate of the population mean using the bootstrap sample, $\left(\hat{\mu}_{jps,1}^*, \ldots, \hat{\mu}_{jps,B}^*\right)$. Finally, the $100\left(1 - \alpha\right)\%$ bootstrap expanded percentile confidence interval is obtained as $\left(\hat{\mu}_{jps}^{*,\alpha'/2}, \hat{\mu}_{jps}^{*,1-\alpha'/2}\right)$, where $\hat{\mu}_{jps}^{*,p}$ is the pth quantile of the bootstrap sample $\left(\hat{\mu}_{jps,1}^*, \ldots, \hat{\mu}_{jps,B}^*\right)$.

To generate a JPS sample, we assume that the ranking in each set of size m is done using perceptual linear ranking model due to Dell and Clutter (1972). In this model, it is assumed that in each set of size m, the actual rank of the concomitant variable Y is assigned as the judgment rank of the interest variable X, where the following relation between X and Y holds

$$Y = \lambda\left(\frac{X - \mu_x}{\sigma_x}\right) + \sqrt{1 - \lambda^2}Z,$$

where μ_x and σ_x are mean and standard deviation of X, respectively, Z is independent from X and follows a standard normal distribution and the parameter λ is the correlation coefficient between X and Y which controls the quality of ranking.

We set $n \in \{10, 20, 30, 50, 100\}$, $m \in \{3, 4, 5\}$, $\lambda \in \{0.5, 0.7, 0.9, 1\}$ and for each combination of (n, m, λ), we have generated 10,000 JPS random samples from three symmetric distributions, i.e. standard normal distribution ($N\left(0, 1\right)$), standard uniform
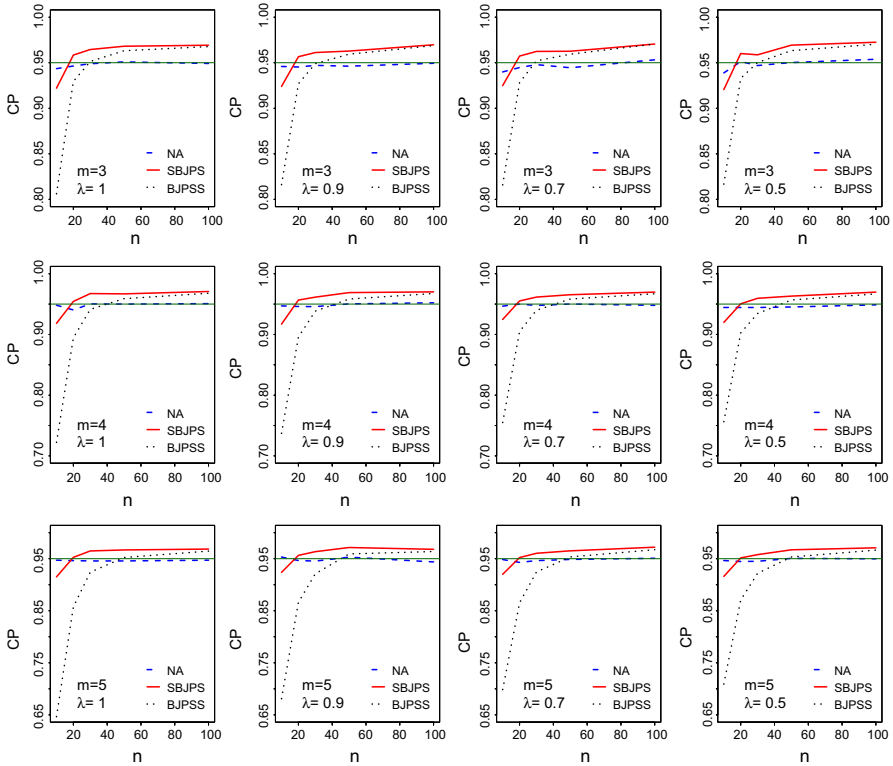
Fig. 1 Estimated coverage probability (CP) of different confidence intervals based on NA (represented by two dash line, and blue color), SBJPS (represented by solid line and red color) and BJPSS (represented by dotted line and black color) methods as a function of sample size n for $m \in \{3, 4, 5\}$, $\lambda \in \{0.5, 0.7, 0.9, 1\}$ when the parent distribution is $N(0, 1)$. This figure appears in color in the electronic version of this paper. (Color figure online)

