Two-Stage Fixed-Width and Bounded-Width Confidence Interval Estimation Methodologies for the Common Correlation in an Equi-Correlated Multivariate Normal Distribution

Shyamal K. De[1] **and Nitis Mukhopadhyay**[2]
[1]School of Mathematical Sciences, National Institute of Science Education and Research,
HBNI, Jatni, Odisha, India
[2]Department of Statistics, University of Connecticut,
Storrs, Connecticut, USA

Abstract: In this paper, two-stage and fixed-sample-size procedures are developed for constructing confidence intervals for the common correlation ρ of an equi-correlated multivariate normal distribution. Two different approaches for estimation are considered, namely fixed-width and bounded-width interval estimation. In the fixed-width confidence interval estimation problem, a two-stage procedure is developed and the exact distribution of the corresponding stopping variable and the exact coverage probability of the interval estimator are derived. Asymptotic optimality properties such as asymptotic first-order efficiency and asymptotic consistency properties of the two-stage procedure are established. Bounded-width confidence intervals are obtained by applying fixed-accuracy estimation methodologies of Mukhopadhyay and Banerjee (2014,2015a,b) and Banerjee and Mukhopadhyay (2016) using different transformations on ρ. Both fixed-sample-size and two-stage sampling methodologies for bounded-width confidence interval estimation of ρ are developed incorporating such transformations. Finally, performances of all the procedures are compared via extensive sets of simulation studies.

Keywords: Asymptotic consistency; Asymptotic efficiency; Exact computations; Fixed-accuracy; Proportional closeness; Simulations; Stopping variable; Two-stage sampling.

Subject Classifications: 60G51; 60K15; 60K40.

1. INTRODUCTION

Correlated multivariate normal distributions appear in many areas of statistics such as *analysis of variance* (ANOVA) design with repeated measurements. Suppose that there are n experimental units each measured on a certain variable at m time periods. Thus, m measurements from a single unit are correlated. However, the measurements from different units may be assumed independent.

Address correspondence to Shyamal K. De, School of Mathematical Sciences, National Institute of Science Education and Research, Jatni, Odisha 752050, India; Tel: +91 (674) 249-4083; Fax: +91 (674) 249-4081; E-mails: sde@niser.ac.in; nitis.mukhopadhyay@uconn.edu

In an ANOVA design with repeated measurements, we assume that the error part can be modeled by an m-variate normal distribution with mean vector $\mathbf{0}$ and covariance matrix Σ. We investigate the problem of estimation of the common correlation coefficient ρ when Σ has the following special structure:

$$\begin{gathered}\Sigma = \sigma^2 R \text{ where the correlation matrix is } R = (1-\rho)I_{m\times m} + \rho J_{m\times m}, \\ I_{m\times m} \text{ is an identity matrix and } J_{m\times m} = \mathbf{1}\mathbf{1}^T \text{ is a matrix of 1's.}\end{gathered} \tag{1.1}$$

Here, both parameters σ^2 and ρ are assumed unknown with $\sigma \in \mathbb{R}^+$ and $\rho \in (-(m-1)^{-1}, 1)$. Let us denote the parameter vector $\theta = (\rho, \sigma^2)$. We revisit this equi-correlated model considered by many authors including Rao (1973, p. 67), Mukhopadhyay (1982), SenGufta (1987), Zacks and Ramig (1987), Ghezzi and Zacks (2005), Haner and Zacks (2013), De (2014), and De and Mukhopadhyay (2015), among others.

Let us now precisely formulate our specific problem on hand. Consider a sequence of *independently and identically distributed* (i.i.d.) m-dimensional vectors $\mathbf{X}_1, \mathbf{X}_2, \ldots$ from a multivariate normal distribution with mean vector $\mathbf{0}$ and variance-covariance matrix $\Sigma = \sigma^2 R$ as given in (1.1). Based on these multivariate observations, we would like to develop two-stage sampling strategies as well as fixed-sample-size methodologies for *fixed-width* and *bounded-width* confidence interval estimation of the common correlation ρ in the presence of the nuisance parameter σ^2.

Multistage sampling designs were first developed by Mahalanobis (1940) followed by path-breaking constructions of two-stage methodologies due to Stein (1945,1949). For a review, one may refer to Siegmund (1985), Ghosh and Sen (1991), Ghosh et al. (1997), Mukhopadhyay and de Silva (2009), and Zacks (2009a,b), among other sources.

Mukhopadhyay and Banerjee (2014,2015a,b) and Banerjee and Mukhopadhyay (2016) drew attention to certain important drawbacks inherent in the constructions of *fixed-width* (Chow and Robbins, 1965; Khan, 1969) and *proportional accuracy* (Nadás, 1969; Willson and Folks, 1983; Zacks, 1966) confidence intervals for an unknown parameter whose support was, for example, $\mathbb{R}^+$ or $(0,1)$. Mukhopadhyay and Banerjee (2014) and Banerjee and Mukhopadhyay (2016) then resolved such drawbacks by introducing the idea of a *fixed-accuracy* confidence interval which always lies within the relevant parameter space under consideration. Under the model (1.1), De and Mukhopadhyay (2015) estimated σ^2 with prescribed fixed-accuracy in the presence of the nuisance parameter ρ in the spirit of Mukhopadhyay and Banerjee (2014) and Banerjee and Mukhopadhyay (2016).

In this paper, our goal is to construct both fixed-width and bounded-width confidence intervals for the common correlation parameter ρ. The analyses would be vastly different, especially from those in Mukhopadhyay and Banerjee (2015), since the parameter space of ρ happens to be $(-(m-1)^{-1}, 1)$ instead of $\mathbb{R}^+$ or $(0,1)$. We will include the *exact* distributions of the stopping variables under the proposed two-stage sampling strategies and the exact coverage probabilities associated with the confidence interval estimators.

It is helpful to note that the exact formulas for evaluating risk functions and other useful functionals in the context of a number of two-stage and purely sequential methodologies (especially in the cases of gamma and exponential distributions) were systematically developed by Zacks (2009a,b,2017), Zacks and Mukhopadhyay (2006a,b), and Zacks and Khan (2011).

Fixed-sample-size estimation of the common variance of correlated multinormal distributions (1.1) was first considered by Zacks and Ramig (1987). They developed the *uniform minimum variance unbiased estimators* (UMVUEs) and the *maximum likelihood estimators* (MLEs) of σ^2 and ρ.

The dissertation of Haner (2012) and the subsequent paper by Haner and Zacks (2013) developed theory and methodologies for fixed-width confidence interval estimation of σ^2. De (2014) developed a modified three-stage procedure for fixed-width interval estimation of σ^2.

At the present juncture, we proceed with the estimation problems surrounding the common correlation parameter ρ. In what follows, we use the phrases *with probability 1* (w.p.1) and *almost sure* or *almost surely* (a.s.) interchangeably. In this article, we begin with some preliminaries (Section 2) and then build the theory and methodology of two-stage fixed-width confidence interval estimation (Section 3).

Section 3.1 includes sample size calculation for prescribed width, Section 3.2 gives the formulation of our two-stage methodology, 3.3 shows the exact distribution of final sample size N and the corresponding estimator $\widehat{\rho}_N$, Section 3.4 discusses a choice of pilot sample size and asymptotic properties, followed by Section 3.5 highlighting numerical studies.

Section 4 develops the theory and methodology of confidence interval estimation of ρ with fixed-accuracy approach. Section 4.1 introduces a class of one-to-one transformations $g_\nu(.)$ on ρ leading to bounded-width intervals, Section 4.2 lays down the fixed-sample-size methods under transformation $g_1(.)$ on ρ, Section 4.3 introduces two-stage methodologies with associated asymptotics under two other promising transformations $g_2(.)$ and $g_3(.)$ on ρ, Section 4.4 introduces a two-stage methodology with associated asymptotic properties under the next promising transformation $g_4(.)$ on ρ, followed by Section 4.5 highlighting analysis of simulated datasets.

2. PRELIMINARY THEORY

Let us now discuss few preliminary results that are crucial for developing the two-stage and fixed-sample-size methodologies of this paper. These results can also be found in other articles such as Haner and Zacks (2013), De (2014), De and Mukhopadhyay (2015).

We first consider a very special orthogonal transformation, called the Helmert transformation, applied to the original set of observations $\boldsymbol{X}_1, \boldsymbol{X}_2 \ldots$. One may refer to Mukhopadhyay (2000, pp. 197-201) for additional insights.

Thereby, we obtain a transformed sequence of uncorrelated multivariate normal vectors $\boldsymbol{Y}_1, \boldsymbol{Y}_2 \ldots$, where $\mathbf{Y}_i = H\mathbf{X}_i$, for $i = 1, 2, \ldots$, and H is the Helmert matrix given as:

$$H = \begin{bmatrix} \frac{1}{\sqrt{m}} & \frac{1}{\sqrt{m}} & \cdots & & & \frac{1}{\sqrt{m}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & & \cdots & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ \frac{1}{\sqrt{m(m-1)}} & & \cdots & \frac{1}{\sqrt{m(m-1)}} & & -\frac{(m-1)}{\sqrt{m(m-1)}} \end{bmatrix}. \tag{2.1}$$

The transformed observations $\boldsymbol{Y}_1, \boldsymbol{Y}_2 \dots$ are i.i.d. $N_m(\mathbf{0},\, \sigma^2\Lambda)$ random variables where

$$\Lambda = HRH^T = \text{diag}(\lambda_1, \dots, \lambda_m)$$

is a diagonal matrix, and $\lambda_1, \dots, \lambda_m$ are the eigenvalues of the correlation matrix R. It follows from Rao (1973, p. 67) that the eigenvalues are

$$\lambda_1 = 1 + (m-1)\rho \quad \text{and} \quad \lambda_2 = \cdots = \lambda_m = 1 - \rho.$$

To ensure that the multivariate normal distribution of $\boldsymbol{Y}_i$ is non-singular, $\lambda_1, \dots, \lambda_m$ must be positive, and thus the parameter space of the common correlation parameter ρ must be $\left(-\frac{1}{m-1}, 1\right)$. Given n i.i.d. observations $\boldsymbol{Y}_1, \dots, \boldsymbol{Y}_n$ from $N_m(\mathbf{0},\, \sigma^2\Lambda)$, the likelihood function of (ρ, σ^2) is

$$L(\rho, \sigma^2 | V_{1n}, V_{2n}) \propto (\sigma^2)^{-\frac{nm}{2}} (1-\rho)^{-\frac{n(m-1)}{2}} (1+(m-1)\rho)^{-\frac{n}{2}} \exp\left\{ -\frac{1}{2\sigma^2} \left(\frac{V_{1n}}{1+(m-1)\rho} + \frac{V_{2n}}{1-\rho} \right) \right\},$$

where

$$V_{1n} = \textstyle\sum_{i=1}^n Y_{i1}^2 \text{ and } V_{2n} = \sum_{i=1}^n \sum_{j=2}^m Y_{ij}^2.$$

It follows from Lehmann (1997) that (V_{1n}, V_{2n}) is a complete sufficient statistic of $\theta = (\rho, \sigma^2)$, and thus, one should consider estimators of θ based on (V_{1n}, V_{2n}). Properties of V_{1n} and V_{2n} are:

(i) $V_{1n} \;\sim\; \sigma^2 \left(1 + (m-1)\rho\right) \chi_n^2$,

(ii) $V_{2n} \;\sim\; \sigma^2 \left(1 - \rho\right) \chi_{n(m-1)}^2$, and

(iii) V_{1n} and V_{2n} are independent.

The MLEs of σ^2 and ρ, as given in Zacks and Ramig (1987), are:

$$\widehat{\sigma}_n^2 = \frac{V_{1n} + V_{2n}}{nm} \text{ and} \tag{2.2}$$

$$\widehat{\rho}_n = \frac{V_{1n} - V_{2n}(m-1)^{-1}}{V_{1n} + V_{2n}}. \tag{2.3}$$

Both these estimators are shown to be strongly consistent and asymptotically normal. The asymptotic normality of $\widehat{\rho}_n$ as $n \to \infty$ is given in Zacks and Ramig (1987) as:

$$\sqrt{n}\,(\widehat{\rho}_n - \rho) \xrightarrow{\mathcal{L}} N\left(0, \sigma_\rho^2\right) \quad \text{where } \sigma_\rho^2 = \frac{2(1-\rho)^2(1+(m-1)\rho)^2}{m(m-1)}. \tag{2.4}$$

This result is crucial for developing both fixed-width and bounded-width confidence intervals of ρ.

3. FIXED-WIDTH INTERVAL ESTIMATION

In this section, we develop the theory and methodology for fixed-width interval estimation of ρ and establish various asymptotic and exact properties of the procedure. Prior to developing such methodology, we provide the following motivation for considering a two-stage procedure.

3.1. Sample Size Calculation for Prescribed Width: A Motivation

A large sample confidence interval for ρ can be constructed based on the asymptotic normality of $\widehat{\rho}_n$. Since $\widehat{\rho}_n$ is a strong consistent estimator of ρ and σ_ρ^2 is a continuous function of ρ, $\sigma_{\widehat{\rho}_n}^2$ converges almost surely to σ_ρ^2 and

$$\sqrt{n}\left(\widehat{\rho}_n - \rho\right)/\sigma_{\widehat{\rho}_n} \xrightarrow{\mathcal{L}} N(0,1) \text{ as } n \to \infty.$$

Thus, a large sample $(1-\alpha)$ confidence interval for ρ based on n i.i.d. observations is

$$I_n^{(ls)} = \left[\widehat{\rho}_n - z_{\alpha/2}\frac{\sigma_{\widehat{\rho}_n}}{\sqrt{n}},\, \widehat{\rho}_n + z_{\alpha/2}\frac{\sigma_{\widehat{\rho}_n}}{\sqrt{n}}\right], \tag{3.1}$$

where z_α is the upper α^{th} quantile of a standard normal random variable. Let us consider the question of determining the sample size n if an experimenter wants some given w to be the width of this interval. Clearly, the width $W = 2z_{\alpha/2}\sigma_{\widehat{\rho}_n}/\sqrt{n}$ is random, and in order to achieve a prescribed width w, we need n such that $n = \left(2z_{\alpha/2}\sigma_{\widehat{\rho}_n}/w\right)^2$. As the estimator $\widehat{\rho}_n$ involves n, it is not possible to solve for n explicitly.

This problem is similar to the case of sample size calculation for constructing a large sample confidence interval for p with prefixed width w based on n i.i.d. observations $X_1, \ldots, X_n$ from Bernoulli(p) distribution. It is well known that a $(1-\alpha)$ confidence interval for p, derived from the asymptotic normality of sample proportion $\widehat{p}_n = \frac{1}{n}\sum_{i=1}^n X_i$, is given by $\left[\widehat{p}_n \pm z_{\alpha/2}\sqrt{\widehat{p}_n(1-\widehat{p}_n)/n}\right]$. Therefore, to attain a predetermined width w for the above confidence interval, we need sample size n such that

$$n = \left(\frac{2z_{\alpha/2}}{w}\right)^2 \widehat{p}_n(1-\widehat{p}_n) \le \left(\frac{z_{\alpha/2}}{w}\right)^2$$

using the inequality $p(1-p) \le 1/4$ for all $p \in (0,1)$. Since sample size n cannot be determined explicitly from the above equation, often experimenters use the conservative choice as $n = \lfloor \left(\frac{z_{\alpha/2}}{w}\right)^2 \rfloor + 1$. Here, the notation $\lfloor x \rfloor$ represents the largest integer less than x. Interestingly, we can also have a similar approach of getting a conservative choice of n for estimation of ρ.

Note that $\sigma_\rho^2 = \frac{2}{m(m-1)}(f(\rho))^2$ where $f(\rho) = (1-\rho)(1+(m-1)\rho) > 0$ for all $\rho \in (-1/(m-1), 1)$. The second derivative test yields that $f(\rho)$ attains its global maximum at $\rho = \rho_0$, where we denote

$$\rho_0 = \frac{m-2}{2(m-1)} \in \left(\frac{-1}{m-1}, 1\right), \text{ and thus, } \sup_{\frac{-1}{m-1}<\rho<1} \sigma_\rho^2 = \left(\frac{m}{2(m-1)}\right)^3.$$

Therefore, to ensure a given width w of the interval $I_n^{(ls)}$, the sample size n should satisfy

$$n = \left(2z_{\alpha/2}\sigma_{\widehat{\rho}_n}/w\right)^2 \le \left(\frac{2z_{\alpha/2}}{w}\right)^2 \left(\frac{m}{2(m-1)}\right)^3 \tag{3.2}$$

which yields a conservative choice of sample size as

$$n_{con} = \left\lfloor \left(\frac{2z_{\alpha/2}}{w}\right)^2 \left(\frac{m}{2(m-1)}\right)^3 \right\rfloor + 1.$$

One may be tempted to use this choice of sample size for planning the experiment to estimate ρ. However, we show in Table 1 that n_{con} may be too conservative leading to unnecessarily high cost of the experiment, and thus may not be recommended in practice. A more reasonable method of sample size calculation would be to use a two-stage sampling procedure where ρ is first estimated based on a pilot sample of smaller size. Then, instead of plugging in the maximizer $\rho = \rho_0$ in the left hand side of (3.2) to obtain n_{con}, we can plug in the estimated ρ in (3.2) and obtain an estimate of the required sample size.

Table 1 shows that with n_{con} many observations the average width, represented as $\overline{W}_{con}$, can be much smaller than the required width w, which is being overprotective about the precision of the interval estimator. Instead, a two-stage procedure can lead to quite significant savings (sometimes even more than 86%) in sample size without loosing much on the coverage probability. With this motivation, we formally develop the theory and methodology of two-stage sampling procedure for fixed-width interval estimation of ρ in the following sections.

Table 1. Comparison between the fixed-sample-size method for constructing a $(1-\alpha)$ confidence interval of ρ with sample size n_{con} and the two-stage method of subsection 3.2 for fixed-width interval estimation of ρ with pilot sample size $k = 100$, $\sigma = 1$ and $\alpha = 0.1$

m	ρ	$w = 2d$	n_{con}	n^*	$\widehat{E}(N)$	$\widehat{CP}_{con}$	$\widehat{CP}$	$\overline{W}_{con}$
2	0.8	0.02	27056	3507	3657.61	0.8993	0.8901	0.0072
		0.05	4329	562	586.80	0.8993	0.8881	0.018
	-0.7	0.02	27056	7038	7249.21	0.8997	0.8919	0.0102
		0.05	4329	1126	1159.21	0.8989	0.8919	0.0255
5	0.75	0.02	6606	2706	2757.25	0.8987	0.8966	0.0128
		0.05	1057	433	441.29	0.8991	0.8969	0.0320
	-0.1	0.02	6606	1179	1193.73	0.8985	0.8910	0.0084
		0.05	1057	189	191.35	0.8994	0.8827	0.0211
10	0.75	0.02	4640	2257	2292.28	0.8994	0.8989	0.0140
		0.05	743	362	366.96	0.8995	0.8979	0.0349
	0.1	0.02	4640	1759	1756.69	0.8995	0.8928	0.0123
		0.05	743	282	281.63	0.8990	0.8888	0.0308

3.2. Two-Stage Procedure for a Fixed-Width Confidence Interval

Based on n i.i.d. observations $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ from $N_m(\mathbf{0}, \sigma^2 R)$, the objective of fixed-width interval estimation of common correlation ρ is to find an interval $J_n = [\widehat{\rho}_n - d, \widehat{\rho}_n + d]$ for some prefixed $d > 0$ such that the coverage probability $P(\rho \in J_n) \geq 1 - \alpha$, for some given $\alpha \in (0, 1)$. Using the asymptotic normality of $\widehat{\rho}_n$ as in (2.4), we can approximate the coverage probability,

$$P_\theta(\rho \in J_n) \approx 2\Phi\left(\frac{\sqrt{n}d}{\sigma_\rho}\right) - 1 \quad \text{for large } n,$$

where Φ denotes the *cumulative distribution function* (CDF) of a standard normal random variable. Solving the inequality $2\Phi\left(\frac{\sqrt{n}d}{\sigma_\rho}\right) - 1 \geq 1 - \alpha$, we obtain the sample size n as the smallest integer n such that

$$n \geq \frac{\beta_d (1-\rho)^2 \{1 + (m-1)\rho\}^2}{m(m-1)},$$

where $\beta_d = 2d^{-2} z_{\alpha/2}^2$. Thus, the approximate *minimum sample size* required to achieve $(1-\alpha)$ coverage probability is

$$n^* \equiv n^*(d) = \left\lfloor \frac{\beta_d (1-\rho)^2 \{1 + (m-1)\rho\}^2}{m(m-1)} \right\rfloor + 1, \tag{3.3}$$

provided ρ were known.

Since ρ is an unknown parameter, we must collect samples in at least two stages to obtain a fixed-width interval with $(1-\alpha)$ confidence coefficient. Below, we describe a two-stage sampling procedure in the spirit of Stein (1945,1949) for fixed-width interval estimation of ρ using a fixed pilot sample size k. Later, we will discuss some choices of pilot sample size in order to achieve asymptotic optimality properties of the procedure.

Stage I: For prefixed k, obtain k independent and identically distributed observations $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_k$, known as pilot sample, from $N_m(\mathbf{0}, \sigma^2 R)$. Transform these observations to $\mathbf{Y}_i = H\boldsymbol{X}_i, i = 1, \ldots, k$, and compute the maximum likelihood estimator $\widehat{\rho}_k$ as in (2.3) based on $\mathbf{Y}_1, \ldots, \mathbf{Y}_k$. Estimate the optimal sample size n^* by

$$K^* \equiv K^*(d) = \left\lfloor \frac{\beta_d (1-\widehat{\rho}_k)^2 \{1 + (m-1)\widehat{\rho}_k\}^2}{m(m-1)} \right\rfloor + 1. \tag{3.4}$$

Suppose N denotes the final sample size for the two-stage procedure. If $K^* \leq k$, stop sampling and set $N = k$ and $\widehat{\rho}_N = \widehat{\rho}_k$. Otherwise, proceed to Stage II.

