Modified Three-Stage Sampling for Fixed-Width Interval Estimation of The Common Variance of Equi-correlated Normal Distributions

Shyamal K. De

Department of Mathematical Sciences, Binghamton University,

Binghamton, New York, USA

Abstract: In this work, we present a modified three-stage sampling procedure to construct fixed-width confidence intervals of the common variance of equi-correlated normal distributions. We derive the exact distribution of the stopping variable of this sampling scheme and also the exact distribution of the estimator of the common variance at stopping. The modified three-stage sampling substantially reduces the expected sample size compared to that of the two-stage sampling scheme of Haner and Zacks (2013). The coverage probabilities of the proposed interval estimators are computed exactly and are compared with the coverage probabilities obtained by two-stage sampling. We derive exact formulae for the functionals of the stopping variable and the estimator of the common variance at stopping.

Keywords: Common variance; Equi-correlated normal; Exact distribution of stopping variable; Fixed-Width interval estimation; Three-Stage sampling; Two-Stage sampling.

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

1. INTRODUCTION

Correlated multinormal models, as mentioned by Ghezzi and Zacks (2005), are commonly used in the ANOVA design with repeated measurements. Suppose there are n experimental units, and each unit is

Address correspondence to Shyamal K. De, Department of Mathematical Sciences, Library North 2200, Binghamton University, 4400 Vestal Parkway East, Binghamton, New York 13902-6000, USA; Fax: 607-777-2450; E-mail: shyamalkd@gmail.com

measured on a variable at m time points. Therefore, m measurements from a unit are correlated whereas the measurements from different units may be assumed independent. In an ANOVA design, where the $m \times 1$ measurement vector from a unit is considered as the response variable, one can model the error part by an $m-$ variate normal distribution with mean $\mathbf{0}$ and covariance matrix Σ. In this work, we investigate the estimation of Σ with a special structure. Formally, consider n independently and identically distributed $m-$ dimensional vectors $\mathbf{X_1}, \mathbf{X_2}, \ldots, \mathbf{X_n}$ from multivariate normal with mean $\mathbf{0}$ and covariance matrix $\Sigma = \sigma^2 R$, where R is the unknown correlation matrix. In this paper, we consider the equi-correlated model of Zacks and Ramig (1987), Ghezzi and Zacks (2005), and Haner and Zacks (2013), that is, each entry of R is the common correlation $\rho \in (-\frac{1}{m-1}, 1)$ between any two components of $\mathbf{X_1}$. For a pre-specified $\alpha \in (0, 1)$, the problem is to construct fixed-width $(1 - \alpha)$- level confidence intervals of the common variance σ^2 in the presence of nuisance parameter ρ.

The literature on fixed-width interval estimation and multi stage sampling schemes is quite rich. Dantzig (1940) proved the non-existence of fixed-sample procedures that can produce a fixed-width confidence interval with preassigned confidence level. Around the same time, Mahalanobis (1940) introduced the ground-breaking idea of using pilot samples to control margin of errors. The seminal paper by Stein (1945) first showed that fixed-width confidence intervals with prescribed confidence level can be constructed by two-stage sampling procedures. Later, Mukhopadhyay (1976) introduced the idea of three-stage sampling to construct fixed-width confidence intervals for the mean of a population with unknown variance. Hall (1981), Mukhopadhyay (1990), and others developed asymptotic theory for three-stage procedures. Hall (1981) showed that three-stage sampling combines the simplicity of Stein's two-stage sampling with the efficiency of the fully sequential procedure of Chow and Robbins (1965). The development of multi stage sampling techniques can be found in Ghosh and Mukhopadhyay (1981), Ghosh and Sen (1990), Mukhopadhyay (1990), Mukhopadhyay and Solanky (1994), Ghosh et al. (1997), Mukhopadhyay and de Silva (2009), Zacks (2009), etc.

Interval estimation of the common variance of correlated multinormal distributions was first considered by Zacks and Ramig (1987). They developed the uniform minimum variance unbiased estimators and the maximum likelihood estimators of σ^2 and ρ. Ghezzi and Zacks (2005) developed three equivariant estimators of σ^2 for the equi-correlated model and generalized these estimators for the general case of arbitrary correlation matrix R. Later, the dissertation of Haner (2012) and the subsequent paper by Haner and Zacks (2013) developed a nice theory and methodologies for fixed-width interval estimation of σ^2. Haner (2012) considered a number of two-stage and three-stage sampling schemes based on the estimators proposed by Ghezzi and Zacks (2005). For the equi-correlated model, Haner and Zacks (2013), proposed a two-stage, Stein type, sampling procedure and derived the exact distribution of the stopping variable along with the exact distribution of maximum likelihood estimator of σ^2 at stopping.

Motivated by the previous studies, we develop a modified version of three-stage sampling scheme to construct fixed-width confidence intervals of σ^2 in the equi-correlated case. We derive the exact distributions of the stopping variables, the exact distributions of the estimators of σ^2 at the time of stopping, and the exact formula for coverage probabilities. Due to the modification in the sampling scheme, our three-stage procedures require significantly less expected sample size compared to the expected sample size for the two-stage sampling in Haner and Zacks (2013). Thus, our method reduces the expected sampling cost of the experiment. Moreover, the coverage probabilities of the interval estimators obtained by the modified three-stage procedure are close to the ones achieved by the two-stage procedure in Haner and Zacks (2013).