distribution ($U(0, 1)$), beta distribution with parameters 0.5, and 0.5 ($B(0.5, 0.5)$) and three asymmetric distributions, i.e. standard log-normal distribution ($LN(0, 1)$), Gamma distribution with scale parameter 1 and shape parameter 0.5 ($G(0.5)$) and Weibull distribution with scale parameter 1 and shape parameter 0.5 ($W(0.5)$). We have then estimated the coverage probability (CP) of bootstrap and NA 95% confidence intervals using the simulated samples. Also, the bootstrap size is taken to be $B = 1000$. For brevity, we only present the results for standard normal and standard log-normal distributions in Figs. 1 and 2, respectively, in the paper. The results for other distributions can be found in Figs. S1, S2, S3 and S4 in the Supplementary Material.

Figure 1 presents the simulation results when the parent distribution is standard normal. We observe from this figure that the NA confidence interval provides the nearest CP to its nominal level 95% in all considered cases, which is not surprising since the NA method should be the best technique when the parent distribution is truly normal. It should be mentioned that the CP of NA confidence interval does not change much with sample size n. The confidence interval based on SBJPS method
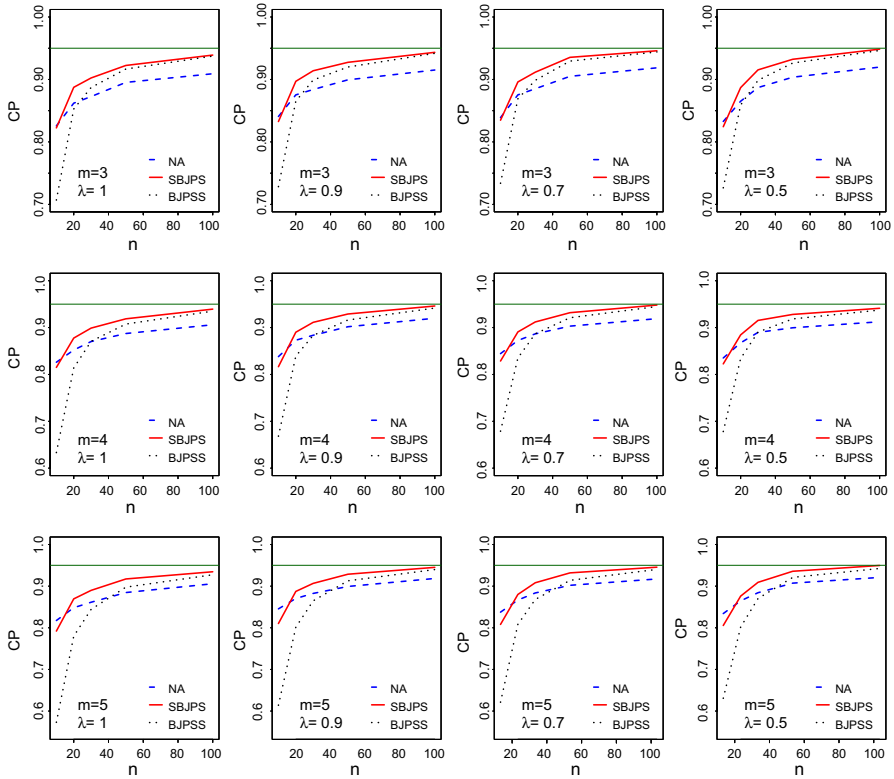
Fig. 2 Estimated coverage probability (CP) of different confidence intervals based on NA (represented by two dash line, and blue color), SBJPS (represented by solid line and red color) and BJPSS (represented by dotted line and black color) methods as a function of sample size n for $m \in \{3, 4, 5\}$, $\lambda \in \{0.5, 0.7, 0.9, 1\}$ when the parent distribution is $LN(0, 1)$. This figure appears in color in the electronic version of this paper. (Color figure online)