Stage II: Obtain additional $(K^* - k)$ i.i.d. observations from $N_m\left(\mathbf{0}, \sigma^2 R\right)$ independent of the initial k samples and label the observations as $\boldsymbol{X}^*_{k+1}, \ldots, \boldsymbol{X}^*_{K^*}$. Transform these observations to $\mathbf{Y}^*_i = H\boldsymbol{X}^*_i$, $i = k+1, \ldots, K^*$. Define

$$V^*_{1(N-k)} = \sum_{i=k+1}^{K^*} Y^{*2}_{i1} \text{ and } V^*_{2(N-k)} = \sum_{i=k+1}^{K^*} \sum_{j=2}^{m} Y^{*2}_{ij},$$

and set $N = K^*$. Thus, the final sample size or the stopping variable for the two-stage procedure is

$$N \equiv N(d) = \max\{k, K^*(d)\}.$$

The final point estimator of ρ is

$$\widehat{\rho}_N = \frac{V_{1k} + V^*_{1(N-k)} - \left(V_{2k} + V^*_{2(N-k)}\right)/(m-1)}{V_{1k} + V^*_{1(N-k)} + V_{2k} + V^*_{2(N-k)}}$$

and the fixed-width $(1-\alpha)$ confidence interval for ρ is $J_N = [\widehat{\rho}_N - d,\ \widehat{\rho}_N + d]$.

3.3. Exact Distribution of N and $\widehat{\rho}_N$

For any fixed pilot sample size k, we derive the exact formulas for the cumulative distribution function of the stopping variable N and its functionals. Moreover, we derive the exact formula for the coverage probability by first deriving the exact distribution of the final estimator $\widehat{\rho}_N$. Let $F_{n,m}$ be a random variable having F-distribution with numerator and denominator degrees of freedom n and m respectively.

Theorem 3.1. *Let k be any fixed positive integer less than u, and $G_k(.)$ be the cumulative distribution function of $F_{k(m-1),k}$. The cumulative distribution function of N, denoted by $F_{N,\theta}(.)$, under θ, is given as*

$$F_{N,\theta}(n) = \begin{cases} 0 & \text{if } n < k, \\ G_k\left(\frac{c_{1n}\{1+(m-1)\rho\}}{(m-1)(1-\rho)}\right) - G_k\left(\frac{c_{2n}\{1+(m-1)\rho\}}{(m-1)(1-\rho)}\right) + 1 & \text{if } k \le n < u, \\ 1 & \text{if } n \ge u, \end{cases}$$

where we denote

$$u = \frac{\beta_d\, m^3}{16(m-1)^3},\ c_{1n} = \frac{2}{1+\sqrt{1-c_n}} - 1,\ c_{2n} = \frac{2}{1-\sqrt{1-c_n}} - 1,\ \text{and}\ c_n = \sqrt{\frac{n}{u}}.$$

Proof. Note that $P(N < k) = 0$ as $N = \max\{k, K^*\}$. For integer valued n in $[k, u)$,

$$\begin{aligned} P_\theta(N \le n) = P_\theta(K^* \le n) &= P_\theta\left(\frac{\beta_d(1-\widehat{\rho}_k)^2\{1+(m-1)\widehat{\rho}_k\}^2}{m(m-1)} \le n\right) \\ &= P_\theta\left((1-\widehat{\rho}_k)\{1+(m-1)\widehat{\rho}_k\} \le \sqrt{\frac{nm(m-1)}{\beta_d}}\right) \\ &= P_\theta\left(\left\{\frac{V_{1k}-V_{2k}}{V_{1k}+V_{2k}}\right\}^2 \ge 1 - c_n\right), \end{aligned}$$

where $c_n = 4\sqrt{\frac{n(m-1)^3}{\beta_d\, m^3}}$. The last equality is obtained by applying (2.3). Clearly, $P_\theta(N \le n) = 1$ if

$c_n \geq 1$, which leads to the condition $n \geq u$. Thus, for integer valued n in $[k, u)$,

$$\begin{aligned} P_\theta(N \leq n) &= P_\theta\left(\frac{V_{1k}-V_{2k}}{V_{1k}+V_{2k}} \geq \sqrt{1-c_n}\right) + P_\theta\left(\frac{V_{1k}-V_{2k}}{V_{1k}+V_{2k}} \leq -\sqrt{1-c_n}\right) \\ &= P_\theta\left(\frac{V_{2k}}{V_{1k}} \leq c_{1n}\right) + P_\theta\left(\frac{V_{2k}}{V_{1k}} \geq c_{2n}\right) \\ &= P_\theta\left(F_{k(m-1),k} \leq \frac{c_{1n}\left\{1+(m-1)\rho\right\}}{(m-1)(1-\rho)}\right) + P_\theta\left(F_{k(m-1),k} \geq \frac{c_{2n}\left\{1+(m-1)\rho\right\}}{(m-1)(1-\rho)}\right). \end{aligned}$$

The last equality is obtained since $V_{1k} \sim \sigma^2\left(1+(m-1)\rho\right)\chi^2_k$, $V_{2k} \sim \sigma^2\left(1-\rho\right)\chi^2_{k(m-1)}$, and V_{1k} and V_{2k} are independent random variables. Hence, the proof is complete. $\square$

Note that in Theorem 3.1, $F_{N,\theta}(.)$ involves only ρ. Once the cumulative distribution function of N is obtained, it is easy to find $E_\theta(N)$ and $V_\theta(N)$. Exact values of these functionals for some selected cases are provided in Tables 2 and 3.

Remark 3.1. A natural upper bound for the choice of pilot sample size k is u. If one chooses $k > u$, then $c_k > 1$, and hence, $P_\theta(N \leq k) = 1$ from the proof of Theorem 3.1. This means that the two-stage procedure will stop in the first stage with probability 1, which makes the procedure a fixed-sample-size procedure. Hence, for all exact computations, we will consider $k < u$.

Next, we derive the exact formula for the distribution function of $\widehat{\rho}_N$. For any positive integer n, let us denote the cumulative distribution functions of V_{in} and V^*_{in} by $F_{V_{in}}(.)$ and $F_{V^*_{in}}(.)$ respectively for $i = 1, 2$. For any $x \in \left[\frac{-1}{m-1},\, 1\right]$, the distribution function of $\widehat{\rho}_N$, denoted by $F_{\widehat{\rho}_N}(.)$ can be written as

$$F_{\widehat{\rho}_N}(x) = P_\theta(\widehat{\rho}_N \leq x) = P_\theta(g\left(\widehat{\rho}_N\right) \geq g(x)) = P_\theta\left(\frac{V_{2N}}{V_{1N}} \geq g(x)\right) \text{ where } g(x) = \frac{(m-1)(1-x)}{1+(m-1)x}.$$

The second equality is obtained since $g(x)$ is a decreasing function of x. Now,

$$P_\theta\left(\frac{V_{2N}}{V_{1N}} \geq g(x)\right) = I_1 + I_2, \text{ where } I_1 = P_\theta\left(\frac{V_{2k}}{V_{1k}} \geq g(x),\, N = k\right) \text{ and } I_2 = P_\theta\left(\frac{V_{2N}}{V_{1N}} \geq g(x),\, N > k\right).$$

Working with I_1, we note from the proof of Theorem 3.1 that

$$\begin{aligned}
I_1 &= P_\theta\left(\frac{V_{2k}}{V_{1k}} \geq g(x),\ K^* \leq k\right) \\
&= P_\theta\left(\frac{V_{2k}}{V_{1k}} \geq g(x),\ \left|\frac{V_{1k}-V_{2k}}{V_{1k}+V_{2k}}\right| \geq \sqrt{1-c_k}\right) \\
&= P_\theta\left(\frac{V_{2k}}{V_{1k}} \geq g(x),\ \frac{V_{1k}-V_{2k}}{V_{1k}+V_{2k}} \geq \sqrt{1-c_k}\right) + P_\theta\left(\frac{V_{2k}}{V_{1k}} \geq g(x),\ \frac{V_{1k}-V_{2k}}{V_{1k}+V_{2k}} \leq -\sqrt{1-c_k}\right) \\
&= P_\theta\left(\frac{V_{2k}}{V_{1k}} \geq g(x),\ \frac{V_{2k}}{V_{1k}} \leq c_{1k}\right) + P_\theta\left(\frac{V_{2k}}{V_{1k}} \geq g(x),\ \frac{V_{2k}}{V_{1k}} \geq c_{2k}\right) \\
&= P_\theta\left(\frac{g(x)}{g(\rho)} \leq F_{k(m-1),\,k} \leq \frac{c_{1k}}{g(\rho)}\right) + P_\theta\left(F_{k(m-1),\,k} \geq \frac{\max\{g(x), c_{2k}\}}{g(\rho)}\right) \\
&= G_k\left(\frac{c_{1k}}{g(\rho)}\right) - G_k\left(\frac{g(x)}{g(\rho)}\right) + 1 - G_k\left(\frac{\max\{g(x), c_{2k}\}}{g(\rho)}\right). \qquad (3.5)
\end{aligned}$$

Now, working with I_2, we have:

$$\begin{aligned}
I_2 &= P_\theta\left(\frac{V_{2N}}{V_{1N}} \geq g(x),\ K^* > k\right) \\
&= P_\theta\left(\frac{V_{2k}+V^*_{2(K^*-k)}}{V_{1k}+V^*_{1(K^*-k)}} \geq g(x),\ \left|\frac{V_{1k}-V_{2k}}{V_{1k}+V_{2k}}\right| < \sqrt{1-c_k}\right) \\
&= P_\theta\left(\frac{V_{2k}+V^*_{2(K^*-k)}}{V_{1k}+V^*_{1(K^*-k)}} \geq g(x),\ c_{1k} < \frac{V_{2k}}{V_{1k}} < c_{2k}\right) \\
&= \int_0^\infty \int_{v_1 c_{1k}}^{v_1 c_{2k}} P\left(V^*_{2j} \geq g(x)(v_1 + V^*_{1j}) - v_2\right) dF_{V_{2k}}(v_2) dF_{V_{1k}}(v_1). \qquad (3.6)
\end{aligned}$$

The last equality is obtained by conditioning on $(V_{1k}, V_{2k}) = (v_1, v_2)$ where j is the value of $K^* - k$ when $(V_{1k}, V_{2k}) = (v_1, v_2)$, that is,

$$j \equiv j(v_1, v_2) = \left\lfloor \frac{\beta_d}{m(m-1)}\left(1 - \frac{v_1 - v_2/(m-1)}{v_1+v_2}\right)^2 \left(1 + (m-1)\frac{v_1 - v_2/(m-1)}{v_1+v_2}\right)^2 \right\rfloor + 1 - k.$$

Since $\left\{c_{1k} < \frac{V_{2k}}{V_{1k}} < c_{2k}\right\} = \{K^* - k > 0\}$, $j = j(v_1, v_2) > 0$ on the set $\left\{(v_1, v_2) : \frac{v_2}{v_1} \in (c_{1k}, c_{2k})\right\}$. In (3.6), we also used independence of (V_{1k}, V_{2k}) and (V^*_{1j}, V^*_{2j}). Finally, by conditioning on $V^*_{1j} = y$ and using the independence of V^*_{1j} and V^*_{2j}, we obtain

$$I_2 = \int_0^\infty \int_{v_1 c_{1k}}^{v_1 c_{2k}} \int_0^\infty \left[1 - F_{V^*_{2j}}\left(g(x)(v_1 + y) - v_2\right)\right] dF_{V^*_{1j}}(y)\, dF_{V_{2k}}(v_2)\, dF_{V_{1k}}(v_1). \qquad (3.7)$$

The cumulative distribution function of $\widehat{\rho}_N$ is obtained as $F_{\widehat{\rho}_N}(x) = I_1 + I_2$ for $x \in \left[\frac{-1}{m-1}, 1\right]$. The exact coverage probability yielded by the fixed-width $(1-\alpha)$ confidence interval J_N is

$$P_\theta\left(\widehat{\rho}_N - d \leq \rho \leq \widehat{\rho}_N + d\right) = F_{\widehat{\rho}_N}(\rho + d) - F_{\widehat{\rho}_N}(\rho - d). \qquad (3.8)$$

It is interesting to note that the distribution of $\widehat{\rho}_N$ does not depend on the true value of σ^2. In Tables 2 and 3, we will present the exact coverage probabilities for different values of ρ, k, m and show that they are very close to the target level $1-\alpha$.

3.4. Choice of Pilot Sample Size and Asymptotic Properties

Performance of two-stage procedures depend significantly on the choice of a pilot sample size k. If k is too small, then we may obtain a poor estimate K^* of the optimal sample size n^*, but if k is too large, then we may oversample too much leading to unnecessarily high cost of the procedure.

In this section, we will establish some desirable asymptotic properties of the proposed two-stage procedure, formally known as *asymptotic consistency* and *asymptotic first-order efficiency* properties, when the fixed width $2d$ is very small, that is, as $d \downarrow 0$. Efficiency of final sample size (stopping variable) N means that the expected sample size should be asymptotically same as the optimal sample size, and consistency of a confidence interval means that it should attain the target confidence level asymptotically, in this case as $d \downarrow 0$. These asymptotic properties were first considered by Chow and Robbins (1965) and further refined by Ghosh and Mukhopadhyay (1981).

Let us denote the optimal sample size n^* and the stopping variable N, as functions of d, by $n^*(d)$ and $N(d)$ respectively. As we note from (3.3) that $n^*(d) \to \infty$ as $d \downarrow 0$, it becomes evident that the pilot sample size k must be chosen as a function of d that also tends to ∞ as $d \downarrow 0$ in order to perform any asymptotic analysis. Furthermore, we show in the following Theorem 3.2 that asymptotic optimality properties can be obtained if we choose $k(d)$ such that the convergence rate of $k(d)$ to ∞ is slower than that of $n^*(d)$. Now, let us establish the well-known *asymptotic first-order efficiency* property of our two-stage methodology, which states that the average final sample size $E(N(d))$ and the unknown optimal sample size $n^*(d)$ are asymptotically equivalent.

Theorem 3.2. *For any choice of pilot sample size* $k(d)$ *such that* $k(d) \to \infty$ and $k(d) = o(n^*(d))$, *the proposed two-stage procedure yields*:

(*i*) $\lim_{d\downarrow 0} N(d)/n^*(d) = 1$ *a.s.*;

(*ii*) $\lim_{d\downarrow 0} E[N(d)]/n^*(d) = 1$ [*Asymptotic first-order efficiency*].

Proof. Let us define the function $h(\rho) = (1-\rho)^2(1+(m-1)\rho)^2$. From the definition of $K^*(d)$ in (3.4), we can write

$$\frac{\beta_d\, h\left(\widehat{\rho}_{k(d)}\right)}{m(m-1)} \le K^*(d) \le \frac{\beta_d\, h\left(\widehat{\rho}_{k(d)}\right)}{m(m-1)} + 1.$$

Dividing all sides of the above inequalities by $n^*(d)$ and applying $\frac{\beta_d\, h(\rho)}{m(m-1)} \le n^*(d) \le \frac{\beta_d\, h(\rho)}{m(m-1)} + 1$ appropriately, we have

$$\frac{h\left(\widehat{\rho}_{k(d)}\right)}{h(\rho) + m(m-1)/\beta_d} \le \frac{K^*(d)}{n^*(d)} \le \frac{h\left(\widehat{\rho}_{k(d)}\right)}{h(\rho)} + \frac{1}{n^*(d)}. \tag{3.9}$$

Since $\widehat{\rho}_n$ is a strong consistent estimator of ρ, taking limit as $d \downarrow 0$ on all sides of the above inequalities, we obtain $\lim_{d\downarrow 0} \frac{K^*(d)}{n^*(d)} = 1$ almost surely. The definition of $N(d)$ yields

$$K^*(d) \le N(d) \le K^*(d) + k(d). \tag{3.10}$$

Now, dividing all sides of the above inequalities by $n^*(d)$, taking limit as $d \downarrow 0$ and using $k(d)/n^*(d) \to 0$, we complete the proof of part *(i)* of Theorem 3.2.

To prove the second part, note that the maximum value of the function $h(\rho)$ in $\left(\frac{-1}{m-1}, 1\right)$ is attained at $\rho = \rho_0 = \frac{m-2}{2(m-1)}$. Since $\widehat{\rho}_n$ is the maximum likelihood estimator of ρ, for all $n = 1, 2, \dots,$ $\widehat{\rho}_n$ lies in $\left(\frac{-1}{m-1}, 1\right)$ with probability 1. Hence, using (3.9), we have

$$\sup_{d>0} \frac{K^*(d)}{n^*(d)} \le \sup_{d>0} \frac{h\left(\widehat{\rho}_{k(d)}\right)}{h(\rho)} + 1 \le \frac{h(\rho_0)}{h(\rho)} + 1 = \frac{m^4}{16\, h(\rho)\, (m-1)^2} + 1 \ \text{ a.s.}$$

Thus, the Lebesgue dominated convergence theorem yields

$$\lim_{d\downarrow 0} E\left[K^*(d)/n^*(d)\right] = 1.$$

Finally, dividing all sides of (3.10) by $n^*(d)$, taking expectation and then limit as $d \downarrow 0$, and using $k(d) = o(n^*(d))$, we complete the proof of asymptotic first-order efficiency. □

Remark 3.2. We will show that one can find pilot sample size $k(d)$ satisfying the conditions of Theorem 3.2, that is, $k(d) \to \infty$ and $k(d) = o(n^*(d))$ as $d \downarrow 0$. Consider an auxiliary stopping variable in the spirit of Mukhopadhyay (1980) as

$$T = \inf\left\{ n \ge 1 : n \ge \frac{\beta_d (1 - \widehat{\rho}_n + n^{-r})^2 \left\{1 + (m-1)\left(\widehat{\rho}_n + n^{-r}\right)\right\}^2}{m(m-1)} \right\} \ \text{ for some } r > 0.$$

where $r > 0$. The auxiliary variable T is constructed in a manner to estimate the optimal sample size n^* in a purely sequential sampling procedure where $1 - \widehat{\rho}_n$ and $1 + (m-1)\widehat{\rho}_n$ are replaced by $1 - \widehat{\rho}_n + n^{-r}$ and $1 + (m-1)\left(\widehat{\rho}_n + n^{-r}\right)$ respectively. This additional n^{-r} term is added appropriately because in some extreme cases $1 - \widehat{\rho}_n$ and $1 + (m-1)\widehat{\rho}_n$ may be close to 0 leading to very early stopping. Noting that $1 - \widehat{\rho}_T \ge 0$ and $1 + (m-1)\widehat{\rho}_T \ge 0$ almost surely, we observe the following lower bounds of T

$$T \ge \beta_d T^{-4r} \left(\frac{m-1}{m}\right), \text{ i.e., } T \ge \left[\frac{\beta_d (m-1)}{m}\right]^{1/(1+4r)} \ \text{ a.s.}$$

Thus, one choice of pilot sample size for asymptotic analysis is

$$k'(d) = \left\lfloor \left(\beta_d \frac{m-1}{m}\right)^{1/(1+4r)} \right\rfloor + 1 \ \text{ for some } r > 0.$$

One can easily check that $k'(d) \uparrow \infty$ and $k'(d)/n^*(d) \to 0$ as $d \downarrow 0$ as required in Theorem 3.2. Remark 3.1 suggests that there is a natural upper bound $u \equiv u(d) = \frac{\beta_d\, m^3}{16(m-1)^3}$ for pilot sample size. Taking into account of this, one can also choose

$$k(d) = \left\lfloor \min\left\{ \frac{\beta_d\, m^3}{16(m-1)^3},\ \left(\beta_d \frac{m-1}{m}\right)^{1/(1+4r)} \right\} \right\rfloor + 1. \tag{3.11}$$

as a pilot sample size. Since $k(d)/n^*(d) \le k'(d)/n^*(d) + 1/n^*(d) \to 0$ as $d \downarrow 0$, $k(d)$ also satisfies the conditions of Theorem 3.2.

Next, we will show that our proposed two-stage procedure is *asymptotically consistent*, that is, the attained coverage probability tends to the target level of significance $(1-\alpha)$ as $d \downarrow 0$. In order to establish this property, let us first observe the following facts and prove necessary auxiliary results. We define

$$U_{1n} = \frac{1}{n}\sum_{i=1}^{n} Y_{i1}^2 = \frac{V_{1n}}{n} \quad \text{and} \quad U_{2n} = \frac{1}{n(m-1)}\sum_{i=1}^{n}\sum_{j=2}^{m} Y_{ij}^2 = \frac{V_{2n}}{n(m-1)}$$

and note that U_{1n} is a sample mean of i.i.d. random variables $Y_{i1}^2 \sim \sigma^2\left(1+(m-1)\rho\right)\chi_1^2$ and U_{2n} is a sample mean of i.i.d. random variables $Y_{i0} \sim \sigma^2 \frac{(1-\rho)}{m-1}\chi_{m-1}^2$ for $i = 1, \dots, n$, where $Y_{i0} = \frac{1}{m-1}\sum_{j=2}^{m} Y_{ij}^2$. For $i = 1, \dots, n$, let us denote

$$\begin{aligned} &\mu_1 = E(U_{1n}) = \sigma^2\left(1+(m-1)\rho\right),\ \mu_2 = E(U_{2n}) = \sigma^2\left(1-\rho\right), \\ &\sigma_1^2 = V(Y_{i1}^2) = 2\sigma^4(1+(m-1)\rho)^2,\ \sigma_2^2 = V(Y_{i0}) = \frac{2\sigma^4}{(m-1)}(1-\rho)^2. \end{aligned} \tag{3.12}$$

The following lemma is needed to prove the asymptotic consistency property of our proposed two-stage procedure.