The next section provides theoretical preliminaries for developing the modified three-stage procedure. In Section 3, we describe the modified three-stage sampling procedure and clarify how it is different from the three-stage procedure. The exact distribution of stopping variable and its expectation and variance are derived in Section 4. In Section 5, we provide the formula to compute the coverage probability by first computing the exact distribution of the maximum likelihood estimator of σ^2 at stopping. We also derive the mean, variance, and mean squared error (MSE) of the estimator of σ^2 at stopping in this section. Performance of our modified three-stage procedure is compared with the two-stage procedure of Haner and Zacks (2013) in Sections 4 and 5. In Appendix, we compare our modified three-stage method with the three-stage procedure and also provide the R codes for exact computation of the CDF of stopping variable, expectation and variance of the stopping variable, the coverage probability, and the distribution of the estimator of σ^2 at stopping. The R codes are also available in the webpage `https://sites.google.com/site/shyamalkd/`.

2. PRELIMINARY THEORY

This section discusses some preliminary theory needed to describe the modified three-stage procedure in the next section. This theory is also stated in Haner and Zacks (2013).

In order to obtain a sequence of uncorrelated multinormal vectors $\mathbf{Y_1}, \mathbf{Y_2}, \ldots$, from the original observations $\mathbf{X_1}, \mathbf{X_2}, \ldots$, let us first use the transformation $\mathbf{Y}_i = H\mathbf{X}_i$, for $i = 1, 2, \ldots$, where H is the $m \times m$

Helmert orthogonal matrix, i.e.,

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

The transformed random vectors $\mathbf{Y_1}, \mathbf{Y_2}, \ldots$ follow $N_m(\mathbf{0},\, \sigma^2\Lambda)$, where $\Lambda = HRH^T = \text{diag}(\lambda_1, \ldots, \lambda_m)$ is a diagonal matrix, and $\lambda_1, \ldots, \lambda_m$ are eigenvalues of R. Since for the equi-correlated model, the correlation matrix $R = (1-\rho)I_m + \rho J$, where I_m is the $m \times m$ identity matrix and $J = \mathbf{1}\mathbf{1}^T$ is a matrix of $1'$ s, the eigenvalues have the following form

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

Note that the condition $-\frac{1}{m-1} < \rho < 1$ is needed to ensure $\lambda_1 > 0$. Given $\mathbf{Y_1}, \ldots, \mathbf{Y_n}$, the likelihood function of (σ^2, ρ) is

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

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

Notice that (V_{1n}, V_{2n}) is a complete sufficient statistics (Lehmann, 1997) of (σ^2, ρ), and thus, it is natural to consider estimators based on (V_{1n}, V_{2n}). The properties of (V_{1n}, V_{2n}) that are crucial for the derivation of exact distributions are as follows:

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

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

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

Zacks and Ramig (1987) show that the maximum likelihood estimator of σ^2 and ρ are

$$\hat{\sigma}_n^2 = \frac{V_{1n} + V_{2n}}{nm} \qquad \text{and} \tag{2.1}$$

$$\hat{\rho}_n = \frac{V_{1n} - V_{2n}(m-1)^{-1}}{V_{1n} + V_{2n}}. \tag{2.2}$$

These estimators are strongly consistent and asymptotically normal. The asymptotic distribution of $\hat{\sigma}_n^2$ is

$$N\left(\sigma^2,\, \frac{2\sigma^4}{nm}\left\{1 + (m-1)\rho^2\right\}\right).$$

For some pre-determined width 2δ, let us consider the fixed-width confidence interval $\left(\hat{\sigma}_n^2 - \delta, \hat{\sigma}_n^2 + \delta\right)$. We need sample size n such that

$$P\left(\hat{\sigma}_n^2 - \delta < \sigma^2 < \hat{\sigma}_n^2 + \delta\right) \geq 1 - \alpha. \tag{2.3}$$

Solving (2.3) based on the asymptotic distribution of $\hat{\sigma}_n^2$, the sample size should be at least

$$n^0\left(\delta, \sigma, \rho\right) = \left\lfloor \frac{2\sigma^4 Z_{\alpha/2}^2\left(1 + (m-1)\rho^2\right)}{m\delta^2} \right\rfloor + 1, \tag{2.4}$$

where $Z_{\alpha/2}$ is the $\alpha/2$- quantile of standard normal distribution. We can replace $Z_{\alpha/2}^2$ by a higher value $t_{k,\frac{\alpha}{2}}^2$ in (2.4) to get a conservative choice of sample size as

$$n_1^0\left(\delta, \sigma, \rho\right) = \left\lfloor \frac{2\sigma^4 t_{k,\frac{\alpha}{2}}^2\left(1 + (m-1)\rho^2\right)}{m\delta^2} \right\rfloor + 1 = \left\lfloor \sigma^4\left(1 + (m-1)\rho^2\right)\beta_k(\delta) \right\rfloor + 1, \tag{2.5}$$

where $\beta_k(\delta) = 2t_{k,\frac{\alpha}{2}}^2/m\delta^2$ is a function of k and δ, and $t_{k,\frac{\alpha}{2}}$ is the $\alpha/2$- quantile of $t-$ distribution with k degrees of freedom. For large values of k, $t_{k,\frac{\alpha}{2}}$ is close to $Z_{\alpha/2}$, and thus, n_1^0 is only slightly bigger than n^0. We will use the conservative choice of (2.5) in the next sections.

3. MODIFIED THREE-STAGE SAMPLING

In this section, we present a modified three-stage sampling procedure to approximate the nominal coverage probability. Since we do not know σ^2 and ρ in (2.5) we need at least two-stage sampling to achieve the coverage probability (at least approximately) as $(1-\alpha)$.