provides slightly lower (higher) CP to its nominal level for $n = 10$ (≥ 20). The CP of SBJPS confidence interval increases when sample size goes from $n = 10$ to 20 and levels out for $n \geq 20$. The confidence interval based on BJPSS method has the lowest CP for $n \leq 30$ and the difference between its CP and the nominal level 95% is more pronounced for $n = 10$. This can be justified by the nature of BJPSS algorithm. Note that BJPSS forces the number of units with each judgment rank to be fixed. Since the number of units with a particular rank is actually random in JPS, BJPSS underestimates the amount of variability in the JPS mean estimate. It is worth mentioning that the CP of BJPSS confidence interval increases with set sample size n, for $n \leq 50$ and it remains almost unchanged for $n > 50$. It is also easy to see the the set size m and the quality of ranking λ do not have much effect on CPs of different confidence intervals.

Comparing Fig. 1 with Figs. S1 and S2 in the Supplementary Material, we observe that the patterns of the CPs of different confidence intervals for standard normal distribution remain almost the same as $U(0, 1)$ and $B(0.5, 0.5)$ distributions.

Simulation results for standard log-normal distribution are given in Fig. 2. It is clear from this figure that CPs of all considered confidence intervals are lower than the nominal level, yet they converge to 95% as the sample size goes from $n = 10$ to 100. The confidence interval based on SBJPS method has the nearest CP to the nominal level 95% in all considered cases except for $n = 10$ in which the NA confidence interval usually has slightly better performance. The BJPSS confidence interval has the poorest performance for $n \leq 30$, and the difference between its CP and the nominal level becomes sizeable for $n = 10$, but for $n \geq 50$ its CP usually fall between CPs of the confidence intervals based on SBJPS and NA approaches. Similar to what we have observed for standard normal distribution in Fig. 1, it is evident from Fig. 2 that the set size m and the quality of ranking λ do not have much effect on CPs of the confidence intervals for the standard log-normal distribution, as well.

Comparing Fig. 2 with Figs. S3 and S4 in the Supplementary Material, we find that the patterns of CPs of different confidence intervals for $LN(0, 1)$, $G(0.5)$, $W(0.5)$ distributions are almost the same.

5 A Real Data Example

In this section, a real dataset is used to illustrate the potential application proposed procedures in this paper to construct a confidence interval for the population mean. The real dataset is then used the evaluate the performance of different confidence intervals in the JPS setting.

Bone mineral density (BMD) is the amount of bone mineral in bone tissues, and it is frequently used in medicine as an indicator for detecting osteoporosis. The BMD measurement is usually made over the lumbar spine and over the upper part of the hip using dual-energy X-ray absorptiometry (DEXA) technology, and a person is considered to be suffering from osteoporosis if his/her BMD measurement using DEXA technology is no larger than 0.56. Note that obtaining exact measurement of BMD using DEXA technology is costly, and the technology may not be easily accessible in some developing countries. It is also inconvenient to use because it need a medical expert to manually segment images. But, a medical expert can simply assign judgement ranks to sample units in a set of small size in terms of probability of suffering from osteoporosis. This can be done using the medical expert's personal experience or by checking if the patient has some risk factors of osteoporosis such as family history, cigarette smoking, excessive alcohol and caffeine consumption. Therefore, JPS seems to be a better alternative to SRS for this application.

The dataset used in this section in obtained from the third National Health and Nutrition Examination Survey (NHANES III), and is available online at http://www.cdc.gov/nchs/nhanes/nh3data.htm.[1] We consider the BMD of people who suffer from osteoporosis in NHANES III dataset as our hypothetical population (denoted here after by BMD dataset). The histogram of BMD dataset is presented in Fig. 3. Suppose that we are interested in constructing a confidence interval for mean of BMD in this population. So, we assume that the BMD is our variable of interest. Using $n = 20$ and
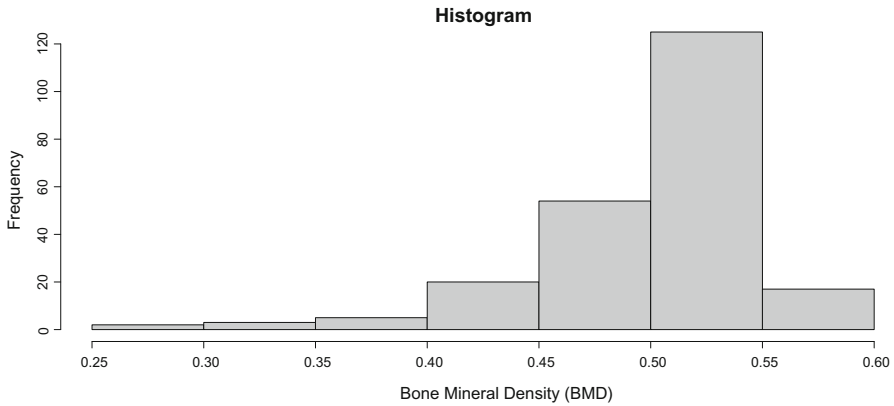
---