Lemma 3.1. *Let* $\boldsymbol{W}_n = \sqrt{n}\,(U_{1n}-\mu_1,\, U_{2n}-\mu_2)$*. For any choice of pilot sample size* $k(d)$ *such that* $k(d) \to \infty$ *and* $k(d) = o(n^*(d))$ *as* $d \downarrow 0$*,* $\boldsymbol{W}_{N(d)}$ *converges in distribution to* $N_2(\mathbf{0}, \Sigma)$ *where the covariance matrix* $\Sigma = diag\,(\sigma_1^2,\, \sigma_2^2)$*.*

Proof. Since $\left(Y_{i1}^2, Y_{i0}\right)$, $i = 1, \dots, n$, are i.i.d. random vectors with mean (μ_1, μ_2) and covariance $\Sigma = diag\,(\sigma_1^2,\, \sigma_2^2)$, the bivariate central limit theorem yields: $\boldsymbol{W}_n \xrightarrow{\mathcal{L}} N_2(\mathbf{0},\, \Sigma)$ as $n \to \infty$. To prove Lemma 3.1, we need to show, by Cramér-Wold theorem (see Cramér and Wold, 1936):

$$\boldsymbol{t}^T \boldsymbol{W}_{N(d)} \xrightarrow{\mathcal{L}} N(0,\, \boldsymbol{t}^T \Sigma\, \boldsymbol{t}) \text{ for all } \boldsymbol{t} \in \mathbb{R}^2.$$

Fix some $\boldsymbol{t} = (t_1, t_2) \in \mathbb{R}^2$ and write $\boldsymbol{t}^T \boldsymbol{W}_{N(d)} = \boldsymbol{t}^T \boldsymbol{W}_{n^*(d)} + \left(\boldsymbol{t}^T \boldsymbol{W}_{N(d)} - \boldsymbol{t}^T \boldsymbol{W}_{n^*(d)}\right)$. Since we know that $\boldsymbol{t}^T \boldsymbol{W}_{n^*(d)} \xrightarrow{\mathcal{L}} N(0,\, \boldsymbol{t}^T \Sigma\, \boldsymbol{t})$ as $d \downarrow 0$, by Slutsky's theorem it is enough to show that

$$\boldsymbol{t}^T \boldsymbol{W}_{N(d)} - \boldsymbol{t}^T \boldsymbol{W}_{n^*(d)} \xrightarrow{P} 0 \text{ as } d \downarrow 0.$$

For convenience, let us now omit d from $N(d)$ and $n^*(d)$. Note that

$$\begin{aligned}
\boldsymbol{t}^T\left(\boldsymbol{W}_N-\boldsymbol{W}_{n^*}\right) &= t_1\sqrt{N}\left((U_{1N}-\mu_1)-\sqrt{\frac{n^*}{N}}(U_{1n^*}-\mu_1)\right)+t_2\sqrt{N}\left((U_{2N}-\mu_2)-\sqrt{\frac{n^*}{N}}(U_{2n^*}-\mu_2)\right)\\
&= t_1\sqrt{N}(U_{1N}-U_{1n^*})-t_1\sqrt{N}\left(\sqrt{\frac{n^*}{N}}-1\right)(U_{1n^*}-\mu_1)\\
&\quad +\ t_2\sqrt{N}(U_{2N}-U_{2n^*})-t_2\sqrt{N}\left(\sqrt{\frac{n^*}{N}}-1\right)(U_{2n^*}-\mu_2)\\
&= t_1\sqrt{N}(U_{1N}-U_{1n^*})+t_2\sqrt{N}(U_{2N}-U_{2n^*})-\left(1-\sqrt{\frac{n^*}{N}}\right)\boldsymbol{t}^T\boldsymbol{W}_{n^*}. \qquad (3.13)
\end{aligned}$$

A sequence of random variables $\{T_n\}$ satisfies *uniform continuity in probability (u.c.i.p.)* condition of Anscombe (1952) for some sequence $\{w_n\}$ of positive real numbers, if for any arbitrarily small $\epsilon > 0$ and $\eta > 0$, there exists small $\delta > 0$ such that

$$\limsup_{n} P\left(\max_{n(1-\delta)\le n'\le n(1+\delta)}|T_n-T_{n'}|>\epsilon w_n\right)<\eta.$$

This tightness condition can be found in Anscombe (1952), Ghosh et al. (1997), Gut (2009), Mukhopadhyay and Chattopadhyay (2012), Sproule (1974), among others.

Now, fix some arbitrarily small $\epsilon > 0$ and define $n_{1\delta} = (1-\delta)n^*$ and $n_{2\delta} = (1+\delta)n^*$ for some $\delta > 0$ as in the *u.c.i.p.* condition above. Since U_{1n} and U_{2n} are sample means of i.i.d. random variables, following Anscombe (1952, p. 602), U_{1n} and U_{2n} satisfy the *u.c.i.p.* condition with $w_n = n^{-1/2}$, which implies

$$\limsup_{d\downarrow 0} P_\theta\left(\max_{n_{1\delta}\le n\le n_{2\delta}}\sqrt{n}|U_{in^*(d)}-U_{in}|>\epsilon\right)<\epsilon \quad \text{for } i=1,2. \qquad (3.14)$$

Note that

$$\begin{aligned}
&P_\theta\left(|t_1\sqrt{N}(U_{1N}-U_{1n^*})+t_2\sqrt{N}(U_{2N}-U_{2n^*})|>\epsilon\right)\\
&\le P_\theta\left(|t_1\sqrt{N}(U_{1N}-U_{1n^*})+t_2\sqrt{N}(U_{2N}-U_{2n^*})|>\epsilon,\ |N-n^*|\le\delta n^*\right)+P_\theta(|N-n^*|>\delta n^*)\\
&\le \sum_{i=1,2}P_\theta\left(\max_{n_{1\delta}\le n\le n_{2\delta}}\sqrt{n}|U_{in^*}-U_{in}|>\frac{\epsilon}{2|t_i|}\right)+P_\theta\left(\left|\frac{N}{n^*}-1\right|>\delta\right).
\end{aligned}$$

Taking $\limsup_{d\downarrow 0}$, using (3.14) and the fact that $N(d)/n^*(d) \xrightarrow{P} 1$ as $d\downarrow 0$, we have

$$\limsup_{d\downarrow 0} P_\theta\left(|t_1\sqrt{N}(U_{1N}-U_{1n^*})+t_2\sqrt{N}(U_{2N}-U_{2n^*})|>\epsilon\right)\le\frac{\epsilon}{2}(|t_1|^{-1}+|t_2|^{-1}).$$

Since $\epsilon > 0$ is arbitrary,

$$t_1\sqrt{N(d)}(U_{1N(d)} - U_{1n^*(d)}) + t_2\sqrt{N(d)}(U_{2N(d)} - U_{2n^*(d)}) \xrightarrow{P} 0 \quad \text{as } d \downarrow 0.$$

Note that $\left(1 - \sqrt{\frac{n^*(d)}{N(d)}}\right) \boldsymbol{t}^T \boldsymbol{W}_{n^*(d)} \xrightarrow{P} 0$ since $n^*(d)/N(d) \xrightarrow{P} 1$ and $\boldsymbol{t}^T\boldsymbol{W}_{n^*(d)} \xrightarrow{\mathcal{L}} N(0,\, \boldsymbol{t}^T\Sigma\,\boldsymbol{t})$. Hence, we complete the proof of Lemma 3.1 by applying Slutsky's theorem in (3.13). □

Now, we are ready to prove the asymptotic consistency property of the proposed two-stage procedure with the help of the above Lemma 3.1.

Theorem 3.3. *For any choice of pilot sample size $k(d)$ such that $k(d) \to \infty$ and $k(d) = o(n^*(d))$ as $d \downarrow 0$, the proposed two-stage procedure yields*

$$\limsup_{d\downarrow 0} P\left(\widehat{\rho}_{N(d)} - d \le \rho \le \widehat{\rho}_{N(d)} + d\right) = 1 - \alpha \;\; [\textit{asymptotic consistency}].$$

Proof. Let us define the function

$$M(u_1, u_2) = \frac{u_1 - u_2}{u_1 + (m-1)u_2}, \text{ and observe that } \widehat{\rho}_n = M(U_{1n}, U_{2n}) \text{ and } \rho = M(\mu_1, \mu_2),$$

where μ_1 and μ_2 are given in (3.12). Let us define $\boldsymbol{U}_n = (U_{1n}, U_{2n})$ and $\boldsymbol{\mu} = (\mu_1, \mu_2)$. Since from Lemma 3.1, we know that $\sqrt{N(d)}\left(U_{1N(d)} - \mu_1,\, U_{2N(d)} - \mu_2\right) \xrightarrow{\mathcal{L}} N_2(\mathbf{0},\, \Sigma)$, the bivariate delta method gives

$$\sqrt{N(d)}(M(\boldsymbol{U}_{N(d)}) - M(\boldsymbol{\mu})) \xrightarrow{\mathcal{L}} N_2\left(\mathbf{0},\, [\nabla M(\boldsymbol{\mu})]^T \Sigma\, [\nabla M(\boldsymbol{\mu})]\right) \text{ as } d \downarrow 0,$$

where

$$\nabla M(\boldsymbol{\mu}) = \left(\frac{\partial M}{\partial u_1}, \frac{\partial M}{\partial u_2}\right)^T \Bigg|_{(u_1,u_2)=(\mu_1,\,\mu_2)} = (M_1(\boldsymbol{\mu}), M_2(\boldsymbol{\mu}))^T,$$

with

$$M_1(\boldsymbol{\mu}) = \frac{m\mu_2}{(\mu_1 + (m-1)\mu_2)^2} \quad \text{and} \quad M_2(\boldsymbol{\mu}) = \frac{-m\mu_1}{(\mu_1 + (m-1)\mu_2)^2}.$$

Algebraic calculations yield $[\nabla M(\boldsymbol{\mu})]^T \Sigma\, [\nabla M(\boldsymbol{\mu})] = \sigma_\rho^2$, where σ_ρ^2 is given in (2.4). Thus, we have proved

$$\sqrt{N(d)}\left(\widehat{\rho}_{N(d)} - \rho\right) \xrightarrow{\mathcal{L}} N(0, \sigma_\rho^2) \text{ as } d \downarrow 0. \tag{3.15}$$

From (3.3) and (2.4), we have

$$\sigma_\rho\, z_{\alpha/2} \le d\sqrt{n^*(d)} \le \sqrt{\sigma_\rho^2 z_{\alpha/2}^2 + d^2}, \text{ which implies } \limsup_{d\downarrow 0} d\sqrt{n^*(d)} = \sigma_\rho\, z_{\alpha/2},$$

and hence, $\limsup_{d\downarrow 0} d\sqrt{N(d)} = \left(\limsup_{d\downarrow 0} \sqrt{N(d)/n^*(d)}\right)\left(\limsup_{d\downarrow 0} d\sqrt{n^*(d)}\right) = \sigma_\rho\, z_{\alpha/2}$ a.s. from part (i) of Theorem 3.2. Finally, using the asymptotic normality of $\widehat{\rho}_{N(d)}$, we have the asymptotic coverage probability

$$\limsup_{d\downarrow 0} P\left(\widehat{\rho}_{N(d)} - d \le \rho \le \widehat{\rho}_{N(d)} + d\right) = 2\limsup_{d\downarrow 0} \Phi\left(\frac{d\sqrt{N(d)}}{\sigma_\rho}\right) - 1 = 1 - \alpha.$$

The proof is now complete. □

Remark 3.3. In addition to possessing the first-order efficiency and consistency properties, the two-stage procedure provides the final estimator $\widehat{\rho}_{N(d)}$ which is *asymptotically unbiased*. That is, $E[\widehat{\rho}_{N(d)}] \to \rho$ as $d \downarrow 0$. This can be easily shown by noting that $|\widehat{\rho}_{N(d)}| \le 1$ a.s., $\limsup_{d\downarrow 0} \widehat{\rho}_{N(d)} = \rho$ a.s. (see Gut, 2009), and applying the Lebesgue dominated convergence theorem.

3.5. Numerical Studies for Fixed-Width Interval

In this section, we evaluate the performance of the proposed two-stage procedure for constructing a fixed-width confidence interval for the common correlation ρ in the presence of nuisance parameter σ^2. It was noted earlier that the performance of the two-stage method did not depend on σ^2. Therefore, we fixed $\sigma^2 = 1$ for all numerical studies of this section. We mainly investigated the following aspects of the procedure.

1. How to compute the *exact* probability distributions of the stopping variable N and the *exact* distribution of $\widehat{\rho}_N$ numerically. We validate all exact formulas for cumulative distribution function, expectation and standard deviation of N, and the exact formula for coverage probability via extensive simulation study.

2. We verify whether the actual coverage probability attained by the two-stage procedure is close enough to the target level of significance $1-\alpha$ for any prescribed $\alpha \in (0,1)$.

3. We check how the overall performance of the procedure changes with the choice of pilot sample size k, and also with different values of m, ρ, α, and d.

4. The asymptotic first-order efficiency and consistency properties were proved in Theorems 3.2 and 3.3. For fixed small $d > 0$, we measure the closeness between the ratio of expected and optimal sample size and 1, and similarly, the closeness between the attained coverage probability and $1-\alpha$ to confirm the validity of the asymptotic results.

In all simulations, we used $M = 10^5$ replications. For exact computation of the cumulative distribution function of $\widehat{\rho}_N$, the triple integral I_2 in (3.7) was computed numerically using the "*integrate*" function from "R" which uses an adaptive Gauss–Kronrod quadrature method. If the integrate function returned an error, we applied Gauss's quadrature method (see Rade and Westergren, 1990 and Abramowitz and Stegun, 1965) as an alternative.

In all the tables, $E(N)$ and $\widehat{E}(N)$ represent the exact and the simulated values (sample average of M observations) of the expectation of N respectively, $SD(N)$ and $\widehat{SD}(N)$ denote the exact and the simulated

Table 2. Comparison of the exact values and the simulated values of expectation, standard deviation, and coverage probability of the two-stage procedure for fixed-width interval estimation of ρ with $\sigma = 1$, $\alpha = 0.1$, and $d = 0.03$

m	ρ	n^*	k	$E(N)$	$\widehat{E}(N)$	$SD(N)$	$\widehat{SD}(N)$	CP	$\widehat{CP}$
3	-0.2	520	50	525.67	525.92	149.15	149.83	0.8870	0.8848
			100	522.94	523.41	106.74	107.11	0.8923	0.8905
			150	521.98	521.92	87.49	87.89	0.8924	0.8919
3	0	1003	50	982.95	983.02	155.84	155.97	0.8918	0.8908
			100	992.64	993.03	112.84	112.89	0.8960	0.8931
			150	995.92	996.13	92.89	93.02	0.8958	0.8952
3	0.7	520	50	537.71	537.75	152.37	152.72	0.8917	0.8920
			100	529.08	528.85	108.02	108.38	0.8947	0.8942
			150	526.09	525.68	88.22	88.53	0.8960	0.8954
5	-0.1	131	30	136.57	136.60	58.30	58.46	0.8619	0.8639
			50	134.61	134.52	44.86	44.95	0.8713	0.8714
			70	133.89	133.79	37.52	37.64	0.8785	0.8755
5	0	301	50	301.60	301.59	78.62	78.94	0.8855	0.8844
			100	301.39	301.24	56.31	56.33	0.8926	0.8890
			150	301.30	301.09	46.17	46.06	0.8918	0.8901
5	0.7	391	50	400.02	400.07	88.89	88.96	0.8962	0.8948
			100	395.78	395.51	63.62	63.74	0.8964	0.8970
			150	394.28	394.17	52.13	52.26	0.8970	0.8965
10	0.05	127	30	128.93	128.80	48.13	48.29	0.8649	0.8651
			50	128.32	128.30	37.50	37.58	0.8764	0.8733
			70	128.15	128.14	31.55	31.57	0.8772	0.8768
10	0.25	397	50	389.33	389.51	57.11	57.17	0.8912	0.8894
			100	393.30	393.36	40.90	40.99	0.8972	0.8939
			150	394.65	394.67	33.54	33.56	0.8952	0.8940
10	0.7	321	50	326.02	325.94	59.98	60.02	0.8964	0.8959
			100	323.59	323.54	43.21	43.30	0.8992	0.8981
			150	322.72	322.74	35.49	35.54	0.8980	0.8973

values (sample standard deviations of M observations) of the standard deviation of N respectively, and CP and $\widehat{CP}$ represent the exact and the simulated values of attained coverage probability respectively.

Table 2 compares the exact and the simulated values of expectations and standard deviations of N and

Table 3. Comparison of the exact values and the simulated values of expectation, standard deviation, and coverage probability of the two-stage procedure for fixed-width interval estimation of ρ with $\sigma = 1$, $\rho = 0.2$, and $m = 6$

α	d	n^*	k	$E(N)$	$\widehat{E}(N)$	$SD(N)$	$\widehat{SD}(N)$	CP	$\widehat{CP}$
0.1	0.025	739	50	724.05	724.27	104.89	104.92	0.8920	0.8907
			100	731.54	732.02	75.22	75.39	0.8964	0.8952
			150	734.10	734.15	61.71	61.86	0.8963	0.8953
	0.05	185	50	181.39	181.39	26.22	26.25	0.8883	0.8867
			100	183.26	183.19	18.81	18.83	0.8938	0.8894
			150	184.04	184.01	15.07	15.12	0.8893	0.8891
0.05	0.025	1049	50	1027.82	1027.62	148.93	149.15	0.9437	0.9436
			100	1038.47	1038.57	106.80	107.11	0.9471	0.9466
			200	1043.93	1044.03	76.06	76.34	0.9469	0.9469
	0.05	263	50	257.33	257.29	37.23	37.26	0.9411	0.9396
			100	259.99	259.98	26.70	26.72	0.9453	0.9437
			150	260.90	260.81	21.91	21.96	0.9437	0.9433
0.01	0.025	1812	50	1774.88	1774.68	257.22	257.88	0.9871	0.9869
			100	1793.27	1792.98	184.47	184.99	0.9887	0.9883
			300	1805.86	1805.34	107.52	108.03	0.9890	0.9892
	0.05	453	50	444.09	444.11	64.31	64.38	0.9860	0.9850
			100	448.69	448.87	46.12	46.20	0.9880	0.9868
			150	450.26	450.33	37.83	37.90	0.9879	0.9873

coverage probabilities achieved by the two-stage procedure for different choices of m and ρ while α and d are fixed at 0.1 and 0.03 respectively. The numerically obtained exact values of expectation, standard deviation and coverage probabilities are almost same as the simulated estimates, which confirms that the numerical approximations of exact formulas are nearly accurate.

The achieved coverage probabilities are also close to the target level of 90%. Performance of the two-stage procedure depends on a choice of the pilot sample size k. Specifically, as k increases (from 50 to 150 or from 30 to 70), we see that the attained coverage probabilities get closer to 0.9, expected sample sizes get closer to the optimal sample sizes n^*, and standard deviations of N reduce.

In Table 3, we compare the exact and the simulated values of expectations and standard deviations of N and coverage probabilities for different choices of α and d fixing $m = 6$ and $\rho = 0.2$. The numerical results obtained here are very similar to those reported in Table 2. Additionally, we observe that the optimal and expected sample sizes decrease as d increases, and increase as α decreases. This phenomena naturally follows from the definition of n^* in (3.3).

Table 4 validates the assertion of asymptotic first-order efficiency and consistency properties stated in Theorems 3.2 and 3.3. Moreover, the last column of Table 4 illustrates that the bias of the final point estimator $\widehat{\rho}_{N(d)}$ tends to 0 as $d \downarrow 0$ which was observed in Remark 3.3. Here, the two-stage procedure is implemented with pilot sample size $k(d)$, given in (3.11) with $r = 0.14$.

Table 4. Validating asymptotic properties of the two-stage procedure for fixed-width interval estimation of ρ with $\sigma = 1$, $\alpha = 0.1$, $r = 0.14$, $\rho = 0.6$, and $m = 4$

d	$n^*(d)$	$k(d)$	$\left\lvert\frac{\widehat{E}(N(d))}{n^*(d)} - 1\right\rvert$	$\lvert\widehat{CP} - (1-\alpha)\rvert$	$\widehat{bias}\left(\widehat{\rho}_{N(d)}\right)$
0.1	57	47	0.019783	0.01248	0.002262
0.07	116	75	0.003240	0.00752	0.002074
0.03	625	220	0.001085	0.00238	0.000444
0.01	5657	900	0.000266	0.00001	0.000039

Figures 1 and 2 compare the empirical (simulated) and exact *probability mass functions* (PMFs) of N corresponding to $\rho = 0.2$ and $\rho = 0.7$ respectively fixing $m = 6$, $\sigma = 1$, $\alpha = 0.1$, and $d = 0.05$. Figures 3 and 4 compare the empirical and exact PMFs of N corresponding to $m = 4$ and $m = 10$ respectively fixing $\rho = 0.6$, $\sigma = 1$, $\alpha = 0.1$, and $d = 0.05$. In all these figures, the simulated and the exact PMFs match quite closely which validates the exact formula in Theorem 3.1.

In Figure 5, we plot the exact CDF of N for different values of ρ with $m = 4$, $\sigma = 1$, $k = 50$, $\alpha = 0.1$, and $d = 0.05$. We observe that the stopping variable N for $\rho = 0.33$ is *stochastically larger* than N corresponding to $\rho = 0.06$ and 0.6. This observation is not surprising since n^*, as a function of ρ, attains maximum at $\rho_0 = \frac{m-2}{2(m-1)}$, and thus, for $m = 4$, the optimal and the expected sample sizes tend to be smaller as ρ gets further away from $\rho_0 \approx 0.33$.