In a Stein type two-stage procedure, the first stage estimates the unknown quantities σ^2, ρ, and $n_1^0\left(\delta, \sigma, \rho\right)$ based on some initial sample of size k. In the second stage (if necessary), additional $\left(n_1^0\left(\delta, \hat{\sigma}_k, \hat{\rho}_k\right) - k\right)$ samples are drawn and σ^2 is estimated based on $n_1^0\left(\delta, \hat{\sigma}_k, \hat{\rho}_k\right)$ samples. Haner and Zacks (2013) show that the confidence interval $\left(\hat{\sigma}^2_{n_1^0} - \delta, \hat{\sigma}^2_{n_1^0} + \delta\right)$ from two-stage sampling can yield approximately $(1-\alpha)$ coverage probability.

We extend this idea of two-stage sampling to the following three-stage sampling with some modification.

Stage I: For some fixed k, draw an initial random sample $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_k$ from m-variate normal distribution with mean **0** and covariance matrix $\sigma^2 R$. Based on this sample, compute (V_{1k}, V_{2k}) and the MLEs $\hat{\sigma}_k^2$ and $\hat{\rho}_k$ as in (2.1) and (2.2) respectively. Then compute the random variable (i.e., the estimated sample size)

$$K^* = \hat{\sigma}_k^4 \left\{1 + (m-1)\hat{\rho}_k^2\right\} \beta_k(\delta),$$

where $\beta_k(\delta)$ is defined in (2.5). Suppose N is the stopping variable, i.e., the random sample size for this three-stage procedure.

If $K^* \leq k$, stop sampling and set $N = k$ and $\hat{\sigma}_N^2 = \hat{\sigma}_k^2$. Otherwise, proceed to Stage II.

Stage II: Draw another K_p random samples from $N_m\left(\mathbf{0},\ \sigma^2 R\right)$ independent of the initial k samples, where

$$K_p = \lfloor (K^* - k)\, p \rfloor + 1,$$

and $p \in (0, 1)$ determines the proportion of additional $(K^* - k)$ samples to be considered in the second stage. Label the additional K_p observations as $\boldsymbol{X}^*_{k+1}, \ldots, \boldsymbol{X}^*_{k+K_p}$ and transform them to $\mathbf{Y}_i^* = H\boldsymbol{X}_i^*$, for $i = k+1, \ldots, k+K_p$. Define

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

and compute the estimators of σ^2 and ρ using (2.1) and (2.2) as

$$\hat{\sigma}^2_{k+K_p} = \frac{V_{1k} + V^*_{1K_p} + V_{2k} + V^*_{2K_p}}{m(k+K_p)} \quad \text{and} \quad \hat{\rho}_{k+K_p} = \frac{V_{1k} + V^*_{1K_p} - \left(V_{2k} + V^*_{2K_p}\right)(m-1)^{-1}}{V_{1k} + V^*_{1K_p} + V_{2k} + V^*_{2K_p}}.$$

Then estimate the sample size of (2.5) based on $(k + K_p)$ samples as

$$\widetilde{K} = \left\lfloor \hat{\sigma}^4_{k+K_p} \left\{1 + (m-1)\hat{\rho}^2_{k+K_p}\right\} \beta_{k+K_p}(\delta) \right\rfloor + 1.$$

If $\widetilde{K} \leq k + K_p$, stop sampling and set $N = k + K_p$ and $\hat{\sigma}^2_N = \hat{\sigma}^2_{k+K_p}$. Otherwise, set $N = \widetilde{K}$ and proceed to Stage III.

Stage III: Draw additional $\widetilde{N} = \widetilde{K} - (k + K_p)$ random samples from $N_m\left(\mathbf{0}, \sigma^2 R\right)$ independent of the previous $k + K_p$ samples and label them as $\widetilde{\boldsymbol{X}}_{k+K_p+1}, \ldots, \widetilde{\boldsymbol{X}}_{\widetilde{K}}$. Transform these observations to $\widetilde{\mathbf{Y}}_i = H\widetilde{\boldsymbol{X}}_i$, for $i = k + K_p + 1, \ldots, \widetilde{K}$ and define

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

Finally, estimate σ^2 by $\hat{\sigma}^2_{\widetilde{K}}$ and set $\hat{\sigma}^2_N = \hat{\sigma}^2_{\widetilde{K}}$, where

$$\hat{\sigma}^2_{\widetilde{K}} = \frac{V_{1k} + V^*_{1K_p} + \widetilde{V}_{1\widetilde{N}} + V_{2k} + V^*_{2K_p} + \widetilde{V}_{2\widetilde{N}}}{m\widetilde{K}}.$$

The fixed-width confidence interval for σ^2 is given by $\left(\hat{\sigma}^2_N - \delta, \hat{\sigma}^2_N + \delta\right)$.

Remark 3.1. Throughout this paper, we refer the triple sampling procedure proposed by Mukhopadhyay (1976), and later treated by Hall (1981), Mukhopadhyay (1990), Ghosh et al. (1997, page 159), and Mukhopadhyay and de Silva (2009, page 124) as "three-stage" procedure. According to the three-stage procedure, we should estimate the oracle sample size n_1^0 by K^* in the first stage and then proceed to the second stage if $\lfloor pK^* \rfloor + 1 > k$ where $p \in (0, 1)$ is some fixed proportion. The modified three-stage procedure draws samples in the second stage if $K^* > k$. In the second stage, the three-stage method would consider additional $(\lfloor pK^* \rfloor + 1 - k)$ samples whereas the modified method considers $K_p = \lfloor (K^* - k)\, p \rfloor + 1$ additional samples. The modified three-stage procedure draws approximately $\lfloor k(1 - p) \rfloor$ more samples in the second stage compared to the three-stage scheme. Due to this extra sampling, our modified scheme requires slightly higher expected sample size, but increases the coverage probability, reduces the variance of the stopping time, and also reduces the bias in the estimation of σ^2. The reader is referred to Table 4 in appendix for a detailed comparison of the modified three-stage and the three-stage procedure. We derive the exact formulas for the distributions of stopping variable and the estimator of σ^2 at stopping for the modified three-stage procedure. Such exact formulas can also be derived for the three-stage procedure in a similar way.