[1] Access date March 2020.

Histogram

Fig. 3 The histogram of BMD dataset

Table 1 A JPS sample of size $n = 20$ using set size $m = 3$ from BMD dataset

(Y_i, R_i)	(Y_i, R_i)	(Y_i, R_i)	(Y_i, R_i)	(Y_i, R_i)
(0.405, 1)	(0.538, 1)	(0.516, 2)	(0.445, 1)	(0.517, 2)
(0.527, 2)	(0.536, 2)	(0.459, 2)	(0.548, 3)	(0.538, 3)
(0.545, 3)	(0.497, 3)	(0.445, 1)	(0.536, 2)	(0.554, 3)
(0.543, 2)	(0.552, 3)	(0.539, 2)	(0.458, 2)	(0.541, 1)

Table 2 95% JPS confidence intervals for the mean of the BMD using the data in Table 1

Method	Lower bound	Upper bound
$SBJPS$	0.487	0.534
$BJPSS$	0.489	0.530
NA	0.488	0.530

$m = 3$, we draw a JPS sample from the given population. To do so, we first draw a SRS sample of size 20 from the BMD dataset and measure all of them. Then we each measured unit, we draw 2 additional units to create a set of size $m = 3$. Sampling with replacement is considered, so the assumption of independence is guaranteed. The rank of each measured unit in the set of size $m = 3$ is determined using the linear ranking model described in Sect. 4 with $\lambda = 0.7$. The sample units with their ranks are presented in Table 1 and the corresponding confidence intervals for the population mean are given in Table 2.

We next use this BMD dataset to compare performance of different confidence intervals. In doing so, we set $n \in \{10, 20, 30, 50\}$, $m \in \{3, 4, 5\}$ and for each combination of (n, m), we draw 10,000 JPS samples from the BMD dataset, where all samplings are done with replacement. The ranking is done using the linear ranking model as described in Sect. 4 with $\lambda \in \{1, 0.9, 0.7, 0.5\}$. Finally, we estimate CPs of bootstrap expanded percentile and NA confidence intervals for the population mean in the JPS setting using 10,000 samples. The bootstrap size is taken as $B = 1000$. The results are presented in Table 3.

Table 3 Estimated coverage probability (CP) of different confidence intervals of the population mean in the JPS setting at nominal level 95% using BMD dataset when the ranking is done using linear ranking model with $\lambda \in \{1, 0.9, 0.7, 0.5\}$

n	m	$\lambda = 1$			$\lambda = 0.9$			$\lambda = 0.7$			$\lambda = 0.5$		
		SBJPS	BJPSS	NA	SBJPS	BJPSS	NA	SBJPS	BJPSS	NA	SBJPS	BJPSS	NA
10	3	**0.912**	0.761	0.901	**0.911**	0.776	0.901	**0.912**	0.779	0.901	**0.908**	0.782	0.906
	4	**0.903**	0.668	0.902	**0.908**	0.677	0.907	0.908	0.701	**0.910**	0.903	0.697	**0.904**
	5	0.899	0.565	**0.901**	**0.906**	0.567	0.905	0.905	0.625	**0.913**	0.908	0.636	**0.910**
20	3	**0.932**	0.902	0.918	**0.940**	0.909	0.923	**0.936**	0.910	0.924	**0.939**	0.913	0.927
	4	**0.931**	0.860	0.910	**0.934**	0.873	0.918	**0.936**	0.880	0.926	**0.936**	0.880	0.922
	5	**0.931**	0.823	0.915	**0.930**	0.822	0.910	**0.932**	0.841	0.918	**0.929**	0.839	0.914
30	3	**0.946**	0.932	0.927	**0.948**	0.935	0.931	**0.951**	0.940	0.933	**0.950**	0.937	0.930
	4	**0.945**	0.912	0.919	**0.954**	0.929	0.934	**0.951**	0.928	0.932	**0.950**	0.930	0.932
	5	**0.945**	0.894	0.917	**0.944**	0.897	0.920	**0.950**	0.908	0.927	**0.950**	0.911	0.931
50	3	**0.952**	0.947	0.930	**0.951**	0.947	0.936	0.956	**0.955**	0.940	0.960	**0.956**	0.939
	4	**0.952**	0.944	0.933	**0.951**	0.954	0.940	0.955	**0.952**	0.941	0.960	**0.952**	0.940
	5	**0.949**	0.933	0.926	**0.953**	0.937	0.930	**0.953**	0.943	0.935	0.958	**0.945**	0.936