Figure 6 compares the exact CDFs of N for different values of m while $\rho = 0.6$, $\sigma = 1$, $k = 50$, $\alpha = 0.1$, and $d = 0.05$ are fixed. We observe that the stopping variable N corresponding to $m = 3$ is *stochastically larger* than that of $m = 4$, and N corresponding to $m = 4$ is *stochastically larger* than that of $m = 5$. This is expected since n^* is a decreasing function of m, and hence, both optimal and expected sample sizes decrease as m increase. The last two facts from Figures 5 and 6 are also observed in Table 2.

4. INTERVAL ESTIMATION WITH FIXED-ACCURACY APPROACH

Even though fixed-width confidence intervals enjoy the attractive property of having nearly any *prescribed precision* or *accuracy* of estimation in terms of having a fixed-width $2d$, they may suffer from some serious drawbacks in practice. In the problem of estimating common correlation ρ in a equi-correlated multivariate normal model, the parameter ρ and its maximum likelihood estimator $\widehat{\rho}_n$ belong to $\left(\frac{-1}{m-1}, 1\right)$ a.s.

For any fixed $d > 0$, there is a positive probability that the lower confidence limit $\widehat{\rho}_n - d \leq \frac{-1}{m-1}$ and/or the upper confidence limit $\widehat{\rho}_n + d \geq 1$. Depending on the true value of ρ and the prescribed value of d, these positive probabilities may be significantly high and the attained coverage probability may fall well short of

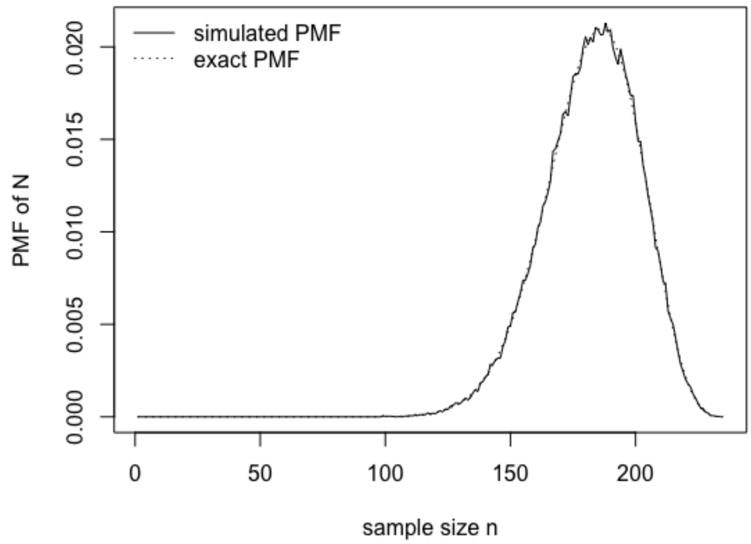

Figure 1. Comparison of empirical PMF and exact PMF of two-stage stopping variable N for $\rho = 0.2$, $m = 6$, $\sigma = 1$, $k = 100$, $\alpha = 0.1$, and $d = 0.05$.

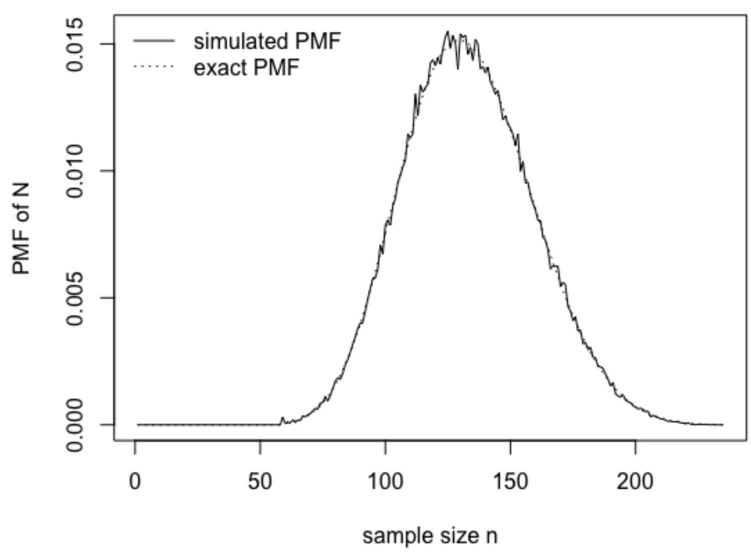

Figure 2. Comparison of empirical PMF and exact PMF of two-stage stopping variable N for $\rho = 0.7$, $m = 6$, $\sigma = 1$, $k = 60$, $\alpha = 0.1$, and $d = 0.05$.

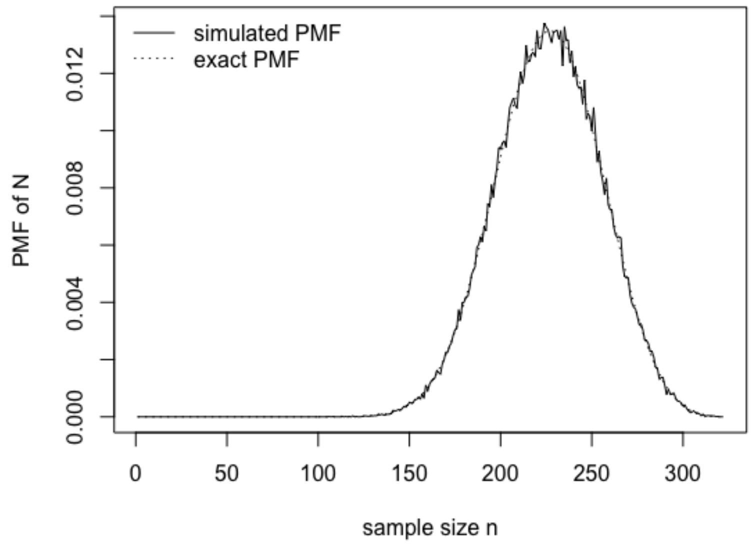

Figure 3. Comparison of empirical PMF and exact PMF of two-stage stopping variable N for $\rho = 0.6$, $m = 4$, $\sigma = 1$, $k = 100$, $\alpha = 0.1$, and $d = 0.05$.

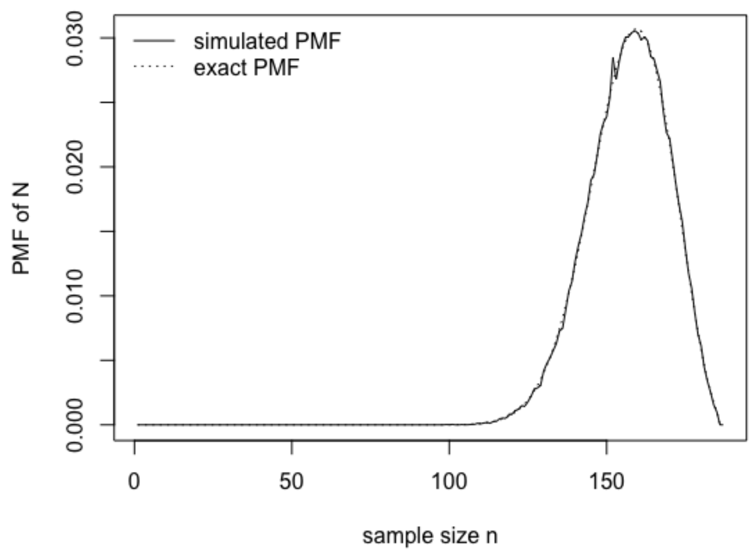

Figure 4. Comparison of empirical PMF and exact PMF of two-stage stopping variable N for $\rho = 0.6$, $m = 10$, $\sigma = 1$, $k = 100$, $\alpha = 0.1$, and $d = 0.05$.

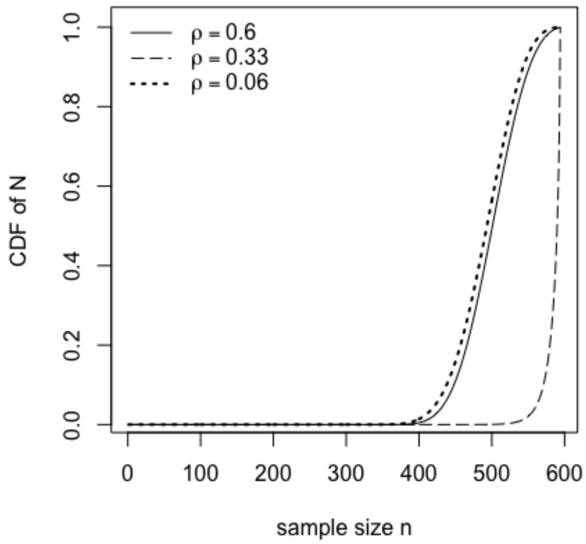

Figure 5. Comparison of the CDFs of two-stage stopping variable N for different values of ρ using $m = 4$, $\sigma = 1$, $k = 50$, $\alpha = 0.1$, and $d = 0.05$.

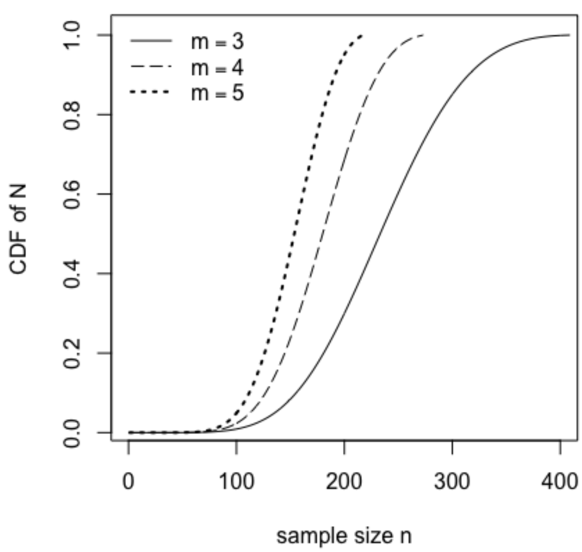

Figure 6. Comparison of the CDFs of two-stage stopping variable N for different values of m using $\rho = 0.6$, $\sigma = 1$, $k = 50$, $\alpha = 0.1$, and $d = 0.05$.

the target level of confidence $1-\alpha$. Chance of observing such phenomena is higher as the true value of ρ is closer to the boundaries of the parameter space $\left(\frac{-1}{m-1}, 1\right)$. For instance, if ρ happens to be 0.9 and $d = 0.05$, then there is non-negligible probability (≈ 0.086) that the upper confidence limit $\widehat{\rho}_n + d$ will exceed 1, and the coverage probability will be smaller than that required.

Table 5 illustrates few such cases where ρ is close to one of the boundaries of the parameter space and the *missed confidence limit*, defined as $misCL = P(\widehat{\rho}_n + d \geq 1 \cup \widehat{\rho}_n - d \leq \frac{-1}{m-1})$, does not appear negligible. This eventually causes the coverage probability to fall well short of the target level of confidence $1-\alpha$. Thus, it is difficult, in general, to choose the margin of error d appropriately unless the true value of ρ is known. So, a fundamentally different approach is called for.

In view of these difficulties, we propose an alternative approach of constructing a confidence interval for ρ by providing a *fixed-accuracy* confidence interval of a suitable transformation of ρ. The idea of fixed-accuracy confidence intervals was originally proposed by Mukhopadhyay and Banerjee (2014) and developed the approach further in Banerjee and Mukhopadhyay (2016), De and Mukhopadhyay (2015), and Mukhopadhyay and Banerjee (2015a,b).

Suppose that the parameter space for a generic parameter λ is $(0, \infty)$, and some estimator $\widehat{\lambda}_n$ of λ belongs to $(0, \infty)$ a.s. With a prefixed $\delta > 1$, the *fixed-accuracy confidence interval* for λ, denoted by $I_n = \left[\delta^{-1}\widehat{\lambda}_n, \, \delta\widehat{\lambda}_n\right]$, is such that

$$P_\lambda\left(\delta^{-1}\widehat{\lambda}_n \leq \lambda \leq \delta\widehat{\lambda}_n\right) \geq 1-\alpha \quad \text{for all } \lambda \text{ and given } \alpha \in (0,1). \tag{4.1}$$

Here, δ denotes the measure of fixed-accuracy to be prescribed by an experimenter.

Table 5. Illustrating some drawbacks of the fixed-width confidence intervals in terms of producing smaller coverage probabilities than desired for pilot sample size $k = 10$ for different choices of m, ρ, α, and d

m	ρ	$1-\alpha$	d	n^*	$\widehat{E}(N)$	$\widehat{CP}$	$misCL$
3	0.9	0.95	0.05	41	64.96	0.8971	0.0863
	−0.42	0.90	0.03	52	72.83	0.8059	0.0534
5	−0.18	0.95	0.03	47	58.46	0.8470	0.0779
	−0.14	0.90	0.04	43	50.42	0.8032	0.0371
10	0	0.95	0.04	54	58.93	0.8558	0.0325
	−0.03	0.90	0.03	38	43.51	0.7935	0.0369

As an alternative to the fixed-accuracy approach, however, one may also consider a prescribed proportional closeness approach discussed in Ehrenfeld and Littauer (1964, p. 339), Simons and Zacks (1967), Zacks (1966,2001,2009a,b), among others. For a parameter λ, not necessarily having its parameter space $(0, \infty)$, the idea is to find an estimator of λ that does not differ from the absolute value of λ by more than a certain percentage of λ with high probability. Formally, for fixed $\gamma \in (0,1)$ and $\alpha \in (0,1)$, an estimator of

λ, say $\widehat{\lambda}_n$, is said to have a *prescribed proportional closeness* $(\gamma, 1-\alpha)$ if

$$P_\lambda\left(|\widehat{\lambda}_n - \lambda| < \gamma|\lambda|\right) \geq 1-\alpha \quad \text{for all } \lambda. \tag{4.2}$$

Here, γ represents the measure of fixed proportional closeness to be prescribed by an experimenter. The probability in (4.2) measures the extent of proportional closeness of $\widehat{\lambda}_n$ to λ, and thus, the corresponding interval

$$\widetilde{I}_n = \left[(1+\gamma)^{-1}\,\widehat{\lambda}_n,\ (1-\gamma)^{-1}\,\widehat{\lambda}_n\right] \tag{4.3}$$

is called the *proportional closeness interval* estimator of λ (see Zacks, 2001). For *small* $\gamma > 0$, $(1-\gamma)^{-1} = 1+\gamma+\mathcal{O}(\gamma^2) \approx 1+\gamma$, which implies the interval $\widetilde{I}_n \approx \left[(1+\gamma)^{-1}\,\widehat{\lambda}_n,\ (1+\gamma)\,\widehat{\lambda}_n\right]$ where $(1+\gamma)$ plays the role of δ given in the fixed-accuracy interval I_n.

So it is interesting to note that the idea of fixed-accuracy interval may be closely related, but more general in the sense that unlike γ, δ does not have to lie between 0 and 1, to the idea of a prescribed proportional closeness interval for some parameter $\lambda \in (0,\infty)$. While this may sound alright on the surface, one still may not know how small $\gamma(>0)$ needs to be chosen in the face of having no idea about the magnitude of λ.

In the remaining sections of this article, we will consider the fixed-accuracy approach to construct bounded-width confidence intervals for ρ. Specifically, our strategy will be to first re-parameterize ρ to ρ^* by taking some suitable one-to-one transformation $g\colon (\frac{-1}{m-1}, 1) \to (0,\infty)$, then construct a $1-\alpha$ fixed-accuracy confidence interval

$$I_n = \left[\delta^{-1}\widehat{\rho}^*_n,\ \delta\widehat{\rho}^*_n\right] \text{ for } \rho^* = g(\rho),$$

and finally invert the interval I_n in order to come up with a $1-\alpha$ confidence interval for ρ by obtaining $J_n = g^{-1}(I_n)$. Observe that we use the notation $\widehat{\rho}^*_n$ to denote an estimator of ρ^* based upon n multivariate observation vectors. Then, we would show in the sequel that the width of J_n is bounded above by a function of δ which is less than 1. In what follows, we consider a class of reasonable transformations and develop the associated theory and methodology for constructing such bounded-width $1-\alpha$ confidence intervals for ρ.

4.1. A Class of Transformations Leading to Bounded-Width Intervals

In this section, we consider a class of one-to-one transformations from the parameter space $\left(\frac{-1}{m-1}, 1\right)$ to $(0,\infty)$ and construct corresponding bounded-width confidence intervals for ρ following the fixed-accuracy approach explained earlier. This class of transformations is proposed through a chain of transformations.

Recall that we have $\frac{-1}{m-1} < \rho < 1$. So, we first transform ρ to $\rho' = \frac{\rho+(m-1)^{-1}}{1+(m-1)^{-1}}$ that belongs to $(0,1)$, and then transform ρ' to $(1-\rho')^{-\nu}$ with some $\nu > 0$, which finally yields the class of transformations $\{\widetilde{g}_\nu : \nu > 0\}$ as

$$\widetilde{g}_\nu(\rho) = (1-\rho')^{-\nu} - 1 = \left[\frac{m}{(m-1)(1-\rho)}\right]^\nu - 1 \in (0,\infty).$$

Our first objective is to construct a fixed-accuracy confidence interval for the transformed parameter

$$\rho_\nu := \widetilde{g}_\nu(\rho),\ \text{for } \nu > 0.$$

Recall that $\widehat{\rho}_n$ denotes the MLE of ρ defined in (2.3). By invariance property, the MLE of ρ_ν, denoted by $\widehat{\rho}_{\nu n}$, is $\widetilde{g}_\nu(\widehat{\rho}_n)$ which belongs to $(0,\infty)$ a.s. With some prescribed fixed-accuracy measure $\delta > 1$, we would like to construct a fixed-accuracy confidence interval

$$I_{\nu n} = \left[\delta^{-1}\widehat{\rho}_{\nu n},\, \delta\widehat{\rho}_{\nu n}\right] \text{ for } \rho_\nu,$$

such that the coverage probability $P_\theta(\rho_\nu \in I_{\nu n}) \geq 1-\alpha$ for any prefixed $\alpha \in (0,1)$ and $\nu > 0$.

Now, we define $\psi_\nu(x) = \ln \widetilde{g}_\nu(x)$ and apply the delta method from (2.4) as $n \to \infty$ to write

$$\sqrt{n}\left(\psi_\nu(\widehat{\rho}_n) - \psi_\nu(\rho)\right) \xrightarrow{\mathcal{L}} N\left(0, \xi^2(\nu,\rho)\right) \text{ with } \xi^2(\nu,\rho) = \frac{2\nu^2 m^{2\nu-1}}{m-1}\left[\frac{m-(m-1)(1-\rho)}{m^\nu-(m-1)^\nu(1-\rho)^\nu}\right]^2. \tag{4.4}$$

Using (4.4), the coverage probability can be approximated for large n as

$$P_\theta(\rho_\nu \in I_{\nu n}) = P_\theta\{-\ln\delta \leq \psi_\nu(\widehat{\rho}_n) - \psi_\nu(\rho) \leq \ln\delta\} \approx 2\Phi\left(\frac{\sqrt{n}\ln\delta}{\xi(\nu,\rho)}\right) - 1.$$

Then, by solving $2\Phi\left(\sqrt{n}\ln\delta/\xi(\nu,\rho)\right) - 1 \geq 1-\alpha$ for (large) n, we obtain $n \geq (\ln\delta)^{-2} z_{\alpha/2}^2 \xi^2(\nu,\rho)$. Therefore, the approximate optimal sample size required for constructing a $1-\alpha$ fixed-accuracy confidence interval for ρ_ν will be

$$n^*(\nu,\rho) = \left\lfloor \beta(\delta)\,\xi^2(\nu,\rho)\right\rfloor + 1, \text{ where } \beta(\delta) = \frac{z_{\alpha/2}^2}{(\ln\delta)^2}, \tag{4.5}$$

where $\lfloor u \rfloor$ stands for the largest integer smaller than $u(>0)$.

Note that the optimal sample size $n^*(\nu,\rho)$ involves the unknown parameter ρ, and hence, its magnitude remains unknown in practice. Similar to the approach of the fixed-width confidence interval estimation, we must gather observations at least in two-stages, where $n^*(\nu,\rho)$ would be estimated in the first stage by plugging in an estimator of ρ, say $\widehat{\rho}_k$, obtained from a pilot data of size k. If we do not have the required estimated sample size $n^*(\nu,\widehat{\rho}_k)$ in the first stage, we draw additional $n^*(\nu,\widehat{\rho}_k) - k$ observations in the second stage. Formally, the two-stage procedure is executed as follows:

Stage I: For some prefixed pilot sample size k, draw k i.i.d. observations $\boldsymbol{X}_1,\ldots,\boldsymbol{X}_k$ from $N_m(\mathbf{0},\,\sigma^2 R)$ and compute $\widehat{\rho}_k$ as in (2.3) based on the transformed observations $\boldsymbol{Y}_i = H\boldsymbol{X}_i$, $i = 1,\ldots,k$, where H is the Helmert matrix defined in (2.1). Estimate the optimal sample size $n^*(\nu,\rho)$ as

$$K^*(\nu,k) = \left\lfloor \beta(\delta)\,\xi^2(\nu,\widehat{\rho}_k)\right\rfloor + 1. \tag{4.6}$$

If $K^*(\nu,k) \leq k$, stop sampling and set the final sample size $N(\nu,k) = k$ and $\widehat{\rho}_{N(\nu,k)} = \widehat{\rho}_k$. Else, proceed to Stage II.