Remark 3.2. If $p = 1$, we have a pure three-stage procedure, and if $p = 0$, our method is almost like a two-stage procedure. By selecting an appropriate value of p (neither too small nor too large), we gain the advantage of pure three-stage sampling, i.e, higher coverage probability and, at the same time, the benefit of two-stage sampling, i.e, lower expected sample size compared to that of a pure three-stage sampling. In this paper, we choose moderate values of p such as 0.3, 0.5, and 0.7 and show that they lead to substantial reduction of sample size compared to that of a two-stage procedure while the coverage probability for the modified three-stage method remains almost the same as the two-stage procedure. In numerous three-stage procedures, a popular choice of p is 0.5.

Remark 3.3. The modified three-stage procedure, presented in Sections 3, 4, and 5, is based on the MLEs of σ^2 and ρ. However, our procedure can be applied using any other estimators of σ^2 and ρ as long as the asymptotic distribution of the estimators are known and the formula of sample size n to achieve $(1-\alpha)$ level of confidence can be explicitly written as a function of σ^2 and ρ.

4. THE EXACT DISTRIBUTION OF N AND ITS FUNCTIONALS

In this section, we derive the exact formula for the distribution function of N and its expectation and variance. The stopping variable is

$$N = \begin{cases} k & \text{if } K^* \leq k, \\ k + K_p & \text{if } K^* > k \ \text{ and } \ \widetilde{K} \leq k + K_p, \\ \widetilde{K} & \text{if } K^* > k \ \text{ and } \ \widetilde{K} > k + K_p. \end{cases} \tag{4.1}$$

4.1. Derivation of The CDF of N and Its Functionals

First, let us introduce some notations to describe the results of this section. Suppose $F_{V_{1n}}$ and $F_{V_{2n}}$ are the cumulative distribution functions of V_{1n} and V_{2n} respectively. Since $V_{1n} \sim \sigma^2 (1 + (m-1)\rho) \chi_n^2$ and $V_{2n} \sim \sigma^2 (1-\rho) \chi^2_{n(m-1)}$,

$$F_{V_{1n}}(x) = F_{\chi_n^2}\left[\frac{x}{\sigma^2 \{1 + (m-1)\rho\}}\right] \quad \text{and} \quad F_{V_{2n}}(x) = F_{\chi^2_{n(m-1)}}\left[\frac{x}{\sigma^2 (1-\rho)}\right],$$

where $F_{\chi_n^2}$ is the CDF of χ_n^2 distribution. Note that, conditional on K_p, $V^*_{1K_p}$ and $V^*_{2K_p}$ are independent, and

$$V^*_{1K_p} \sim \sigma^2 \{1 + (m-1)\rho\} \chi^2_{K_p} \quad \text{and} \quad V^*_{2K_p} \sim \sigma^2 (1-\rho) \chi^2_{K_p(m-1)}.$$

Therefore, given $K_p = j$, the cumulative distribution functions of V^*_{1j} and V^*_{2j} are $F_{V_{1j}}$ and $F_{V_{2j}}$ respectively. Similarly, conditional on $\widetilde{N} = \widetilde{n}$, $\widetilde{V}_{1\widetilde{N}}$ and $\widetilde{V}_{2\widetilde{N}}$ are independent,

$$V_{1\widetilde{N}} \sim \sigma^2 \{1 + (m-1)\rho\} \chi^2_{\widetilde{N}} \quad \text{and} \quad V_{2\widetilde{N}} \sim \sigma^2 (1-\rho) \chi^2_{\widetilde{N}(m-1)},$$

and the CDFs of $\widetilde{V}_{1\widetilde{N}}$ and $\widetilde{V}_{2\widetilde{N}}$ are $F_{V_{1\widetilde{n}}}$ and $F_{V_{2\widetilde{n}}}$ respectively.

Second, note that for any sample size n, $\hat{\sigma}_n^4 \left\{1 + (m-1)\hat{\rho}_n^2\right\} = \left\{(m-1)V_{1n}^2 + V_{2n}^2\right\}/m(m-1)n^2$. Therefore,

$$K^* = \frac{(m-1)V_{1k}^2 + V_{2k}^2}{\pi_k(\delta)} \quad \text{and} \quad \widetilde{K} = \left\lfloor \frac{(m-1)\left(V_{1k} + V^*_{1K_p}\right)^2 + \left(V_{2k} + V^*_{2K_p}\right)^2}{\pi_{k+K_p}(\delta)} \right\rfloor + 1, \qquad (4.2)$$

where $\pi_n(\delta) = \left\{m(m-1)n^2\right\}/\beta_n(\delta)$ is a function of n and δ. Now, we present the formula for the distribution function of N in the following theorem.