The confidence interval with the nearest estimated CP to the nominal level 95% is highlighted in bold face

The results in Table 3 are consistent with what we observed in Sect. 4. The confidence interval based on SBJPS method often provides the nearest CP to the nominal level 95%, which can be justified by the fact that the distribution of BMD dataset is skewed (see Fig. 3).

6 Conclusion

Judgment post stratification (JPS) is a useful sampling scheme in applications requiring cost efficiency. The JPS adds additional judgment ranking information to a simple random sample to improve estimation of the population parameters. The judgment ranks are obtained through subjective judgment of an expert, concomitant variable, or a combination of them and need not to be accurate.

Since JPS is a cost efficient sampling method and is applicable in the settings in which obtaining exact values of sample units is much harder than finding their ranks in a set of small size, it is very common that a researcher cannot obtain a JPS sample with *large enough* size to use asymptotic distribution of the estimators. Therefore, existence an alternative method to draw statistical inference based on a JPS sample is essential.

In this paper, we described two bootstrap methods based on a JPS sample, i.e. simple bootstrap JPS (SBJPS) and bootstrap JPS by stratum (BJPSS). SBJPS has been already used in the literature without studying its consistency and BJPSS is our proposal. We then showed that the bootstrap procedures are consistent. Finally, as an application of the bootstrap methods, we discussed construction of bootstrap expanded percentile confidence intervals for the population mean using empirical mean estimator in JPS and compared their coverage probabilities with the confidence interval obtained via normal approximation (NA) using Monte Carlo simulation study for variety choices of sample size, set size and parent distribution. Based on the simulation study, we found that each of NA and SBJPS techniques can be the best method in terms of coverage probability (CP) for construction of a confidence interval for the population mean in the most considered cases. So, we recommend using SBJPS technique to construct a confidence interval for the population mean based on bootstrap methods when the sample size is small and normality assumption is in suspect.

Although this work studies two resampling methods for a JPS sample, it remains an ample space for future research in this field. For example, several improved mean estimators over the empirical mean estimator have been proposed in the literature by Wang et al. (2008), Frey and Feeman (2012) and Frey (2016), and exact expressions for variances of some of those mean estimators are not analytically available. Thus, one can use the bootstrap methods described in this paper to construct confidence intervals for the population mean using the improved mean estimators and it is intuitively expected that those confidence intervals have better performance than confidence intervals based on the empirical mean estimator. It is known that the bias-corrected and accelerated and bootstrap-t methods often have a competitive performance as compared to bootstrap expanded percentile technique (Hall 1988). Therefore, it is of interest to discuss construction of those confidence intervals based on a JPS sample. One can also think

of constructing confidence intervals for other population attributes rather than the population mean.

Finally, we would like to mention that for a JPS sample of small size, it is very common to observe empty strata (i.e. $n_r = 0$ for some $r \in \{1, \ldots, m\}$). Statistical inference based on a JPS sample with at least one empty stratum may be different from JPS sample with no empty strata. For example, Zamanzade and Wang (2017) showed that when there is at least one empty stratum, maximum likelihood estimator of the population proportion can be much more efficient than other estimators in some certain circumstances. Thus, another interesting topic for future research is to study of the effect of empty strata on the performance of bootstrap confidence intervals in JPS.