Stage II: Draw additional $(K^*(\nu,k) - k)$ i.i.d. observations $\boldsymbol{X}^*_{k+1},\ldots,\boldsymbol{X}^*_{K^*(\nu,k)}$ from $N_m\left(\mathbf{0},\,\sigma^2 R\right)$ independent of the initial k observations and stop sampling. Set $N(\nu,k) = K^*(\nu,k)$. Then transform these observations to $\mathbf{Y}_i^* = H\boldsymbol{X}_i^*$, for $i = k+1, \ldots, K^*(\nu, k)$, and compute the final point estimator of ρ as

$$\widehat{\rho}_{N(\nu,k)} = \frac{V_{1k} + V^*_{1(N(\nu,k)-k)} - \left(V_{2k} + V^*_{2(N(\nu,k)-k)}\right)/(m-1)}{V_{1k} + V^*_{1(N(\nu,k)-k)} + V_{2k} + V^*_{2(N(\nu,k)-k)}},$$

where

$$V^*_{1(N(\nu,k)-k)} = \sum_{i=k+1}^{K^*(\nu,k)} Y_{i1}^{*2} \text{ and } V^*_{2(N(\nu,k)-k)} = \sum_{i=k+1}^{K^*(\nu,k)} \sum_{j=2}^{m} Y_{ij}^{*2}.$$

The final sample size for the two-stage procedure is

$$N(\nu, k) = \max\{k, K^*(\nu, k)\}, \tag{4.7}$$

and the $1-\alpha$ fixed-accuracy confidence interval for ρ_ν is given as

$$I_{\nu N(\nu,k)} = \left[\delta^{-1}\widehat{\rho}_{\nu N(\nu,k)},\, \delta\widehat{\rho}_{\nu N(\nu,k)}\right] \text{ where } \widehat{\rho}_{\nu N(\nu,k)} = \widetilde{g}_\nu\left(\widehat{\rho}_{N(\nu,k)}\right).$$

Next, we construct the confidence interval for the parameter of interest ρ. Note that $\widetilde{g}_\nu$ is a strictly increasing function. Thus, the inverse function $\widetilde{g}_\nu^{-1} : (0, \infty) \to (-1/(m-1), 1)$ exists and is given by

$$\widetilde{g}_\nu^{-1}(x) = 1 - \left(\frac{m}{m-1}\right)(1+x)^{-\frac{1}{\nu}}.$$

Now, since $\widetilde{g}_\nu^{-1}(x)$ is also a strictly increasing function of x, the inverse image of the interval $I_{\nu N(\nu,k)}$, expressed as

$$J_{\nu N(\nu,k)} = \widetilde{g}_\nu^{-1}(I_{\nu N(\nu,k)}) = \left[\widetilde{g}_\nu^{-1}\left(\delta^{-1}\widehat{\rho}_{\nu N(\nu,k)}\right),\, \widetilde{g}_\nu^{-1}\left(\delta\widehat{\rho}_{\nu N(\nu,k)}\right)\right],$$

becomes an approximate $1-\alpha$ confidence interval for ρ. Needless to say that both the upper and the lower confidence limits of $J_{\nu N(\nu,k)}$ are now well inside the parameter space $\left(\frac{-1}{m-1}, 1\right)$ a.s. which is the primary motivation for developing this approach. However, note that the width of $J_{\nu N(\nu,k)}$, given by

$$\widetilde{W}_{\nu N(\nu,k)} = \widetilde{g}_\nu^{-1}\left(\delta\widehat{\rho}_{\nu N(\nu,k)}\right) - \widetilde{g}_\nu^{-1}\left(\delta^{-1}\widehat{\rho}_{\nu N(\nu,k)}\right),$$

is now a random variable. In the next lemma, we prove that the width $\widetilde{W}_{\nu N(\nu,k)}$ is bounded from above by some quantity that is smaller than the length of the parameter space $(-1/(m-1), 1)$ of ρ.

Lemma 4.1. *The width $\widetilde{W}_{\nu N(\nu,k)}$ of the interval $J_{\nu N(\nu,k)}$ is bounded from above by*

$$\left(\frac{m}{m-1}\right) u(\nu, \delta) \text{ a.s.}$$

where

$$u(\nu, \delta) = \left[\left(\delta^{\frac{2}{\nu+1}} - 1\right)^{1+\frac{1}{\nu}} / \left(\delta^2 - 1\right)^{\frac{1}{\nu}}\right] < 1 \ \ \forall\, \nu > 0,\ \delta > 1.$$

Proof. Noting that $\widehat{\rho}_{\nu N(\nu,k)} > 0$ a.s., consider the following function

$$\widetilde{g}_\nu^{-1}(\delta x) - \widetilde{g}_\nu^{-1}\left(\delta^{-1}x\right) = \left(\frac{m}{m-1}\right) f(x), \text{ where } f(x) = \left(1+\delta^{-1}x\right)^{-\frac{1}{\nu}} - (1+\delta x)^{-\frac{1}{\nu}},$$

with $x > 0$. We need to maximize $f(x)$ for $x > 0$ to obtain an upper bound for $\widetilde{W}_{\nu n}$. The inequality $f'(x) > 0$ yields

$$\delta(1+\delta x)^{-1-\frac{1}{\nu}} - \frac{1}{\delta}\left(1+\delta^{-1}x\right)^{-1-\frac{1}{\nu}} > 0, \text{ that is, } x < x_0 = \frac{\delta^{\frac{2\nu}{\nu+1}} - 1}{\delta - \delta^{\frac{\nu-1}{\nu+1}}},$$

using the fact that $\delta > \delta^{(\nu-1)/(\nu+1)}$.

So, $f'(x) > 0$ when $x < x_0$, $f'(x) < 0$ when $x > x_0$, and x_0 is a critical point that could be a local maximum. We conclude that x_0 is actually a global maximum by observing

$$\lim_{x\to 0^+} f(x) = \lim_{x\to\infty} f(x) = 0 \text{ and } f(x_0) = \left(\delta^{\frac{2}{\nu+1}} - 1\right)^{1+\frac{1}{\nu}} / \left(\delta^2 - 1\right)^{\frac{1}{\nu}} > 0.$$

Now, it remains to show that $u(\nu,\delta) < 1$. Consider the function $g(y) = (y-1)^{\nu+1} - y^{\nu+1} + 1$ where $y = \delta^{2/(\nu+1)} > 1$. Note that $g'(y) < 0$ for all $y > 1$ which implies $\sup_{y>1} g(y) = \lim_{y\to 1^+} g(y) = 0$. Therefore, $g(y) < 0$, which implies $\left(\delta^{2/(\nu+1)} - 1\right)^{\nu+1} < \delta^2 - 1$. This completes the proof. □

Lemma 4.1 ensures that the class of intervals $\{J_{\nu N(\nu,k)} : \nu > 0\}$ provide $(1-\alpha)$ *bounded-width confidence intervals* for ρ with non-trivial upper bounds. Since widths of these intervals are only guaranteed to be bounded, we should evaluate the performance of these intervals in terms of the estimated expected length $E_\theta\left[\widetilde{W}_{\nu N(\nu,k)}\right]$ and coverage probabilities through extensive sets of simulation studies.

However, it is not feasible to evaluate the performance of $J_{\nu N(\nu,k)}$ for all values of $\nu > 0$. Moreover, we will be interested in exploring the desirable asymptotic efficiency and consistency properties of the proposed two-stage procedure as $\delta \downarrow 1$. Investigation of these properties will be unnecessarily difficult unless we deal with the cases separately for specific values of ν. Before deciding about the possible values of ν to be considered in this article, let us observe the following interesting result.

Lemma 4.2. *The optimal sample size $n^*(\nu,\rho)$ is an increasing function of ν for fixed ρ and $n^*(\nu,\rho)$ is free of ρ if and only if $\nu = 1$.*

Proof. From the expression of $\xi^2(\nu,\rho)$ in (4.4), let us consider the part involving ν and take logarithm to define the function

$$h(\nu) = \ln\left[\frac{\nu^2 m^{2\nu-1}}{(m^\nu - (m-1)^\nu(1-\rho)^\nu)^2}\right].$$

It suffices to show that $h'(\nu) > 0$ for all $\nu > 0$. To this end, we first obtain

$$h'(\nu) = \frac{2}{\nu}\left[1 - \frac{\ln\{m/(m-1)(1-\rho)\}^\nu}{\{m/(m-1)(1-\rho)\}^\nu - 1}\right].$$

Note that the inequality $\ln x \leq x - 1$ holds for all $x > 0$ where the equality holds if and only if when $x = 1$. Now, $\rho > -1/(m-1)$ implies that $m/(m-1)(1-\rho) > 1$. Thus, $h'(\nu) > 0$ for all $\nu > 0$ is established by applying the previous inequality with $x = \{m/(m-1)(1-\rho)\}^{\nu}$ which is strictly larger than 1 for all $\nu > 0$. The fact that $n^*(\nu, \rho)$ is free of ρ if and only if $\nu = 1$ is evident from the expressions in (4.4) and (4.5). □

Lemma 4.2 suggests that the optimal sample size will be smaller if ν is chosen closer to 0, but it will be higher if ν is chosen bigger than 1. The case $\nu = 1$ is very interesting as the optimal sample size in this case will be completely known, and hence, we do not need a two-stage procedure for constructing a $1 - \alpha$ confidence interval of ρ.

However, for values of ν other than 1, we do need a two-stage procedure. Again, it is clearly not possible to explore all values of $\nu \neq 1$ and pick out the best one. In this paper, we make a modest attempt to find a good interval estimator of ρ by considering $\nu = 1, \frac{1}{2}$, and 2 which lead to handful of simple enough one-to-one transformations. Moreover, we will see that, in general, the investigation of the theoretical properties of the estimation procedure corresponding to $\nu = 1, \frac{1}{2}$, and 2 will be substantially interesting. Also, then, it may not be too difficult to pursue asymptotic second-order efficiency property in the sense of Ghosh and Mukhopadhyay (1981).

For convenience, we adopt the following notations for the three transformations corresponding to $\nu = 1$, $\frac{1}{2}$, and 2 as follows:

$$g_1(\rho) = \widetilde{g}_1(\rho),\ \ g_2(\rho) = \widetilde{g}_{\frac{1}{2}}(\rho), \text{ and } g_3(\rho) = \widetilde{g}_2(\rho). \tag{4.8}$$

We will investigate the theoretical properties of the corresponding bounded-width confidence interval methodologies separately in the next sections.

4.2. Fixed-Sample-Size Methods with g_1 Transformation

In this section, we consider the transformation $g_1\colon (\frac{-1}{m-1}, 1) \to (0, \infty)$ that leads to a fixed sample size method for constructing an *exact* $1 - \alpha$ confidence interval based on a *known optimal sample size*. Let us denote the transformed parameter and its MLE based on n observations as

$$\rho_1 = g_1(\rho) = \frac{1 + (m-1)\rho}{(m-1)(1-\rho)} \ \text{ and } \ \widehat{\rho}_{1n} = g_1(\widehat{\rho}_n).$$

With a fixed-accuracy measure $\delta > 1$, let

$$I_{1n} = \left[\delta^{-1}\widehat{\rho}_{1n},\, \delta\widehat{\rho}_{1n}\right]$$

be a $1 - \alpha$ fixed-accuracy confidence interval for ρ_1. Simple algebraic manipulations yield $\widehat{\rho}_{1n} = V_{1n}/V_{2n}$ and a "pivotal" quantity for ρ is given by

$$\frac{\widehat{\rho}_{1n}}{\rho_1} = \frac{V_{1n}/(1 + (m-1)\rho)}{V_{2n}/(m-1)(1-\rho)} = \frac{\chi^2_n/n}{\chi^2_{n(m-1)}/n(m-1)} \sim F_{n,\,n(m-1)}, \tag{4.9}$$

since V_{1n}, V_{2n} are independent.

Recall from Theorem 3.1 that $G_n(.)$ denotes the CDF of a $F_{n(m-1),\,n}$ random variable. The coverage probability for ρ_1 is then given by

$$P_\theta(\rho_1 \in I_{1n}) = P_\theta\left(\delta^{-1} < \frac{\rho_1}{\widehat{\rho}_{1n}} < \delta\right) = G_n(\delta) - G_n(\delta^{-1}).$$

The minimal sample size required for $1-\alpha$ fixed-accuracy confidence interval for ρ_1 can be obtained by solving

$$G_n(\delta) - G_n(\delta^{-1}) = 1 - \alpha \text{ for } n.$$

Thus, the optimal fixed-sample-size is given as

$$n_{1e}^* \equiv n_{1e}^*(\delta) = \lfloor n_1^0(\delta) \rfloor + 1, \text{ where } n_1^0(\delta) \text{ is a solution of } G_n(\delta) - G_n(\delta^{-1}) = 1 - \alpha. \tag{4.10}$$

Since the analytic form of the solution of (4.10) is not available in a close form, the solution $n_1^0(\delta)$ would be obtained by an appropriate numerical technique. For numerical study, we used Broyden's quasi-Newton method (see Broyden, 1965) for solving nonlinear equations available in statistical software "R".

As an alternative approach, we can find an approximate expression for the optimal sample size using the asymptotic distribution of $\widehat{\rho}_n$ as described in Section 4.1. That is, an approximate but analytic expression for the optimal sample size is obtained from (4.5) as

$$n_1^* = n^*(1, \rho) = \left\lfloor \beta(\delta)\left(\frac{2m}{m-1}\right) \right\rfloor + 1.$$

Extensive numerical study shows that the exact, but numerically computed, optimal sample size n_{1e}^* and the approximately optimal sample size n_1^* are almost the same for all $m > 1$ and $\delta > 1$. In all the cases we studied, they only differed by at most 1.

Remark 4.1. Very similar observations were reported in a seemingly totally unrelated paper of Mukhopadhyay and Zhuang (2016, Section 3) which dealt with the construction of a fixed-accuracy confidence interval problem in the context of a useful functional within the framework described as Fisher's "Nile" example. In that problem also, the pivot had a F-distribution leading to an optimal sample size that became completely free from the unknown parameter whereas numerically computed, optimal sample size and the approximately optimal sample sizes were all nearly same!

Summing up, we have two fixed-sample-size procedures that would produce $1-\alpha$ fixed-accuracy confidence intervals $I_{1n_{1e}^*} = \left[\delta^{-1}\widehat{\rho}_{1n_{1e}^*},\, \delta\widehat{\rho}_{1n_{1e}^*}\right]$ and $I_{1n_1^*} = \left[\delta^{-1}\widehat{\rho}_{1n_1^*},\, \delta\widehat{\rho}_{1n_1^*}\right]$ for ρ_1 based on *known* exact optimal sample size n_{1e}^* and approximate optimal sample size n_1^*. Finally, $1-\alpha$ bounded-width confidence intervals of the original parameter ρ are obtained from the inverse images of $I_{1n_{1e}^*}$ and $I_{1n_1^*}$ as follows:

$$J_{1n_{1e}^*} = \left[g_1^{-1}\left(\delta^{-1}\widehat{\rho}_{1n_{1e}^*}\right),\, g_1^{-1}\left(\delta\widehat{\rho}_{1n_{1e}^*}\right)\right], \text{ and } J_{1n_1^*} = \left[g_1^{-1}\left(\delta^{-1}\widehat{\rho}_{1n_1^*}\right),\, g_1^{-1}\left(\delta\widehat{\rho}_{1n_1^*}\right)\right],$$

where $g_1^{-1}(x) = (x(m-1)-1)/(m-1)(1+x)$ is the inverse function of g_1.

Lemma 4.1 ensures that the widths of these intervals $J_{1n_{1e}^*}$ and $J_{1n_1^*}$, denoted as $W_{1n_{1e}^*}$ and $W_{1n_1^*}$

respectively, are bounded from above by

$$W_{1n^*_{1e}} \leq \left(\frac{m}{m-1}\right)\left(\frac{\delta-1}{\delta+1}\right) \text{ and } W_{1n^*_1} \leq \left(\frac{m}{m-1}\right)\left(\frac{\delta-1}{\delta+1}\right).$$

Though g_1 transformation leads to a pivotal quantity for ρ which makes it possible to have a fixed-sample-size procedure, there is no theoretical guarantee up-front that this transformation will produce the best interval in terms of having the minimum average width among all such intervals. The performances of these bounded-width confidence intervals will be evaluated in Section 4.5 in terms of the attained coverage probability and the average width via extensive sets of simulation studies.

4.3. Two-Stage Procedures with g_2 and g_3 Transformations

In this section, we consider transformations g_2 and g_3 given in (4.8) and construct corresponding $1-\alpha$ confidence intervals for ρ via fixed-accuracy approach. Additionally, we will establish some asymptotic optimality properties of these interval estimators.

Let us define the transformed parameters ρ_2 and ρ_3 that belong to $(0,\infty)$ as

$$\rho_2 = g_2(\rho) = \left(\frac{m}{(m-1)(1-\rho)}\right)^{1/2} - 1 \text{ and } \rho_3 = g_3(\rho) = \left(\frac{m}{(m-1)(1-\rho)}\right)^{2} - 1,$$

and their maximum likelihood estimators based on n observations as $\widehat{\rho}_{2n} = g_2(\widehat{\rho}_n)$ and $\widehat{\rho}_{3n} = g_3(\widehat{\rho}_n)$. For any prefixed fixed-accuracy measure $\delta > 1$, let

$$I_{2n} = \left[\delta^{-1}\widehat{\rho}_{2n},\, \delta\widehat{\rho}_{2n}\right] \text{ and } I_{3n} = \left[\delta^{-1}\widehat{\rho}_{3n},\, \delta\widehat{\rho}_{3n}\right]$$

be $1-\alpha$ fixed-accuracy confidence intervals for transformed parameters ρ_2 and ρ_3 respectively. Following the arguments in Section 4.1, the optimal sample size required for constructing I_{2n} and I_{3n}, obtained from (4.5), are respectively given by

$$n^*_2 \equiv n^*_2(\delta) = n^*(1/2,\rho) = \lfloor \beta(\delta)\,\xi^2_2(\rho)\rfloor + 1 \text{ and } n^*_3 \equiv n^*_3(\delta) = n^*(2,\rho) = \lfloor \beta(\delta)\,\xi^2_3(\rho)\rfloor + 1, \quad (4.11)$$

where $\xi^2_2(\rho) = \xi^2(1/2,\rho)$ and $\xi^2_3(\rho) = \xi^2(2,\rho)$, obtained from (4.4) as

$$\xi^2_2(\rho) = \frac{1}{2(m-1)}\left[\sqrt{m} + \sqrt{(m-1)(1-\rho)}\right]^2 \text{ and } \xi^2_3(\rho) = \frac{8m^3}{m-1}\left[m+(m-1)(1-\rho)\right]^{-2}.$$

Unlike the optimal sample sizes n^*_{1e} and n^*_1 corresponding to g_1 transformation, now the expressions of the optimal sample sizes n^*_2 and n^*_3 involve unknown parameter ρ, and thus, they are unknown in practice. So, there cannot be any fixed sample size procedure for constructing $1-\alpha$ fixed-accuracy confidence intervals for ρ_2 and ρ_3.

The two-stage procedure described in Section 4.1 should be applied in these two cases. Before describing the two-stage procedures, let us obtain some reasonable choices of pilot sample sizes for both cases.

Note that $\xi^2_2(\rho)$ is a strictly decreasing function of ρ which implies that $\inf_\rho \xi^2_2(\rho) = \frac{1}{2}m/(m-1)$.

Therefore, we have $n_2^* \geq k_2$ where

$$k_2 \equiv k_2(\delta) = \left\lfloor \beta(\delta)\, \frac{m}{2(m-1)} \right\rfloor + 1, \tag{4.12}$$

which is a non-trivial lower bound for the optimal sample size n_2^*. Similarly, $\xi_3^2(\rho)$ is a strictly increasing function of ρ which yields $\inf_\rho \xi_3^2(\rho) = 2m/(m-1)$. Thus, a non-trivial lower bound for the optimal sample size n_3^* becomes

$$k_3 \equiv k_3(\delta) = \left\lfloor \beta(\delta)\, \frac{2m}{m-1} \right\rfloor + 1. \tag{4.13}$$

Since such lower bounds k_2 and k_3 are free of any unknown parameters, they become natural choices of pilot sample sizes for the two-stage procedures for the estimation of ρ_2 and ρ_3 respectively. Next, we describe the two-stage fixed-accuracy interval estimation procedures for ρ_2 and ρ_3 separately using k_2 and k_3 as respective pilot sample sizes.

Two-Stage Procedure for ρ_2

Let us first consider the two-stage procedure for $1-\alpha$ fixed-accuracy interval estimation of ρ_2 with k_2 as the pilot sample size. In the first stage, the estimator of optimal sample size n_2^* is obtained from (4.6) as

$$K_2^* = K^*(1/2, k_2) = \left\lfloor \beta(\delta)\, \xi_2^2(\widehat{\rho}_{k_2}) \right\rfloor + 1. \tag{4.14}$$

Since $\xi_2^2(\widehat{\rho}_{k_2}) \geq m/2(m-1)$ a.s., we have $K_2^* \geq k_2$ w.p.1. Thus, the final sample size or the stopping variable of the two-stage procedure, obtained from (4.7), will look like

$$N_2 = N(1/2, k_2) = \max\{k_2, K^*(1/2, k_2)\} = \max\{k_2, K_2^*\} = K_2^*, \tag{4.15}$$

with probability 1. This implies that sampling will never stop in the first stage if k_2 is used as the pilot sample size.

The derived $1-\alpha$ fixed-accuracy confidence interval for ρ_2 based on N_2 observations is given by

$$I_{2N_2} = \left[\delta^{-1}\widehat{\rho}_{2N_2},\, \delta\widehat{\rho}_{2N_2}\right],$$

where $\widehat{\rho}_{2N_2} = g_2\left(\widehat{\rho}_{N_2}\right)$. Finally, the corresponding $1-\alpha$ bounded-width confidence interval of ρ is going to be

$$J_{2N_2} = \left[g_2^{-1}\left(\delta^{-1}\widehat{\rho}_{2N_2}\right),\, g_2^{-1}\left(\delta\widehat{\rho}_{2N_2}\right)\right] \text{ where } g_2^{-1}(x) = 1 - \frac{m}{m-1}(1+x)^{-2}.$$

Lemma 4.1 guarantees that the width of J_{2N_2}, denoted as W_{2N_2}, is bounded from above, that is, we claim

$$W_{2N_2} \leq \left(\frac{m}{m-1}\right) \frac{\left(\delta^{4/3}-1\right)^3}{\left(\delta^2-1\right)^2}$$

with probability 1.