Theorem 4.1. *Let z^+ denotes $\max\{0, z\}$. The cumulative distribution function of N is given by*

$$F_N(k) = P(N \le k) = \int_0^{\sqrt{k\pi_k(\delta)}} F_{V_{1k}}\left[\sqrt{\frac{k\pi_k(\delta) - y^2}{m-1}}\right] dF_{V_{2k}}(y), \qquad (4.3)$$

and for $n = k+1, k+2, \ldots$, $F_N(n) = P(N \le n)$ equals

$$F_N(k) + \int_0^{u_1} \int_{l_2}^{u_2} \int_0^{u_3} F_{V_{1j}} \left[\sqrt{\frac{n\pi_{k+j}(\delta) - (v_2+y)^2}{m-1}} - v_1 \right] dF_{V_{2j}}(y) dF_{V_{2k}}(v_2) dF_{V_{1k}}(v_1), \quad (4.4)$$

where $u_1 = u_1(n) = \sqrt{(k + \frac{n-k}{p})\pi_k(\delta)(m-1)^{-1}}$, $u_2 = u_2(n, v_1) = \sqrt{(k + \frac{n-k}{p})\pi_k(\delta) - (m-1)v_1^2}$, $l_2 = l_2(v_1) = \sqrt{\{k\pi_k(\delta) - (m-1)v_1^2\}^+}$, $u_3 = u_3(n, v_1, v_2) = \left[\sqrt{n\pi_{k+j}(\delta) - (m-1)v_1^2} - v_2\right]^+$, *and*

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

Proof. Note that

$$\begin{aligned} F_N(k) = P(K^* \le k) &= P\left\{(m-1)V_{1k}^2 + V_{2k}^2 \le k\pi_k(\delta)\right\} \quad \text{using (4.2)} \\ &= \int_0^{\sqrt{k\pi_k(\delta)}} F_{V_{1k}} \left[\sqrt{\frac{k\pi_k(\delta) - y^2}{m-1}} \right] dF_{V_{2k}}(y). \end{aligned}$$

The last equality is obtained by conditioning on V_{2k} and using independence of V_{1k} and V_{2k}. Now, using the definition (4.1) of N, the CDF is $F_N(n) = F_N(k) + F_2 + F_3$, for $n = k+1, k+2, \ldots$, where

$$F_2 = P(k + K_p \le n,\ K^* > k,\ \widetilde{K} \le k + K_p) \quad \text{and} \quad F_3 = P(\widetilde{K} \le n,\ K^* > k,\ \widetilde{K} > k + K_p).$$

Note that $\{k + K_p \le n\} = \{K^* < k + (n-k)/p\}$. Therefore,

$$F_2 = P\left(\widetilde{K} \le k + K_p\,,\ k < K^* < k + (n-k)/p\right) \quad (4.5)$$

We can write the third part F_3 as

$$\begin{aligned} F_3 &= P\left(k + K_p < \widetilde{K} \le n,\ K^* > k,\ k + K_p \le n\right) \\ &= P\left(k + K_p < \widetilde{K} \le n,\ k < K^* < k + (n-k)/p\right) \end{aligned}$$

Therefore, using (4.5), we have $F_2 + F_3 = P\left(\widetilde{K} \le n,\ k < K^* < k + (n-k)/p\right)$. Now, conditioning on

$(V_{1k}, V_{2k}) = (v_1, v_2)$, and using (4.2), we can write

$$F_2 + F_3 = \iint_D P\left\{(m-1)(v_1 + V_{1j}^*)^2 + (v_2 + V_{2j}^*)^2 \leq n\pi_{k+j}(\delta)\right\} dF_{V_{2k}}(v_2) dF_{V_{1k}}(v_1), \tag{4.6}$$

where j is the value of K_p when $(V_{1k}, V_{2k}) = (v_1, v_2)$, i.e., $j = j(v_1, v_2) = \left\lfloor \left(\frac{(m-1)v_1^2 + v_2^2}{\pi_k(\delta)} - k\right) p \right\rfloor + 1.$ We also use independence of (V_{1k}, V_{2k}) and (V_{1j}^*, V_{2j}^*) in (4.6). Now, the region of integration in (4.6) is

$$\begin{aligned} D &= \left\{k \leq \frac{(m-1)v_1^2 + v_2^2}{\pi_k(\delta)} \leq k + (n-k)/p\right\} \\ &= \left\{c_1 \leq (m-1)v_1^2 + v_2^2 \leq c_2\right\} \quad \text{where } c_1 = k\pi_k(\delta) \text{ and } c_2 = (k + (n-k)/p)\pi_k(\delta) \\ &= \left\{0 \leq v_1 \leq \sqrt{\frac{c_2}{m-1}},\ \sqrt{\left\{c_1 - (m-1)v_1^2\right\}^+} \leq v_2 \leq \sqrt{c_2 - (m-1)v_1^2}\right\}. \end{aligned}$$

Figure 1 illustrates the region of integration D. Now, conditioning on $V_{2j}^* = y$ in (4.6) and using independence of V_{2j}^* and V_{1j}^*, we have

$$F_2 + F_3 = \iint_D \int_0^{u_3} F_{V_{1j}}\left[\sqrt{\frac{n\pi_{k+j}(\delta) - (v_2 + y)^2}{m-1}} - v_1\right] dF_{V_{2j}}(y) dF_{V_{2k}}(v_2) dF_{V_{1k}}(v_1),$$

where $u_3 = \left[\sqrt{n\pi_{k+j}(\delta) - (m-1)v_1^2} - v_2\right]^+$ is obtained by choosing nonnegative values of y such that $\sqrt{\left\{n\pi_{k+j}(\delta) - (v_2 + y)^2\right\}/(m-1)} - v_1 \geq 0$. This completes the proof.