References

Bickel PJ, Freedman DA (1981) Some asymptotic theory for bootstrap. Ann Stat 9:1196–1217

Dastbaravarde A, Zamanzade E (2020) On estimation of $P(X > Y)$ based on judgment post stratification. Stat Pap 61:767–785

Dastbaravarde A, Arghami NR, Sarmad M (2016) Some theoretical results concerning non parametric estimation by using a judgment poststratification sample. Commun Stat Theory Methods 45(8):2181–2203

Dell TR, Clutter JL (1972) Ranked set sampling theory with order statistics background. Biometrics 28(2):545–555

Duembgen L, Zamanzade E (2020) Inference on a distribution function from ranked set samples. Ann Inst Stat Math 72:157–185

Efron B (1979) Bootstrap methods: another look at the jackknife. Ann Stat 7:1–26

Frey J (2016) A more efficient mean estimator for judgment post-stratification. J Stat Comput Simul 86(7):1404–1414

Frey J, Feeman TG (2012) An improved mean estimator for judgment post-stratification. Comput Stat Data Anal 56(2):418–426

Frey J, Feeman TG (2013) Variance estimation using judgment post-stratification. Ann Inst Stat Math 65(3):551–569

Frey J, Ozturk O (2011) Constrained estimation using judgment post-stratification. Ann Inst Stat Math 63(4):769–789

Hall P (1988) Theoretical comparison of bootstrap confidence intervals. Ann Stat 16:927–953

He X, Chen W, Qian W (2018) Maximum likelihood estimators of the parameters of the log-logistic distribution. Stat Pap. https://doi.org/10.1007/s00362-018-1011-3

Hesterberg T (2015) What teachers should know about the bootstrap: resampling in the undergraduate statistics curriculum. Am Stat 69(4):371–386

MacEachern SN, Stasny EA, Wolfe DA (2004) Judgment post stratification with imprecise rankings. Biometrics 60(1):207–215

Mahdizadeh M, Zamanzade E (2019) Efficient body fat estimation using multistage pair ranked set sampling. Stat Methods Med Res 28(1):223–234

Omidvar S, Jafari Jozani M, Nematollahi N (2018) Judgment post-stratification in finite mixture modeling: an example in estimating the prevalence of osteoporosis. Stat Med 37(30):4823–4836

Ozturk O (2014) Statistical inference for population quantiles and variance in judgment post-stratified samples. Comput Stat Data Anal 77:188–205

Ozturk O (2015) Distribution free two-sample methods for judgment post-stratified data. Stat Sin 25(4):1691–1712

Ozturk O (2016) Statistical inference based on judgment post-stratified samples in finite population. Surv Methodol 42(2):239–262

Presnell B, Bohn LL (1999) U-statistics and imperfect ranking in ranked set sampling. J Nonparametr Stat 10(2):111–126

Qian W, Chen W, He X (2019) Parameter estimation for the Pareto distribution based on ranked set sampling. Stat Pap. https://doi.org/10.1007/s00362-019-01102-1

Wang X, Lim J, Stokes L (2008) A nonparametric mean estimator for judgment poststratified data. Biometrics 64(2):355–363

Wang X, Wang K, Lim J (2012) Isotonized CDF estimation from judgment poststratification data with empty strata. Biometrics 68(1):194–202

Wang X, Lim J, Stokes SL (2016) Using ranked set sampling with cluster randomized designs for improved inference on treatment effects. J Am Stat Assoc 111(516):1576–1590

Zamanzade E (2016) An isotonized variance estimator for judgment post stratified data. J Korean Stat Soc 45(1):25–32

Zamanzade E, Mahdizadeh M (2020) Using ranked set sampling with extreme ranks in estimating the population proportion. Stat Methods Med Res 29(1):165–177

Zamanzade E, Vock M (2015) Variance estimation in ranked set sampling using a concomitant variable. Stat Probab Lett 105:1–5

Zamanzade E, Vock M (2018) Some nonparametric tests of perfect judgment ranking for judgment post stratification. Stat Pap 59(3):1085–1100

Zamanzade E, Wang X (2017) Estimation of population proportion for judgment post-stratification. Comput Stat Data Anal 112:257–269