Two-Stage Procedure for ρ_3

Now, we consider the two-stage procedure for fixed-accuracy interval estimation of ρ_3 where k_3 is the pilot sample size. After collecting data in the first stage, the estimator of n_3^*, calculated from (4.6), becomes

$$K_3^* = K^*(2, k_3) = \left\lfloor \beta(\delta)\, \xi_3^2(\widehat{\rho}_{k_3}) \right\rfloor + 1.$$

As in the previous case, $K_3^* \geq k_3$ w.p.1 since $\xi_3^2(\widehat{\rho}_{k_3}) \geq 2m/(m-1)$ w.p.1. Therefore, the final sample size, obtained from (4.7), reduces to

$$N_3 = N(2, k_3) = \max\{k_3, K^*(2, k_3)\} = \max\{k_3, K_3^*\} = K_3^* \text{ a.s.} \tag{4.16}$$

Then, $1-\alpha$ fixed-accuracy confidence interval for ρ_3 based on N_3 observations is

$$I_{3N_3} = \left[\delta^{-1}\widehat{\rho}_{3N_3},\, \delta\widehat{\rho}_{3N_3}\right]$$

where $\widehat{\rho}_{3N_3} = g_3\left(\widehat{\rho}_{N_3}\right)$, and the corresponding $1-\alpha$ bounded-width confidence interval of ρ is

$$J_{3N_3} = \left[g_3^{-1}\left(\delta^{-1}\widehat{\rho}_{3N_3}\right),\, g_3^{-1}\left(\delta\widehat{\rho}_{3N_3}\right)\right] \text{ where } g_3^{-1}(x) = 1 - \frac{m}{m-1}(1+x)^{-1/2}.$$

We can conclude from Lemma 4.1 that the width of J_{3N_3}, denoted as W_{3N_3} is bounded from above, that is, we claim

$$W_{3N_3} \leq \left(\frac{m}{m-1}\right)\left[(\delta^{2/3}-1)^3/(\delta^2-1)\right]^{1/2} \text{ w.p.1.}$$

For these two-stage procedures, we observe that $N_2 = K_2^*$ and $N_3 = K_3^*$ w.p.1, which implies that sampling will continue beyond the first stage if k_2 and k_3, given in (4.12) and (4.13), are used as pilot sample sizes. Other forms of pilot sample size may also be considered if a given practical scenario so demands. We will evaluate the performances of these procedures through extensive sets of simulation studies in Section 4.5.

At this point, however, it is worth mentioning that the stopping variable N_2 with pilot sample size k_2 and N_3 with pilot sample size k_3 enjoy a very attractive asymptotic optimality property, namely the *second-order asymptotic efficiency* property. This is a stronger theoretical property than the ordinary first-order asymptotic efficiency. Next, we establish some desirable asymptotic properties along with this second-order property of the proposed two-stage procedures obtained under the transformations g_2 and g_3.

4.3.1. Asymptotic Properties of N_2 and N_3

As the unknown optimal sample sizes n_2^* and n_3^* are estimated by N_2 and N_3 respectively using two-stage procedures, it is natural to ask whether such adaptations are good enough, at least asymptotically. In the literature on multi-stage sampling, this is typically verified by proving asymptotic first-order or second-order efficiency properties of the stopping variable.

These efficiency properties make sure that the expected sample sizes $E_\theta[N_i] \equiv E_\theta[N_i(\delta)]$ and the optimal sample sizes $n_i^* \equiv n_i^*(\delta)$, for $i = 2, 3$, are asymptotically close enough in the sense of either the ratio converging to 1 or the difference being bounded when the fixed-accuracy measure $\delta \downarrow 1$. Another crucial property for a confidence interval method is that the coverage probability achieved by the interval is close enough to the target level of confidence, at least asymptotically. Formally, this is known as asymptotic consistency property, which we will investigate for the proposed confidence intervals $J_{2N_2(\delta)}$ and $J_{3N_3(\delta)}$. Such theoretical results are highly desirable for a well-designed multi-stage sampling strategy.

Before we formally proceed to the asymptotic analysis, we note that the optimal sample sizes $n_2^*(\delta)$ and $n_3^*(\delta)$ grow to ∞ as $\delta \downarrow 1$. Therefore, to develop the asymptotic theory for $N_2(\delta)$ and $N_3(\delta)$, the corresponding pilot sample sizes, as functions of δ, must also tend to ∞. In the present scenario, the proposed pilot sample sizes $k_2(\delta)$ and $k_3(\delta)$ both tend to ∞ as $\delta \downarrow 1$. Thus, we proceed with the asymptotic analysis by first proving a lemma which will be used to establish second-order asymptotic efficiency property of the stopping variables.

Lemma 4.3. *For the maximum likelihood estimator $\widehat{\rho}_n$ in* (2.3), *there exists constants L_2, U_2, L_3 and U_3 such that*

$$\begin{aligned} (i) \quad & \frac{L_2}{n} + \mathcal{O}(n^{-2}) \le E_\theta\left[\xi_2^2(\widehat{\rho}_n) - \xi_2^2(\rho)\right] \le \frac{U_2}{n} + \mathcal{O}(n^{-2}); \\ (ii) \quad & \frac{L_3}{n} + \mathcal{O}(n^{-2}) \le E_\theta\left[\xi_3^2(\widehat{\rho}_n) - \xi_3^2(\rho)\right] \le \frac{U_3}{n} + \mathcal{O}(n^{-2}). \end{aligned}$$

Proof. (i) Suppose $F_{n,\,n(m-1)}$ denotes a F random variable with the numerator and denominator degrees of freedom n and $n(m-1)$ respectively. From (4.9) and the definition of $\widehat{\rho}_n$ in (2.3), we have

$$F_{n,\,n(m-1)} = \left[\frac{(1-\rho)(m-1)}{1+(m-1)\rho}\right]\frac{V_{1n}}{V_{2n}} \text{ and } \widehat{\rho}_n = \frac{(1+(m-1)\rho)F_{n,\,n(m-1)} - (1-\rho)}{(1+(m-1)\rho)F_{n,\,n(m-1)} + (m-1)(1-\rho)}.$$

Let us define a function, for $x \ge 0$

$$q_2(x) = \xi_2^2\left(\frac{ax-b}{ax+c}\right), \text{ where } a = 1+(m-1)\rho,\ b = 1-\rho, \text{ and } c = (m-1)(1-\rho).$$

Then, $q_2\left(F_{n,\,n(m-1)}\right) = \xi_2^2(\widehat{\rho}_n)$ and $q_2(1) = \xi_2^2\left(\frac{a-b}{a+c}\right) = \xi_2^2(\rho)$. The first two derivatives of $q_2(x)$ are given by

$$q_2'(x) = -\frac{a\sqrt{b+c}}{2\sqrt{m-1}}\left[\frac{\sqrt{m}}{(ax+c)^{3/2}} + \frac{\sqrt{(m-1)(b+c)}}{(ax+c)^2}\right] < 0 \text{ and}$$

$$q_2''(x) = \frac{a\sqrt{b+c}}{2\sqrt{m-1}}\left[\frac{3\sqrt{m}}{2(ax+c)^{5/2}} + \frac{2\sqrt{(m-1)(b+c)}}{(ax+c)^3}\right] > 0.$$

By Taylor series expansion of $q_2\left(F_{n,\,n(m-1)}\right)$ around 1, we express

$$q_2\left(F_{n,\,n(m-1)}\right) - q_2(1) = \left[F_{n,\,n(m-1)} - 1\right]q_2'(1) + \frac{1}{2}\left[F_{n,\,n(m-1)} - 1\right]^2 q_2''(\zeta_n), \tag{4.17}$$

where ζ_n is a random variable between 1 and $F_{n,\,n(m-1)}$. Asymptotic orders of the following expectations can be observed from Lemma 5.1 of De and Mukhopadhyay (2015) as

$$E\left[F_{n,\,n(m-1)}-1\right]=\frac{2}{n(m-1)}+\mathcal{O}\left(n^{-2}\right)\ \text{ and }\ E\left[F_{n,\,n(m-1)}-1\right]^2=\frac{2m}{n(m-1)}+\mathcal{O}\left(n^{-2}\right). \quad (4.18)$$

Now, note that $q_2'(1)$ does not depend on n and $q_2''(x)>0$ for all $x\geq 0$. Taking expectations on both sides of (4.17) and using the asymptotic orders from (4.18), we obtain the following lower bound

$$E_\theta\left[\xi_2^2(\widehat{\rho}_n)-\xi_2^2(\rho)\right]\geq\frac{L_2}{n}+\mathcal{O}\left(n^{-2}\right)\ \text{ where } L_2=\frac{2q_2'(1)}{m-1}<0. \quad (4.19)$$

Since $q_2''(x)$ is a strictly decreasing function of x, we claim that $q_2''(x)\leq q_2''(0)$ for all $x\in[0,\infty)$, where $q_2''(0)$ is a positive constant free of n. Thus, taking expectations on both sides of (4.17), using $q_2''(0)$ as the upper bound of $q_2''(\zeta_n)$ and the asymptotic orders as shown in (4.18), the following upper bound is derived

$$\begin{aligned}E_\theta\left[\xi_2^2(\widehat{\rho}_n)-\xi_2^2(\rho)\right]&\leq E\left[F_{n,\,n(m-1)}-1\right]q_2'(1)+\frac{1}{2}E\left[F_{n,\,n(m-1)}-1\right]^2 q_2''(0)\\&=\frac{U_2}{n}+\mathcal{O}\left(n^{-2}\right)\ \text{ where } U_2=\frac{2\,q_2'(1)+m\,q_2''(0)}{m-1}. \quad (4.20)\end{aligned}$$

Finally, combining (4.19) and (4.20), we complete the proof of part (i).

We follow a similar line of arguments to prove part (ii). Let us define $q_3(x)=\xi_3^2\left(\frac{ax-b}{ax+c}\right)$ and observe that $q_3\left(F_{n,\,n(m-1)}\right)=\xi_3^2(\widehat{\rho}_n)$ and $q_3(1)=\xi_3^2(\rho)$. The first two derivatives of $q_3(x)$ are

$$q_3'(x)=\frac{16m^3a(b+c)}{(ax+c)^2}\left[m+\frac{(m-1)(b+c)}{ax+c}\right]^{-3}\ \text{ and } q_3''(x)=-48m^3a(b+c)\frac{l'(x)}{(l(x))^4},$$

where

$$l(x)=m(ax+c)^{2/3}+\frac{(m-1)(b+c)}{(ax+c)^{1/3}}\ \text{and } l'(x)=\frac{a}{3}(ax+c)^{-1/3}\left[2m-\frac{(m-1)(b+c)}{ax+c}\right].$$

Next, Taylor expansion of $q_3\left(F_{n,\,n(m-1)}\right)$ around 1 yields

$$q_3\left(F_{n,\,n(m-1)}\right)-q_3(1)=\left[F_{n,\,n(m-1)}-1\right]q_3'(1)+\frac{1}{2}\left[F_{n,\,n(m-1)}-1\right]^2 q_3''(\zeta_{1n}), \quad (4.21)$$

where ζ_{1n} is a random variable between 1 and $F_{n,\,n(m-1)}$. Note that, for all $x\geq 0$, $q_3'(x)$ take positive values that do not involve n and

$$l'(x)\geq\frac{a}{3}(ax+c)^{-1/3}\left\{2m-(m-1)(b+c)/c\right\}=\frac{am}{3}(ax+c)^{-1/3}>0.$$

This implies $q_3''(x)<0$ for all $x\geq 0$. Thus, using $q_3''(\zeta_{1n})\leq 0$ a.s., taking expectations on both sides of (4.21), and applying the asymptotic orders shown in (4.18), the following upper bound is established

$$E_\theta\left[\xi_3^2(\widehat{\rho}_n)-\xi_3^2(\rho)\right]\leq\frac{U_3}{n}+\mathcal{O}\left(n^{-2}\right)\ \text{ where } U_3=\frac{2q_3'(1)}{m-1}>0. \quad (4.22)$$

To obtain a corresponding lower bound, we need to find $\sup_{x\geq 0} l'(x)(l(x))^{-4}$. To this end, let us first find the global maximum of $l'(x)$. A first derivative test yields that $l''(x) > 0$ for all $x < c/a$ and $l''(x) < 0$ for all $x > c/a$. So, $l'(x)$ has a local maximum at $x = c/a$. Note that $l'(x) \to 0$ as $x \to \infty$, $l'(x) \to \frac{1}{3}am(c)^{-1/3}$ as $x \to 0$ and $l'(c/a) = \frac{1}{2}am(2c)^{-1/3} > \frac{1}{3}am(c)^{-1/3}$. So, the global maximum of $l'(x)$ is at $x = c/a$. Since $l'(x) > 0$ for all $x \geq 0$, $l(x)$ attains global minimum at $x = 0$. Therefore, we obtain

$$\sup_{x\geq 0} \frac{l'(x)}{(l(x))^4} = \frac{l'(c/a)}{(l(0))^4} \text{ and } q_3''(x) \geq L_4 \text{ for all } x \geq 0,$$

where

$$L_4 = -48m^3 a(b+c)\frac{l'(c/a)}{(l(0))^4}$$

is a non-random constant.

Then, taking expectations on both sides of (4.21), using $q_3''(\zeta_{1n}) \geq L_4$ a.s. and applying the asymptotic orders in (4.18), we have the following lower bound

$$\begin{aligned} E_\theta\left[\xi_3^2(\widehat{\rho}_n) - \xi_3^2(\rho)\right] &\geq E\left[F_{n,\,n(m-1)} - 1\right] q_3'(1) + \frac{L_4}{2} E\left[\left(F_{n,\,n(m-1)} - 1\right)^2\right] \\ &= \frac{L_3}{n} + \mathcal{O}\left(n^{-2}\right) \text{ where } L_3 = \frac{2\,q_3'(1) + mL_4}{m-1}. \end{aligned} \tag{4.23}$$

Hence, combining (4.22) and (4.23), we complete the proof of Lemma 4.3. □

Now, we are equipped with the tools to prove the asymptotic second-order efficiency property for the stopping variable N_2. Moreover, we prove that the corresponding bounded-width confidence interval $J_{2N_2(\delta)}$ contains the true parameter value of ρ with probability converging to the target level $1-\alpha$.

Theorem 4.1. *The stopping variable $N_2(\delta)$ and the bounded-width confidence interval $J_{2N_2(\delta)}$ obtained from the two-stage procedure with pilot sample size $k_2(\delta)$ as in* (4.12) *yield*

- (i) $\lim_{\delta\downarrow 1} N_2(\delta)/n_2^*(\delta) = 1$ *w.p.*1;
- (ii) $E_\theta\left[N_2(\delta) - n_2^*(\delta)\right] = \mathcal{O}(1)$ [*Asymptotic second-order efficiency*];
- (iii) $\lim_{\delta\downarrow 1} P_\theta\left(\rho \in J_{2N_2(\delta)}\right) = 1-\alpha$ [*Asymptotic consistency*].

Proof. (i) From the definitions in (4.11) and (4.15), we can write the following inequalities (w.p.1):

$$\frac{\xi_2^2\left(\widehat{\rho}_{k_2(\delta)}\right)}{\xi_2^2(\rho) + \beta(\delta)^{-1}} \leq \frac{N_2(\delta)}{n_2^*(\delta)} \leq \frac{\xi_2^2\left(\widehat{\rho}_{k_2(\delta)}\right)}{\xi_2^2(\rho)} + \frac{1}{n_2^*(\delta)}.$$

Since ξ_2^2 is a continuous function and $\widehat{\rho}_n$ is a strong consistent estimator of ρ, it is clear that $\xi_2^2\left(\widehat{\rho}_{k_2(\delta)}\right) \to \xi_2^2(\rho)$ as $\delta \downarrow 1$ w.p.1. The proof of part (i) follows by noting that $\beta(\delta) \to \infty$ and $n_2^*(\delta) \to \infty$ as $\delta \downarrow 1$.

(ii) From (4.14) and (4.15), we have $N_2(\delta) = K_2^*(\delta) \geq \beta(\delta)\xi_2^2\left(\widehat{\rho}_{k_2(\delta)}\right)$ w.p.1. Taking expectations and using the lower bound from the part (i) of Lemma 4.3, we have

$$E_\theta[N_2(\delta)] \geq \beta(\delta)\left[\xi_2^2(\rho) + \frac{L_2}{k_2(\delta)} + \mathcal{O}\left(k_2(\delta)^{-2}\right)\right]$$
$$\geq n_2^*(\delta) - 1 + \frac{2(m-1)L_2}{m} + o(1). \tag{4.24}$$

The last inequality in (4.24) is obtained by using the definition of $n_2^*(\delta)$ in (4.11), $L_2 < 0$, we claim the inequality

$$\frac{2(m-1)}{m} - o(1) \leq \frac{\beta(\delta)}{k_2(\delta)} \leq \frac{2(m-1)}{m}, \text{ and } \frac{\beta(\delta)}{k_2(\delta)^2} \to 0 \text{ as } \delta \downarrow 1. \tag{4.25}$$

From (4.14) and (4.15), we have $N_2(\delta) = K_2^*(\delta) \leq \beta(\delta)\xi_2^2\left(\widehat{\rho}_{k_2(\delta)}\right) + 1$. Thus, taking expectations and using the upper bound of from the part (i) of Lemma 4.3, we write

$$E_\theta[N_2(\delta)] \leq \beta(\delta)\left[\xi_2^2(\rho) + \frac{U_2}{k_2(\delta)} + \mathcal{O}\left(k_2(\delta)^{-2}\right)\right] + 1$$
$$\leq n_2^*(\delta) + 1 + \frac{2(m-1)U_2}{m} + o(1). \tag{4.26}$$

The last inequality in (4.26) is obtained by using the definition of $n_2^*(\delta)$ and (4.25). Combining (4.24) and (4.26), we complete the proof of asymptotic second-order efficiency of the stopping variable N_2.

(iii) Following the proof of Lemma 3.1 and Theorem 3.3, we can establish the asymptotic normality of $\widehat{\rho}_{N_2(\delta)}$ as follows:

$$\sqrt{N_2(\delta)}\left(\widehat{\rho}_{N_2(\delta)} - \rho\right) \xrightarrow{\mathcal{L}} N(0, \sigma_\rho^2) \text{ as } \delta \downarrow 1. \tag{4.27}$$

To prove (4.27), we can simply mimic the proof of (3.15) by replacing $k(d)$, $N(d)$ and $n^*(d)$ by $k_2(\delta)$, $N_2(\delta)$ and $n_2^*(\delta)$ respectively and considering limits as $\delta \downarrow 1$.

Note that Lemma 3.1 and Theorem 3.3 require the pilot sample size $k(d) = o(n^*(d))$, whereas the pilot sample size $k_2(\delta)$, in the present case, does not satisfy this condition. However, this is not a problem since the condition $k(d) = o(n^*(d))$ was only used to ensure that $N(d)/n^*(d) \to 1$ w.p.1. In the present case, we already proved $N_2(\delta)/n_2^*(\delta) \to 1$ a.s. in part (i). Hence, the asymptotic normality in (4.27) follows.

Now, let us define the function $\phi_2(x) = \ln g_2(x)$ and apply the delta method to write

$$\sqrt{N_2(\delta)}\left(\phi_2\left(\widehat{\rho}_{N_2(\delta)}\right) - \phi_2(\rho)\right) \xrightarrow{\mathcal{L}} N(0,\ \xi_2^2(\rho)) \text{ as } \delta \downarrow 1. \tag{4.28}$$

From part (i) of Theorem 4.1, we have

$$\lim_{\delta\downarrow 1} N_2(\delta)(\ln\delta)^2 = \left(\lim_{\delta\downarrow 1}\frac{N_2(\delta)}{n_2^*(\delta)}\right)\left(\lim_{\delta\downarrow 1} n_2^*(\delta)(\ln\delta)^2\right) = z_{\alpha/2}^2\,\xi_2^2(\rho).$$

Finally, using (4.28) and the previous limit, the asymptotic coverage probability turns out to be

$$\begin{aligned}\lim_{\delta\downarrow 1} P_\theta\left(\rho \in J_{2N_2(\delta)}\right) &= \lim_{\delta\downarrow 1} P_\theta\left(g_2(\rho) \in I_{2N_2(\delta)}\right)\\ &= \lim_{\delta\downarrow 1} P_\theta\left(-\ln\delta \le \phi_2\left(\widehat{\rho}_{N_2(\delta)}\right) - \phi_2(\rho) \le \ln\delta\right)\\ &= \lim_{\delta\downarrow 1} 2\Phi\left(\frac{\sqrt{N_2(\delta)}\ln\delta}{\xi_2(\rho)}\right) - 1 = 1-\alpha.\end{aligned}$$

This proves the asymptotic consistency property of the proposed confidence interval J_{2N_2}. □

It is interesting that the two-stage procedure corresponding to transformation g_3 also enjoys the above mentioned asymptotic optimality properties which makes it an attractive alternate procedure for bounded-width confidence interval for ρ. The following theorem establishes the second-order asymptotic efficiency property of the stopping variable N_3 and the asymptotic consistency property of the corresponding confidence interval J_{3N_3}.

Theorem 4.2. *For the stopping variable $N_3(\delta)$ and the bounded-width confidence interval $J_{3N_3(\delta)}$, obtained from the two-stage procedure with pilot sample size $k_3(\delta)$ as in* (4.13), *we have*:

(i) $\lim_{\delta\downarrow 1} N_3(\delta)/n_3^*(\delta) = 1$ *w.p.*1;

(ii) $E_\theta\left[N_3(\delta) - n_3^*(\delta)\right] = \mathcal{O}(1)$ [*Asymptotic second-order efficiency*];

(iii) $\lim_{\delta\downarrow 1} P_\theta\left(\rho \in J_{3N_3(\delta)}\right) = 1-\alpha$ [*Asymptotic consistency*].