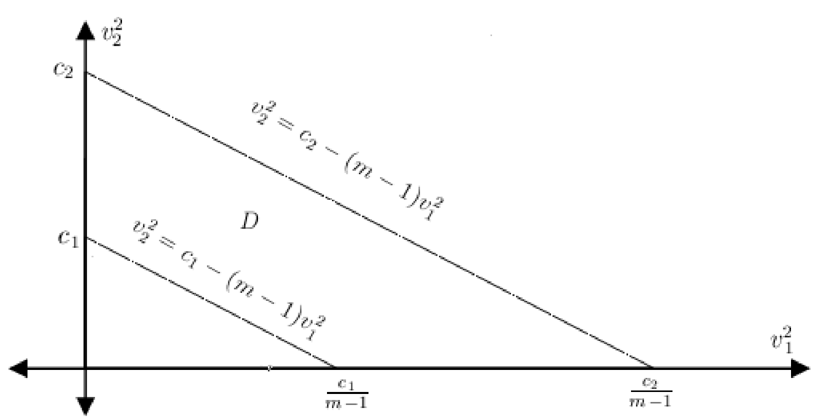

Figure 1. The region of integration D in (4.6)

$\square$

One can easily compute F_N of the above theorem using some numerical integration techniques such as quadrature methods, Monte Carlo method, etc. Using F_N, we can also derive the exact formulas of expectation and variance of N. Since $N \geq k$ almost surely, it is easy to show that $E(N) = k + \sum_{n=k}^{\infty} \bar{F}_N(n)$, where $\bar{F}_N(n) = 1 - F_N(n)$.

To compute the variance of N, we follow the idea of Zacks and Mukhopadhyay (2007) and write

$$\begin{aligned}
E(N^2) &= \sum_{n=k}^{\infty} \left(2 \sum_{i=1}^{n} i - n \right) P(N = n) \\
&= 2 \left[\frac{k(k+1)}{2} + (k+1)P(N \geq k+1) + (k+2)P(N \geq k+2) + \cdots \right] - E(N) \\
&= k^2 + \sum_{n=k}^{\infty} (2n+1)\bar{F}_N(n).
\end{aligned}$$

Therefore,

$$\begin{aligned}
V(N) &= k^2 + \sum_{n=k}^{\infty} (2n+1)\bar{F}_N(n) - \left\{ k + \sum_{n=k}^{\infty} \bar{F}_N(n) \right\}^2 \\
&= \sum_{n=k}^{\infty} (2n - 2k + 1)\bar{F}_N(n) - \left\{ \sum_{n=k}^{\infty} \bar{F}_N(n) \right\}^2 .
\end{aligned}$$

4.2. Numerical Study: Distribution of Stopping Variables

In this numerical study, the following statements are verified.

(1) The *exact* CDF, expected value and variance of N can be computed numerically using Theorem 4.1.

(2) The exact computation of $E(N)$ and $V(N)$ involves numerical approximations. The approximated exact values of $E(N)$ and $V(N)$ match with the simulated values of $E(N)$ and $V(N)$.

(3) On average, the modified three-stage procedure requires less sample size than that of the Stein type two-stage procedure of Haner and Zacks (2013). In fact, for an appropriately chosen value of p, the modified three-stage method yields expected sample size that is close to the oracle sample size of (2.4).

We begin our study with numerical evaluation of the CDF of N as given in Theorem 4.1. In (4.3), $F_N(k)$ is computed using the"*integrate*" function of R which involves an adaptive Gauss–Kronrod quadrature method. For sample size $n \geq k+1$, $F_N(n)$, in (4.4), is obtained by successively computing the three integrals numerically using the "*integrate*" function of R. If the *integrate* function returns an error, we apply Gauss's quadrature method (see Rade and Westergren (1990) and Abramowitz and Stegun (1965)) as an alternative. For all numerical studies in this paper, we set the initial sample size $k = 30$ (as it is selected in Haner and Zacks (2013)), level of confidence $(1-\alpha)100\% = 90\%$, fixed-width $2\delta = 0.2$, and true common variance $\sigma^2 = 1$. Let N_2 be the stopping variable for the two-stage sampling method based on MLE of σ and ρ as given in Haner and Zacks (2013).

Figure 2 and 3 compare the CDF of N_2 with the CDFs of N for $p = 0.3, 0.5$, and 0.7. Both the figures show that N_2 is *stochastically larger* than N to the right of the median of N_2. This indicates that the distribution of N_2 has heavier tail compared to the distribution of N and, therefore, we expect mean and variance of N_2 to be higher than that of N. This is shown in Table 1. For fixed m and ρ, the distribution of N is not too different for $p = 0.3, 0.5$, and 0.7. However, this may not be the case if we choose p close to 0 or 1. In particular, $p = 0$ and $p = 1$ yield two-stage stopping variable N_2 and pure three-stage stopping variable respectively, and both the figures show that the distribution of N_2 is quite different from the distribution of N with $p = 0.3, 0.5$, and 0.7. Moreover, Figure 3 illustrates that N is stochastically bigger for $p = 0.7$ as compared to the case of $p = 0.3$ and 0.5.

Figure 4 compares the distributions of N for different dimensions of the observed vectors such as $m = 2, 3$, and 5 for fixed ρ and p. Higher values of m leads to *stochastically smaller* stopping variable N. Recall that N is determined by stagewise estimation of oracle sample size $n_1^0(\delta, \sigma, \rho)$, where the estimator is a function of V_{1n} and $V_{2n} = \sum_{i=1}^{n} \sum_{j=2}^{m} Y_{ij}^2$. Clearly, for fixed n, as m increases n_1^0 is estimated based on more informative V_{2n}. Therefore, for a higher value of m, stagewise estimation attains certain precision and stops earlier than that of a smaller value of m.

Figure 5 illustrates that if m and p are fixed, stronger correlation between the components of observed vectors $\mathbf{X_1}, \mathbf{X_2}, \ldots$, leads to *stochastically larger* stopping variable N. This is not surprising since stronger correlation between any two components of $\mathbf{X}_i$ implies that individual vectors are less informative than the case of a weaker correlation.

Table 1 compares the expected value and variance of two-stage stopping variable N_2 with the modified three-stage stopping variable N for different values of p, m, and ρ. Here, $E(N)$ and $V(N)$ represent *exact* values whereas $\widehat{E}(N)$ and $\widehat{V}(N)$ represent the *simulated* values of expectation and variance respectively. Note that the simulated values and the exact values of expectation and variance are very similar which validates our numerical approximations of the exact formulas.