Proof. The proof is very similar to that of the Theorem 4.1, and hence we omit it. □

The g_1 transformation led to a fixed-sample-size procedure with a completely known optimal sample size n_{1e}^* and an exact bounded-width confidence interval for ρ. This makes g_1 a very attractive transformation to use in practice. On the other hand, Theorems 4.1 and 4.2 provide very strong theoretical justifications for considering the bounded-width confidence intervals J_{2N_2} and J_{3N_3} obtained from using the transformations g_2 and g_3.

These transformations are special cases of the general class of transformations $\{\widetilde{g}_\nu : \nu > 0\}$. Even though all these transforms exhibit strong theoretical properties, there is no guarantee that a transformation that leads to better performance of the corresponding confidence interval cannot be obtained from outside the class $\{\widetilde{g}_\nu : \nu > 0\}$. However, at this moment, we cannot afford to consider all one-to-one transformations from outside this class. Therefore, we make a modest attempt in this article by considering one simple enough and natural transformation, namely g_4, which is not a member of the class $\{\widetilde{g}_\nu : \nu > 0\}$. In Section 4.4, we provide: (i) some crucial details of the ensuing transformation, (ii) the two-stage procedure emerging from it, and then (iii) the desirable theoretical properties of the methodology.

4.4. Two-Stage Procedure with g_4 Transformation

In this section, we develop the theory and methodology for constructing a bounded-width confidence interval of ρ obtained by using a one-to-one transformation from $\left(\frac{-1}{m-1}, 1\right)$ to $(0, \infty)$, but it is not a member of the previously highlighted class of transformations $\{\widetilde{g}_\nu : \nu > 0\}$. To this end, we first transform ρ to $\rho' = \frac{\rho+(m-1)^{-1}}{1+(m-1)^{-1}}$ that belongs to $(0,1)$, and then transform ρ' to $\ln(1-\rho')^{-1}$, which finally yields the following transformation and the transformed parameter

$$\rho_4 = g_4(\rho) = \ln\left(\frac{m}{(m-1)(1-\rho)}\right) \in (0, \infty).$$

In order to construct a confidence interval for ρ, let us first construct a fixed-accuracy confidence interval for the transformed parameter ρ_4 in the spirit of Section 4.1.

Note that the MLE of ρ_4, denoted by $\widehat{\rho}_{4n}$, is $g_4(\widehat{\rho}_n)$ that belongs to $(0,\infty)$ w.p.1. With any prefixed fixed-accuracy measure $\delta > 1$, we let

$$I_{4n} = \left[\delta^{-1}\widehat{\rho}_{4n},\, \delta\widehat{\rho}_{4n}\right]$$

be the fixed-accuracy confidence interval for ρ_4 such that the coverage probability $P_\theta(\rho_4 \in I_{4n}) \geq 1-\alpha$ for any given $\alpha \in (0,1)$.

Now, define $\psi_4(x) = \ln g_4(x)$ and apply the delta method from (2.4) to obtain

$$\sqrt{n}\left(\psi_4(\widehat{\rho}_n) - \psi_4(\rho)\right) \xrightarrow{\mathcal{L}} N\left(0, \xi_4^2(\rho)\right), \text{where } \xi_4^2(\rho) = \frac{2}{m(m-1)}\left[\frac{m-(m-1)(1-\rho)}{\ln\left(\frac{m}{(m-1)(1-\rho)}\right)}\right]^2, \quad (4.29)$$

which is a function of ρ. From (4.29), the coverage probability can be approximated for large n as

$$P_\theta(\rho_4 \in I_{4n}) = P_\theta(-\ln\delta \leq \psi_4(\widehat{\rho}_n) - \psi_4(\rho) \leq \ln\delta) \approx 2\Phi\left(\frac{\sqrt{n}\ln\delta}{\xi_4(\rho)}\right) - 1.$$

Solving $2\Phi\left(\sqrt{n}\ln\delta/\xi_4(\rho)\right) - 1 \geq 1-\alpha$ for (large) n, we obtain $n \geq (\ln\delta)^{-2} z_{\alpha/2}^2 \xi_4^2(\rho)$. The approximate optimal fixed sample size for constructing a $1-\alpha$ fixed-accuracy confidence interval for ρ_4 will be

$$n_4^*(\delta) \equiv n_4^*(\rho,\delta) = \left\lfloor \beta(\delta)\, \xi_4^2(\rho) \right\rfloor + 1, \text{ where } \beta(\delta) = \frac{z_{\alpha/2}^2}{(\ln\delta)^2}. \quad (4.30)$$

The oracle optimal sample size n_4^* is not known in practice as it involves the unknown parameter ρ. Therefore, we propose a two-stage sampling procedure to construct a confidence interval of ρ in the spirit of the two-stage methodology described in Section 4.1.

In the first stage of sampling, we draw k pilot observations $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_k$ that are i.i.d. from $N_m(\mathbf{0}, \sigma^2 R)$. Based on these pilot samples, we estimate ρ by $\widehat{\rho}_k$ using (2.3) and then estimate the optimal sample size n_4^* by

$$K_4^*(\delta) \equiv K_4^*(k,\delta) = \left\lfloor \beta(\delta)\, \xi_4^2(\widehat{\rho}_k) \right\rfloor + 1. \quad (4.31)$$

If $K_4^* \leq k$, then there is no need to sample further. Otherwise, take additional $K_4^* - k$ observations from

$N_m(\mathbf{0}, \sigma^2 R)$ independent of the pilot samples and stop sampling. The final sample size of this two-stage procedure is

$$N_4(\delta) \equiv N_4(k, \delta) = \max\{k, K_4^*(k, \delta)\}.$$

Suppose $\widehat{\rho}_{N_4}$ denotes the MLE for ρ based on N_4 samples collected from the two-stage procedure and $\widehat{\rho}_{4N_4} = g_4\left(\widehat{\rho}_{N_4}\right)$ is the MLE of ρ_4. Now, a $1-\alpha$ fixed-accuracy confidence interval of ρ_4 is given by

$$I_{4N_4} = \left[\delta^{-1}\widehat{\rho}_{4N_4},\, \delta\widehat{\rho}_{4N_4}\right],$$

and the corresponding $1-\alpha$ confidence interval of ρ is

$$J_{4N_4} = \left[g_4^{-1}\left(\delta^{-1}\widehat{\rho}_{4N_4}\right),\, g_4^{-1}\left(\delta\widehat{\rho}_{4N_4}\right)\right] \text{ where } g_4^{-1}(x) = 1 - \left(\frac{m}{m-1}\right)e^{-x}.$$

Next, we show that J_{4N_4} is also a *bounded-width* confidence interval of ρ with a non-trivial upper bound of the random width of the interval.

Lemma 4.4. *The width W_{4N_4} of the interval J_{4N_4} is bounded above by*:

$$\left(\frac{m}{m-1}\right)u_4(\delta) \text{ a.s.},$$

where

$$u_4(\delta) = (\delta^2 - 1)\delta^{-\left(2\delta^2/(\delta^2-1)\right)} < 1 \text{ for all } \delta > 1.$$

Proof. Note that $\widehat{\rho}_{4N_4} > 0$ w.p.1. This observation leads us to investigate the following function, for $x > 0$,

$$g_4^{-1}(\delta x) - g_4^{-1}\left(\delta^{-1}x\right) = \left(\frac{m}{m-1}\right)f(x), \text{ where } f(x) = e^{-\delta^{-1}x} - e^{-\delta x}.$$

The first derivative $f'(x) = \delta e^{-\delta x} - \delta^{-1}e^{-\delta^{-1}x} = 0$ gives the critical value $x_0 = 2\delta\ln\delta/(\delta^2-1) > 0$. The second derivative $f''(x_0) = (\delta^{-2} - 1)e^{-\left(\delta^{-1}x_0\right)} < 0$. Since $f(x_0) > 0$ and $f(x) \to 0$ when $x \to \infty$ or $x \to 0$, the function $f(x)$ has a global maximum at x_0. Thus, $f\left(\widehat{\rho}_{4N_4}\right) < f(x_0) = u_4(\delta)$ w.p.1, and the upper bound is established. Now, define $g(y) = y\ln y - (y-1)\ln(y-1)$ for $y > 1$. Since $g'(y) = \ln y - \ln(y-1) > 0$ for $y > 1$, $g(y)$ is an increasing function of y and $\inf\{g(y) : y > 1\} = \lim_{y\to 1^+} g(y) = 0$. Thus, $g(y) > 0$ for all $y > 1$ implies $y^y > (y-1)^{y-1}$ for all $y > 1$. Replacing $y = \delta^2$ in this inequality, we prove that $u_4(\delta) < 1$ for all $\delta > 1$. This completes the proof. □

4.4.1. Choice of Pilot Sample Size and Asymptotic Properties of N_4

In this section, we investigate asymptotic properties of the proposed two-stage procedure corresponding to g_4 transformation. We observed in Section 4.3 that both optimal sample sizes $n_2^*(\rho, \delta)$ and $n_3^*(\rho, \delta)$ had non-trivial lower bounds k_2 and k_3 respectively, which were free from unknown parameter ρ. The use of k_2 and k_3 as pilot sample sizes led to stopping only at the second stage, that is, $N_2 = K_2^*$ and $N_3 = K_3^*$, which in turn led to the second-order asymptotic efficiency results for N_2 and N_3.

In the present case of g_4 transformation, such an advantage is lost as the first derivative test suggests that $\xi_4^2(\rho)$ is a strictly decreasing function of ρ, and hence $\inf_\rho n_4^*(\rho,\delta) = 0$ for fixed δ. We no longer have a non-trivial lower bound for $n_4^*(\rho,\delta)$ which can be used as a natural choice for a pilot sample size.

As discussed in Section 4.3.1, the study of asymptotic properties of stopping variables requires pilot sample size $k(\delta)$ to be chosen as a function of δ such that it grows to ∞ as $\delta \downarrow 1$. However, the rate of growth of $k(\delta)$ should be slower than that of the optimal sample size $n_4^*(\delta)$. For such a choice of a pilot sample size, we prove that the stopping variable $N_4(\delta)$ enjoys both asymptotic first-order efficiency and asymptotic consistency properties.

Theorem 4.3. *For any choice of a pilot sample size* $k(\delta)$ *such that* $k(\delta) \to \infty$ *and* $k(\delta) = o(n_4^*(\delta))$ *as* $\delta \downarrow 1$, *the proposed two-stage procedure yields*:

(i) $\lim_{\delta\downarrow 1} N_4(\delta)/n_4^*(\delta) = 1$ *w.p.1*;

(ii) $\lim_{\delta\downarrow 1} E_\theta\left[N_4(\delta)/n_4^*(\delta)\right] = 1$ (*Asymptotic first-order efficiency*);

(iii) $\lim_{\delta\downarrow 1} P_\theta\left(\rho \in J_{4N_4(\delta)}\right) = 1-\alpha$ (*Asymptotic consistency*).

Proof. (i) Following the proof of part (i) of Theorem 4.1, we have $\lim_{\delta\downarrow 1} K_4^*(\delta)/n_4^*(\delta) = 1$ w.p.1. Using the inequality $K_4^*(\delta) \le N_4(\delta) \le K_4^*(\delta) + k(\delta)$ and that $k(\delta) = o(n_4^*(\delta))$, we complete the proof of the first part.

(ii) Using the inequality $\ln(1+x) < x$ for all $x > 0$ with $x = m/\{(m-1)(1-\rho)\} - 1$, we show

$$\frac{d}{d\rho}\xi_4^2(\rho) = \frac{2\sqrt{2}}{\sqrt{m(m-1)}}\,\xi_4(\rho)\left[\frac{\ln\left\{\frac{m}{(m-1)(1-\rho)}\right\} - \frac{1+(m-1)\rho}{(m-1)(1-\rho)}}{(m-1)^{-1}\ln^2\left\{\frac{m}{(m-1)(1-\rho)}\right\}}\right] < 0.$$

Since $\xi_4^2(\rho)$ is a strictly decreasing function of ρ, $\sup_\rho \xi_4^2(\rho) = \lim_{\rho\downarrow\frac{-1}{m-1}} \xi_4^2(\rho) = 2m/(m-1)$ which implies

$$\sup_{\delta>1} K_4^*(\delta)/n_4^*(\delta) \le \frac{1}{\xi_4^2(\rho)}\sup_{\delta>1}\xi_4^2(\widehat{\rho}_{k(\delta)}) + 1 \le \frac{2m}{(m-1)\xi_4^2(\rho)} < \infty.$$

The dominated convergence theorem yields $E_\theta\left[K_4^*(\delta)/n_4^*(\delta)\right] \to 1$ as $\delta \downarrow 1$. We complete the proof of part (ii) by dividing all sides of the inequality $K_4^*(\delta) \le N_4(\delta) \le K_4^*(\delta) + k(\delta)$ by $n_4^*(\delta)$ and then taking expectations and limit respectively.

(iii) We omit the proof of part (iii) as it is very similar to the proofs of part (iii) of Theorems 4.1 and 4.2. □

Next, we show that the choice of a pilot sample size satisfying the conditions of Theorem 4.3 exists. Using the inequality $\ln(x) < x - 1$ for all $x > 1$ with $x = m/\{(m-1)(1-\rho)\}$, we show that the asymptotic variance

$$\xi_4^2(\rho) \ge \frac{2(m-1)}{m}(1-\rho)^2,$$

which yields

$$n_4^*(\rho,\delta) \geq \left\lfloor \frac{2(m-1)}{m}\beta(\delta)(1-\rho)^2 \right\rfloor + 1.$$

Note that this particular lower bound of the optimal sample size is a strictly decreasing function of ρ with zero as its trivial lower bound. So, we follow the approach of Mukhopadhyay (1980) and consider the auxiliary stopping variable for a purely sequential sampling procedure

$$T = \inf\{n \geq 1 : n \geq \frac{2(m-1)}{m}\beta(\delta)(1-\widehat{\rho}_n + n^{-r})^2\} \text{ where } r > 0.$$

In the definition of T, the unknown parameter $1-\rho$ is estimated by $(1-\widehat{\rho}_n) + n^{-r}$ to avoid the undesirable possibility of stopping too early due to having $(1-\widehat{\rho}_n)$ close to zero. The definition of T yields (w.p.1)

$$T \geq \frac{2(m-1)}{m}\beta(\delta)T^{-2r} \quad \text{that is,} \quad T \geq \left[\frac{2(m-1)}{m}\beta(\delta)\right]^{\frac{1}{2r+1}}.$$

Therefore, for some fixed $r > 0$, a possible choice of pilot sample size would be

$$k_4(\delta) \equiv k_4(\delta, r) = \left\lfloor \left\{\frac{2(m-1)}{m}\beta(\delta)\right\}^{\frac{1}{2r+1}} \right\rfloor + 1. \tag{4.32}$$

Clearly, $k_4(\delta) \to \infty$ and $k_4(\delta)/n_4^*(\delta) \to 0$ as $\delta \downarrow 1$.

4.5. Numerical Study for Bounded-Width Intervals

In this section, we evaluate the performances of the proposed bounded-width confidence intervals of ρ via extensive sets of simulation studies. In all simulations, we used 10^5 replications. Usually, a shorter confidence interval is expected with higher sample size (inversely related) if confidence level is fixed, whereas a shorter confidence interval is expected to produce lower coverage probability if sample size is fixed.

Among the proposed methods, some will require more sample sizes than the others leading to possibly shorter intervals. In order to compare performances of various confidence intervals, we propose a new index or measure of performance of a confidence interval.

Associated with a confidence interval J, we define an index as follows:

$$\mathcal{I}_J = \frac{\text{attained coverage probability of } J - (1-\alpha)}{\text{average sample size to obtain } J \times \text{average width of } J}. \tag{4.33}$$

Intuitively, such an index for a confidence interval J represents the *gain in attained coverage probability* compared to the prefixed *nominal target coverage* $1-\alpha$ per *unit length of the interval* and per *unit average sample size*. Clearly, a higher value of this index would indicate better performance of a confidence interval in the sense of producing higher coverage probability with unit resource (sample size) and unit precision (inverse width). Typically, we would expect that the performance index of a good confidence interval would be close to zero.

In Table 6, we compare the performances of our proposed bounded-width confidence intervals $J_{1n_1^*}$,

J_{2N_2}, J_{3N_3} and J_{4N_4} where the pilot sample sizes for J_{2N_2}, J_{3N_3} and J_{4N_4} are chosen from (4.12), (4.13), and (4.32) with $r = 0.15$. Here, $\widehat{E}[N_i]$ represents the estimated expected sample size (estimator of $E_\theta[N_i(\delta)]$) and its standard error $\widehat{sd}(N_i)$ for $i = 2, 3, 4$. $\widehat{CP}_i$ and $\overline{W}_i$ represent the estimated coverage probability and the average width respectively corresponding to confidence interval J_i for $i = 1, 2, 3, 4$. According to (4.33), the performance index $\mathcal{I}_1$ of interval $J_{1n_1^*}$ is $(\widehat{CP}_1 - (1-\alpha))/(n_1^*\overline{W}_1)$, and for $i = 2, 3, 4$, the performance index $\mathcal{I}_i$ of interval J_{iN_i} is $(\widehat{CP}_i - (1-\alpha))/(\widehat{E}[N_i]\overline{W}_i)$.

The following features may be observed from the summaries of the simulation studies presented in Table 6.

1. The performance indexes $\mathcal{I}_1, \mathcal{I}_2, \mathcal{I}_3, \mathcal{I}_4$ of the proposed confidence intervals are close to zero which indicates commendable performance of all the intervals. However, there is no clear winner among these confidence intervals according to the given performance index from (4.33).

2. Under all four procedures, the expected sample sizes are nearly the same as their corresponding optimal sample sizes and the attained coverage probabilities are close to the target level of 90%.

3. The two-stage procedures corresponding to g_3 and g_4 transforms require the highest and the lowest average sample size respectively. As a result, the average width $\overline{W}_3$ is the smallest while the average width $\overline{W}_4$ is the largest among all four methods. The fact that sample size corresponding to the method with g_3 transform is higher than that of the g_1 and g_2 transforms is evident from Lemma 4.2.

4. The standard error of $N_2(\delta)$ is lowest among all two-stage stopping variables.

5. For fixed ρ, as the dimension m increases, all four methods require less sample sizes which can be explained from the respective optimal sample size formulas.

5. AN EPILOGUE

Summing up, there is neither a theoretical evidence nor a numerical evidence from Table 6 that a particular method is uniformly better than the others. It is up to the experimenter to choose one of these methods by factoring in his/her practical logistical constraints.

For example, if one is not comfortable having a method with a random sample size, he/she may use the fixed-sample-size method corresponding to g_1 transform. As δ decreases to 1, Table 8 validates the asymptotic efficiency and asymptotic consistency properties of the proposed two-stage methods given in Theorems 4.1, 4.2 and 4.3.

In Table 5, we illustrated some drawbacks of fixed-width confidence interval

$$I_N = [\widehat{\rho}_N - d, \widehat{\rho}_N + d],$$

obtained from a two-stage procedure, by presenting few cases where I_N failed to produce the desired coverage probability and argued that such a problem can arise if ρ is close to one of the boundaries of its parameter space, and the missed confidence limit $misCL$ may not be negligible.

In these cases, the optimal sample size n^* for the two-stage procedure was not very large, ranging from 38 to 54. Since by construction of the bounded-width confidence intervals J_1, J_2, J_3 and J_4, the corresponding missed confidence limits must be zero, we would like to investigate whether these intervals perform better than I_N at least under the unfriendly cases reported in Table 5.

To this end, notice that all four methods for bounded-width confidence interval estimation involve the fixed-accuracy measure δ, and if we increase (or decrease) δ, the sample sizes $n_1^*(\delta), N_2(\delta), N_3(\delta)$ and $N_4(\delta)$ decrease (or increase) leading to wider (or shorter) confidence intervals. This prompts us to think that δ can possibly play the role of a *tuning parameter* for these methods.

In Table 7, we attempt to show that for some suitably chosen values of δ, any of the bounded-width confidence intervals $J_{1n_1^*(\delta)}$, $J_{2N_2(\delta)}$, $J_{3N_3(\delta)}$, $J_{4N_4(\delta)}$ can perform better than the fixed-width confidence interval $I_N = [\widehat{\rho}_N - d, \widehat{\rho}_N + d]$ in terms of producing higher performance index formulated in (4.33).

In an attempt to find a suitable value of δ, say for the fixed-sample-size confidence interval $J_{1n_1^*(\delta)}$, we tried a number of values of δ and computed their corresponding average widths $\overline{W}_{1n_1^*(\delta)}$, an estimator of the expected width $E_\theta[W_{1n_1^*(\delta)}]$, also given in Table 6 based on 10^5 replications. We finally selected a value of δ such that $\overline{W}_{1n_1^*(\delta)} \approx 2d$. We applied this technique in the cases of all our bounded-width confidence intervals and found out a suitable value of δ for each method.

In Table 7, we present the fixed or average widths of the confidence intervals in column 5, optimal sample sizes in column 6, estimated expected sample sizes and coverage probabilities in columns 7 and 8, and finally the performance indexes in the last column. Interestingly, we now observe that the coverage probabilities of the bounded-width confidence intervals, for suitably chosen δ, are close to the target coverage $1-\alpha$ whereas the coverage probabilities of the fixed-width confidence interval are significantly smaller than $1-\alpha$.

Moreover, the performance indexes of the bounded-width intervals are considerably higher than that of the fixed-width confidence interval. Therefore, in practice, if one is interested to construct a confidence interval for ρ having some prescribed width $2d$ on an average, assuming one may be willing to relax the criteria of having exact $2d$ width, we recommend implementing a simulation technique such as ours. That will help one to locate a suitable value of δ, say δ_0, under any of the four bounded-width confidence intervals, say for $J_{1n_1^*(\delta)}$ if one prefers this particular confidence interval over others. The coverage probability of this interval $J_{1n_1^*(\delta_0)}$ will be close to the target coverage $1-\alpha$.

Now, recall that based on the asymptotic normality of $\widehat{\rho}_n$, we had constructed a large-sample $1-\alpha$ confidence interval $I_n^{(ls)}$ for ρ in (3.1) based on n i.i.d. observations. We showed in Section 3.1 that $\sup_\rho \sigma_\rho^2 = \left[\frac{1}{2}m/(m-1)\right]^3$. This implies that the width of $I_n^{(ls)}$, denoted by $W_n^{(ls)}$, is bounded from above by

$$W_n^{(ls)} = 2z_{\alpha/2}\,\sigma_{\widehat{\rho}_n}/\sqrt{n} \;\leq\; 2\frac{z_{\alpha/2}}{\sqrt{n}}\left[\frac{m}{2(m-1)}\right]^3.$$

Therefore, both $I_n^{(ls)}$ and $J_{1n_{1e}^*(\delta)}$ serve as bounded-width $1-\alpha$ confidence intervals for ρ based on the fixed-sample-size method. It is difficult to compare these two bounded-width confidence intervals theoretically, hence we compared their performances in Table 9 via simulations. In this table, $\widehat{CP}_1$ and $\widehat{CP}^{(ls)}$ represent the estimated coverage probabilities achieved by $J_{1n_{1e}^*(\delta)}$ and $I_n^{(ls)}$ respectively whereas $\overline{W}_1$ and $\overline{W}^{(ls)}$ represent the average widths of $J_{1n_{1e}^*(\delta)}$ and $I_n^{(ls)}$ respectively based on 10^5 replications.

For bounded-width confidence interval $J_{1n^*_{1e}(\delta)}$, again the fixed-accuracy measure δ can act as a *tuning* parameter. We select a value of δ, say δ_0, such that the coverage probability attained by $J_{1n^*_{1e}(\delta_0)}$ is approximately $1-\alpha$. Then, based on the same $n = n^*_{1e}(\delta_0)$ observations, we constructed the interval $I_n^{(ls)}$ and estimated its coverage probability. Since same sample sizes were used under both methods, we compared their performances by modified indexes defined as follows:

$$\mathcal{J}_1 = (\widehat{CP}_1 - (1-\alpha))/\overline{W}_1 \ \text{ and } \ \mathcal{J}^{(ls)} = (\widehat{CP}^{(ls)} - (1-\alpha))/\overline{W}^{(ls)}$$

which represent gain in coverage probabilities per unit average lengths. This should make a good practical sense in view of the earlier index from (4.33)

In Table 9, we summarize a few cases where the optimal sample sizes are small or moderate. Under all these cases, we observe that the performance index of $J_{1n^*_{1e}(\delta)}$ for some suitably chosen δ are significantly higher (sometimes more than 50 times!) than that for the large-sample confidence interval $I^{(ls)}_{n^*_{1e}(\delta)}$. This observation is expected especially since $I_n^{(ls)}$ is a valid $1-\alpha$ confidence interval for ρ only when n is large, whereas $J_{1n^*_{1e}(\delta)}$ is an exact (not a large-sample) $1-\alpha$ confidence interval for ρ with any $\delta > 1$. Therefore, in practice, the fixed-sample-size bounded-width interval J_1 is recommended over the customary large-sample confidence interval at least for those situations where available sample size is either small or moderate.

ACKNOWLEDGMENTS

We are grateful to an Associate Editor and a reviewer for their enthusiastic commentaries. We earnestly thank them. Shyamal K. De received support from grant MTR/2017/000503 under the MATRICS scheme provided by Science and Engineering Research Board, Govt. of India.

REFERENCES

Abramowitz, M. and Stegun, I. A. (1965). *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables*, first edition, New York: Dover.

Anscombe, F. J. (1952). Large-Sample Theory of Sequential Estimation, *Proceedings of Cambridge Philosophical Society* 48: 600–607.

Banerjee, S. and Mukhopadhyay, N. (2016). A General Sequential Fixed-Accuracy Confidence Interval Estimation Methodology for a Positive Parameter: Illustrations Using Health and Safety Data, *Annals of the Institute of Statistical Mathematics* 68: 541–570.

Broyden, C. G. (1965). A Class of Methods for Solving Nonlinear Simultaneous Equations, *Mathematics of Computation. American Mathematical Society* 19: 577–593.

Chow, Y. S. and Robbins, H. (1965). On the Asymptotic Theory of Fixed-Width Sequential Confidence Intervals for the Mean, *Annals of Mathematical Statistics* 36: 457–462.

Cramér, H. and Wold, H. (1936). Some Theorems on Distribution Functions, *Journal of the London Mathematical Society* 11: 290–294.

De, S. K. (2014). Modified Three-Stage Sampling for Fixed-Width Interval Estimation of the Common Variance of Equi-Correlated Normal Distributions, *Sequential Analysis* 33: 87–111.

De, S. K. and Mukhopadhyay, N. (2015). Fixed Accuracy Interval Estimation of the Common Variance in an Equi-Correlated Normal Distribution, *Sequential Analysis* 34: 1–23.

Ehrenfeld, S. and Littauer, S. B. (1964). *Introduction to Statistical Method*, New York: McGraw-Hill.

Ghezzi, D. J. and Zacks, S. (2005). Inference on the Common Variance of Correlated Normal Random Variables, *Communications in Statistics – Theory and Methods* 34: 1517–1531.

Ghosh, B. K. and Sen, P. K. (1991). *Handbook of Sequential Analysis, edited volume*, New York: Dekker.

Ghosh, M. and Mukhopadhyay, N. (1981). Consistency and Asymptotic Efficiency of Two-Stage and Sequential Procedures, *Sankhyā, Series A* 43: 220–227.

Ghosh, M., Mukhopadhyay, N., and Sen, P. K. (1997). *Sequential Estimation*, New York: Wiley.

Gut, A. (2009). *Stopped Random Walks: Limit Theorems and Applications*, second edition, New York: Springer.

Haner, D. M. (2012). On Two-Stage and Three-Stage Sampling Procedures for Fixed-Width Interval Estimation of the Common Variance of Correlated Normal Distributions, *Ph.D. Dissertation*, Binghamton University, New York, USA.

Haner, D. M. and Zacks, S. (2013). On Two-Stage Sampling for Fixed-Width Interval Estimation of the Common Variance of Equi-correlated Normal Distributions, *Sequential Analysis* 32: 1–13.

Khan, R. A. (1969). A General Method of Determining Fixed-Width Confidence Intervals, *Annals of Mathematical Statistics* 40: 704–709.

Lehmann, E. L. (1997). *Testing Statistical Hypotheses*, second edition, New York: Springer.

Mahalanobis, P. C. (1940). A Sample Survey of Acreage Under Jute in Bengal, with Discussion on Planning of Experiments, *Proceedings of Second Indian Statistical Conference*, Calcutta, India: Statistical Publishing Society.

Mukhopadhyay, N. (1980). A Consistent and Asymptotically Efficient Two-Stage Procedure to Construct Fixed Width Confidence Intervals for the Mean, *Metrika* 27: 281–284.

Mukhopadhyay, N. (1982). Stein's Two-Stage Procedure and Exact Consistency, *Scandinavian Actuarial Journal* 1982: 110–122.

Mukhopadhyay, N. (2000). *Probability and Statistical Inference*, New York: Dekker.

Mukhopadhyay, N. and Banerjee, S. (2014). Purely Sequential and Two-Stage Fixed-Accuracy Confidence Interval Estimation Methods for Count Data from Negative Binomial Distributions in Statistical Ecology: One-Sample and Two-Sample Problems, *Sequential Analysis* 33: 251–285.

Mukhopadhyay, N. and Banerjee, S. (2015a). Purely Sequential and Two-Stage Bounded-Length Confidence Intervals for the Bernoulli Parameter with Illustrations from Health Studies and Ecology, in *Ordered Data Analysis, Modeling and Health Research Methods*, *H. N. Nagaraja's 60th Birthday Festschrift Volume*, P. Choudhary, C. H. Nagaraja, and H. K. T. Ng, eds., pp. 211–234, New York: Springer.

Mukhopadhyay, N. and Banerjee, S. (2015b). Sequential Negative Binomial Problems and Statistical Ecology: A Selected Review with New Directions, *Statistical Methodology* 26: 34–60.

Mukhopadhyay, N. and Chattopadhyay, B. (2012). A Tribute to Frank Anscombe and Random Central Limit Theorem from 1952, *Sequential Analysis* 31: 265–277.

Mukhopadhyay, N. and de Silva, B. M. (2009). *Sequential Methods and Their Applications*, Boca Raton: CRC.

Mukhopadhyay, N. and Zhuang, Y. (2016). On Fixed-Accuracy and Bounded Accuracy Confidence Interval Estimation Problems in Fisher's "Nile" Example, *Sequential Analysis* 35: 516–535.

Nadás, A. (1969). An Extension of a Theorem of Chow and Robbins on Sequential Confidence Intervals for the Mean, *Annals of Mathematical Statistics* 40: 667–671.

Rade, L. and Westergren, B. (1990). *Beta Mathematics Handbook: Concepts, Theorems, Methods, Algorithms, Formulas, Graphs, Tables*, second edition, Boca Raton: CRC Press.

Rao, C. R. (1973). *Linear Statistical Inference*, second edition, New York: Wiley.

SenGufta, A. (1987). On Tests for Equicorrelation Coefficient of a Standard Symmetric Multivariate Normal Distribution, *Australian & New Zealand Journal of Statistics* 29: 49–59.

Siegmund, D. (1985). *Sequential Analysis: Tests and Confidence Intervals*, New York: Springer.

Simons, G. and Zacks, S. (1967). A Sequential Estimation of Tail Probabilities in Exponential Distributions With A Prescribed Proportional Closeness, *Technical Report 21*: Stanford University.

Sproule, R. N. (1974). Asymptotic Properties of U-statistics, *Transactions of the American Mathematical Society* 199: 55–64.

Stein, C. (1945). A Two Sample Test for a Linear Hypothesis Whose Power Is Independent of the Variance, *Annals of Mathematical Statistics* 16: 243–258.

Stein, C. (1949). Some Problems in Sequential Estimation, *Econometrica* 17: 77–78.

Willson, L. J. and Folks, J. L. (1983). Sequential Estimation of the Mean of the Negative Binomial Distribution, *Communications in Statistics-Sequential Analysis* 2: 55–70.

Zacks, S. (1966). Sequential Estimation of the Mean of a Log-Normal Distribution Having a Prescribed Proportional Closeness, *Annals of Mathematical Statistics* 37: 1688–1696.

Zacks, S. (2001). The Operating Characteristics of Sequential Procedures in Reliability, Handbook of Statistics 20: 789–811.

Zacks, S. (2009a). *Stage-Wise Adaptive Designs*, New York: Wiley.

Zacks, S. (2009b). The Exact Distributions of the Stopping Times and Their Functionals in Two-Stage and Sequential Fixed-Width Confidence Intervals of the Exponential Parameter, *Sequential Analysis* 28: 69–81.

Zacks, S. (2017). *Sample Path Analysis and Distributions of Boundary Crossing Times*, Lecture Notes in Mathematics, New York: Springer.

Zacks, S. and Khan, A. (2011). Two-Stage and Sequential Estimation of the Scale Parameter of a Gamma Distribution with Fixed-Width Intervals, *Sequential Analysis* 30: 297–307.

Zacks, S. and Mukhopadhyay, N. (2006a). Exact Risks of Sequential Point Estimators of the Exponential Parameter, *Sequential Analysis* 25: 203–220.

Zacks, S. and Mukhopadhyay, N. (2006b). Bounded Risk Estimation of the Exponential Parameter in a Two-Stage Sampling, *Sequential Analysis* 25: 437–452.

Zacks, S. and Ramig, P. F. (1987). Confidence Intervals for the Common Variance of Equicorrelated Normal Random Variables, in *Contributions to the Theory and Applications of Statistics*, volume in honor of Herbert Solomon, A. E. Gelfand, ed., pp. 511–544, New York: Academic Press.

Table 6. Comparing the performances of the four bounded-width intervals $J_{1n_1^*}$, J_{2N_2}, J_{3N_3} and J_{4N_4} when $1-\alpha=0.9$, $\sigma=1$, $\delta=1.1$, and $r=0.15$ for different choices of m and ρ

m	n_1^*	ρ	n_2^* k_2	n_3^* k_3	n_4^* k_4	$\widehat{E}(N_2)$ $\widehat{sd}(N_2)$	$\widehat{E}(N_3)$ $\widehat{sd}(N_3)$	$\widehat{E}(N_4)$ $\widehat{sd}(N_4)$	$\widehat{CP}_1$ $\overline{W_1}$	$\widehat{CP}_2$ $\overline{W_2}$	$\widehat{CP}_3$ $\overline{W_3}$	$\widehat{CP}_4$ $\overline{W_4}$	$\mathcal{I}_1\times10^5$	$\mathcal{I}_2\times10^5$	$\mathcal{I}_3\times10^5$	$\mathcal{I}_4\times10^5$
3	894	-0.25	818 224	1064 894	747 100	817.70 7.53	1063.93 9.36	746.54 20.94	0.9004 0.0397	0.8998 0.0415	0.8983 0.0364	0.8996 0.0435	1.267	-0.5004	-4.366	-1.201
		0.25	651 224	1589 894	465 100	651.62 15.55	1589.18 30.65	466.63 35.60	0.8987 0.0714	0.9009 0.0836	0.8998 0.0536	0.8991 0.0989	-2.069	1.671	-0.282	-1.863
		0.75	443 224	2626 894	194 100	443.99 12.45	2625.68 36.18	195.33 20.29	0.8990 0.0397	0.8999 0.0564	0.8992 0.0232	0.8988 0.0849	-2.870	-0.399	-1.314	-7.413
5	745	-0.10	700 187	843 745	657 115	699.64 4.71	843.25 5.52	656.47 11.37	0.8994 0.0252	0.8973 0.0259	0.8987 0.0236	0.9004 0.0268	-2.986	-14.740	-6.569	2.384
		0.38	541 187	1331 745	385 115	541.52 13.08	1331.36 25.74	386.25 25.21	0.9010 0.0595	0.8985 0.0698	0.8987 0.0445	0.8987 0.0827	2.234	-3.918	-2.142	-3.975
		0.80	365 187	2214 745	157 115	365.91 10.18	2213.05 29.84	158.17 14.04	0.9001 0.0321	0.9006 0.0458	0.8997 0.0186	0.8993 0.0696	0.377	3.823	-0.680	-6.818
10	662	0	629 166	734 662	597 126	628.83 3.56	733.86 4.02	596.72 7.78	0.8993 0.0191	0.9001 0.0196	0.9012 0.0181	0.8997 0.0201	-5.466	0.894	9.182	-2.751
		0.44	484 166	1171 662	347 126	484.72 11.57	1170.83 22.53	348.50 20.28	0.8988 0.0529	0.8998 0.0618	0.9004 0.0398	0.8995 0.0730	-3.315	-0.534	0.816	-2.044
		0.80	336 166	1902 662	152 126	336.81 9.59	1900.92 27.67	152.93 12.16	0.9009 0.0313	0.9004 0.0439	0.8983 0.0185	0.8997 0.0651	4.151	2.570	-4.957	-2.611

Table 7. Performances of bounded-width intervals $J_{1n_1^*}$, J_{2N_2}, J_{3N_3} and J_{4N_4} with $k_4 = 10$ for the cases in Table 5 where the fixed-width interval $[\widehat{\rho}_N - d, \widehat{\rho}_N + d]$ for pilot sample size $k = 10$ was producing less coverage probabilities than targeted

interval	m	ρ	$1-\alpha$	fixed or avg width	optimal ss	expected ss	coverage probability	$\mathcal{I}_J \times 10^3$
FW interval I	3	0.9	0.95	$2d = 0.1$	41	64.96	0.8971	-8.143
J_1 with $\delta = 1.66$				0.1000	45	45	0.9467	-0.738
J_2 with $\delta = 1.38$				0.0983	44	44.86	0.9505	0.122
J_3 with $\delta = 2.54$				0.1003	47	46.51	0.9431	- 1.489
J_4 with $\delta = 1.20$				0.0980	42	45.59	0.9444	-1.249
FW interval I	3	-0.42	0.90	$2d = 0.06$	52	72.83	0.8059	-21.528
J_1 with $\delta = 1.47$				0.0600	55	55	0.8982	-0.540
J_2 with $\delta = 1.46$				0.0600	56	55.56	0.8987	-0.402
J_3 with $\delta = 1.48$				0.0590	56	56.30	0.8989	- 0.337
J_4 with $\delta = 1.45$				0.0599	56	55.90	0.8962	-1.123
FW interval I	5	-0.18	0.95	$2d = 0.06$	47	58.46	0.8470	-29.367
J_1 with $\delta = 1.55$				0.0594	51	51	0.9502	0.0594
J_2 with $\delta = 1.54$				0.0596	51	50.52	0.9465	-1.149
J_3 with $\delta = 1.58$				0.0598	49	49.11	0.9495	- 0.163
J_4 with $\delta = 1.53$				0.0598	51	50.54	0.9476	-0.781
FW interval I	5	-0.14	0.90	$2d = 0.08$	43	50.42	0.8032	-23.999
J_1 with $\delta = 1.48$				0.0797	45	45	0.8977	-0.639
J_2 with $\delta = 1.46$				0.0790	46	45.57	0.8966	-0.952
J_3 with $\delta = 1.51$				0.0794	44	44.12	0.8985	- 0.442
J_4 with $\delta = 1.45$				0.0799	45	45.11	0.8967	-0.912
FW interval I	10	0	0.95	$2d = 0.08$	54	58.93	0.8558	-19.974
J_1 with $\delta = 1.48$				0.0790	56	56	0.9487	-0.287
J_2 with $\delta = 1.46$				0.0786	57	57.07	0.9481	-0.421
J_3 with $\delta = 1.52$				0.0795	54	54.47	0.9489	- 0.254
J_4 with $\delta = 1.45$				0.0797	56	56.21	0.9480	-0.456
FW interval I	10	-0.03	0.90	$2d = 0.06$	38	43.51	0.7935	-40.797
J_1 with $\delta = 1.48$				0.0597	40	40	0.8994	-0.239
J_2 with $\delta = 1.47$				0.0600	40	39.50	0.8977	-0.974
J_3 with $\delta = 1.51$				0.0600	39	38.65	0.8980	- 0.872
J_4 with $\delta = 1.46$				0.0604	39	39.42	0.8980	-0.861

Table 8. Validating asymptotic efficiency and consistency properties the bounded-width confidence intervals J_{2N_2}, J_{3N_3} and J_{4N_4} with $1-\alpha=0.9$, $m=3$, $\rho=0.25$, and $r=0.15$

δ	$n_2^*(\delta)$	$n_3^*(\delta)$	$n_4^*(\delta)$	$k_4(\delta)$	$\left\lvert\frac{\widehat{E}(N_2)}{n_2^*}-1\right\rvert$	$\left\lvert\frac{\widehat{E}(N_3)}{n_3^*}-1\right\rvert$	$\left\lvert\frac{\widehat{E}(N_4)}{n_4^*}-1\right\rvert$	$\lvert\widehat{CP}_2-0.9\rvert$	$\lvert\widehat{CP}_3-0.9\rvert$	$\lvert\widehat{CP}_4-0.9\rvert$
1.60	27	66	20	9	0.01543	0.00034	0.00568	0.00384	0.00420	0.01044
1.45	43	105	31	13	0.01149	0.00150	0.02023	0.00287	0.00173	0.00926
1.30	86	210	62	22	0.00652	0.00115	0.00788	0.00234	0.00171	0.00381
1.15	303	739	217	56	0.00134	0.00021	0.00284	0.00002	0.00046	0.00314
1.05	2485	6062	1775	280	0.00006	8.8×10^{-7}	0.00054	0.00016	0.00018	0.00014

Table 9. Comparing performances of large sample confidence interval $I_n^{(ls)}$ and bounded-width confidence interval J_{1n} with same sample size $n=n_{1e}^*$ for confidence level $1-\alpha=0.90$ and different choices of m, ρ, and δ

m	ρ	δ	n_{1e}^*	$\widehat{CP}_1$	$\widehat{CP}^{(ls)}$	$\overline{W}_1$	$\overline{W}^{(ls)}$	$\mathcal{J}_1$	$\mathcal{J}^{(ls)}$	$\frac{\lvert\mathcal{J}^{(ls)}\rvert}{\mathcal{J}_1}$
3	-0.2	1.53	46	0.9009	0.8872	0.2030	0.2004	0.0044	-0.0639	14.58
		1.95	19	0.9005	0.8724	0.3161	0.3096	0.0017	-0.0889	53.08
	0.25	1.56	42	0.9007	0.8861	0.3226	0.3239	0.0020	-0.0428	20.90
		1.86	22	0.9010	0.8739	0.4373	0.4405	0.0023	-0.0591	25.36
5	-0.1	1.50	42	0.9003	0.8877	0.1078	0.1057	0.0024	-0.1168	48.44
		1.92	17	0.9026	0.8697	0.1749	0.1655	0.0149	-0.1831	12.22
	0.375	1.50	42	0.9003	0.8879	0.2464	0.2471	0.0011	-0.0491	46.49
		1.85	19	0.9019	0.8733	0.3617	0.3608	0.0051	-0.0739	14.46
10	0	1.43	48	0.9005	0.8883	0.0719	0.0705	0.0072	-0.1661	22.95
		1.85	17	0.9022	0.8673	0.1248	0.1176	0.0177	-0.2776	15.68
	0.45	1.46	43	0.9005	0.8894	0.2052	0.2051	0.0023	-0.0519	22.65
		1.93	15	0.9014	0.8689	0.3409	0.3388	0.0041	-0.0919	22.22