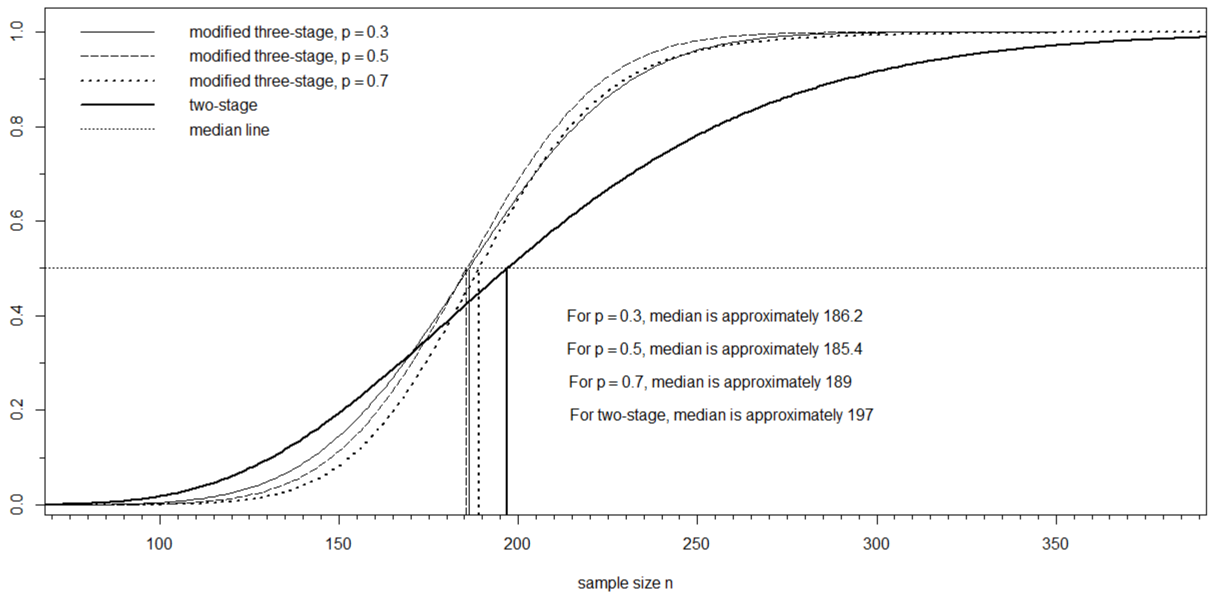

Figure 2. Comparison of the CDF of two-stage stopping variable N_2 with the CDFs of N for different values of p using $k = 30$, $\alpha = 0.1$, $\delta = 0.1$, $\sigma = 1$, $m = 3$, and $\rho = 0.1$

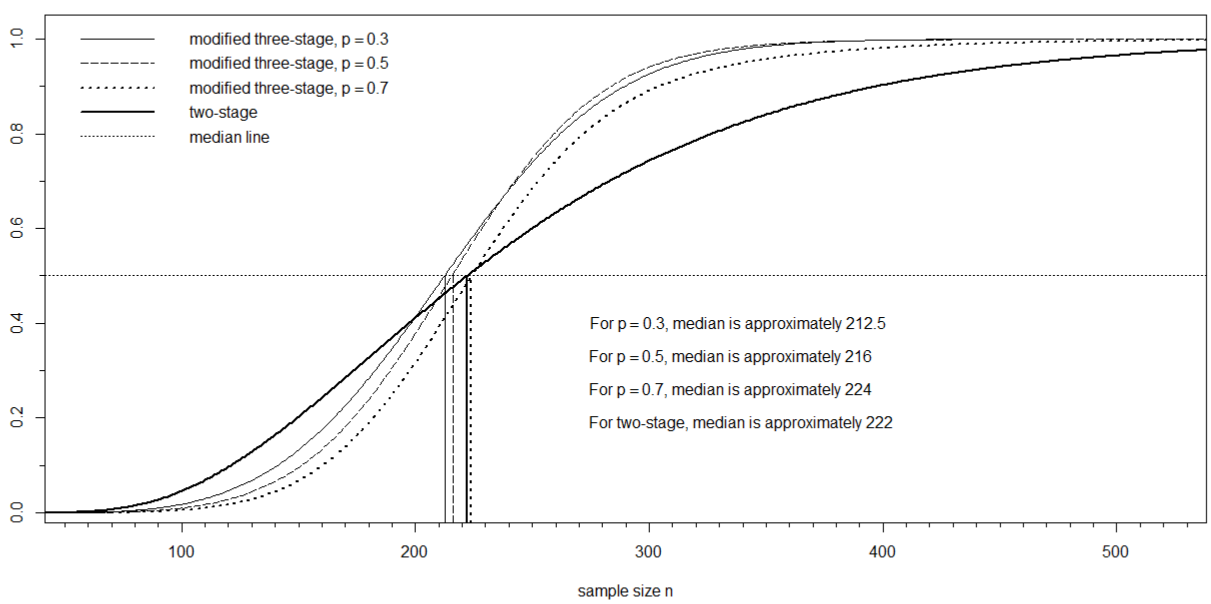

Figure 3. Comparison of the CDF of two-stage stopping variable N_2 with the CDFs of N for different values of p using $k = 30$, $\alpha = 0.1$, $\delta = 0.1$, $\sigma = 1$, $m = 5$, and $\rho = 0.5$

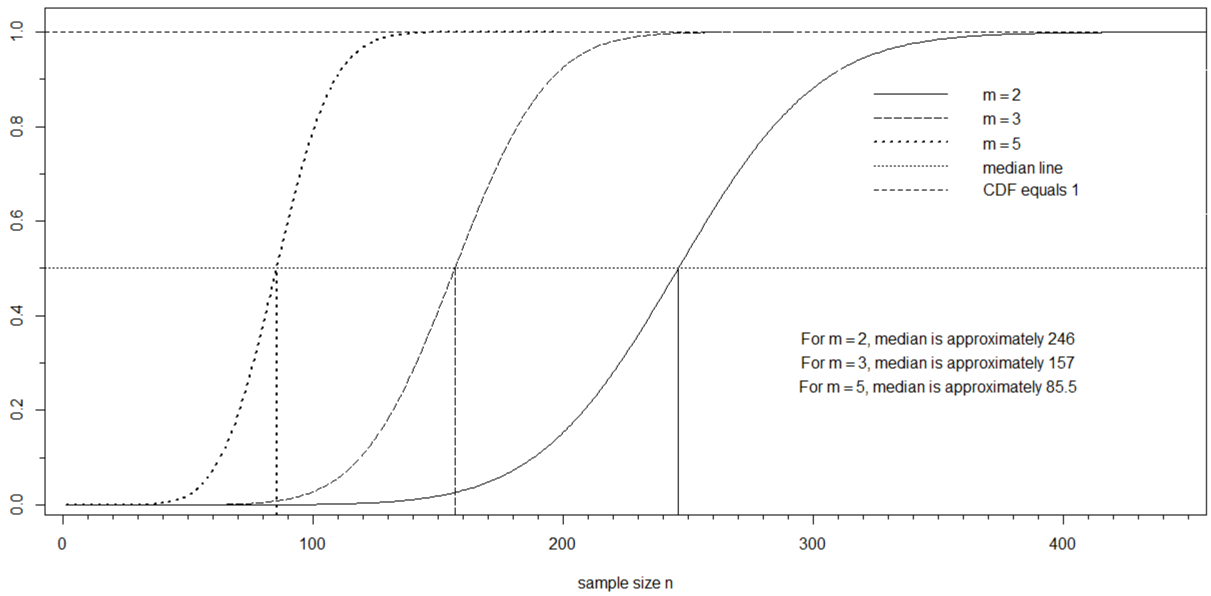

Figure 4. Comparison of the CDFs of N for different values of m using $\sigma = 1, \alpha = 0.1, \delta = 0.1, \rho = 0.1$, and $p = 0.5$

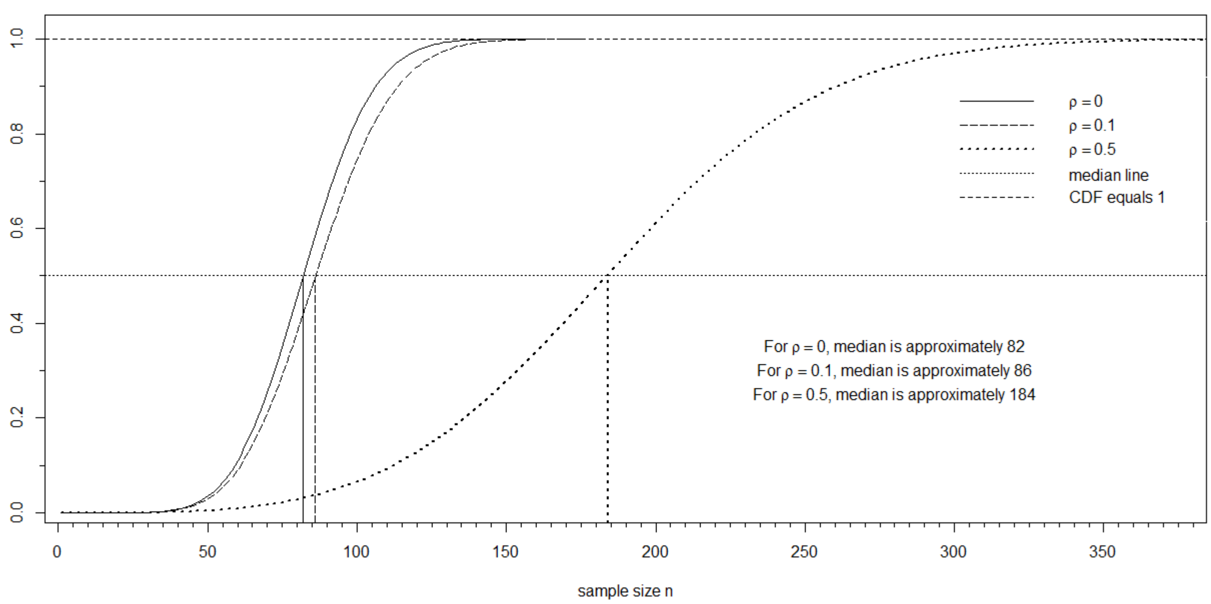

Figure 5. Comparison of the CDFs of N for different values of ρ using $\sigma = 1, \alpha = 0.1, \delta = 0.1, m = 5$, and $p = 0.3$

Table 1. Comparison of expectation and variance of N_2 with the exact values and simulated values of expectation and variance of N for $k = 30$, $\alpha = 0.1$, $\sigma = 1$, and $\delta = 0.1$

p	m	ρ	n^0	$E(N_2)$	$\widehat{E}(N)$	$E(N)$	$V(N_2)$	$\widehat{V}(N)$	$V(N)$
0.3	2	0	271	307.77	272.85	272.50	12982.47	2784.45	2677.53
		0.1	274	310.85	275.32	275.20	13762.71	2918.99	2870.86
		0.5	339	384.59	335.49	335.50	33267.45	6074.35	6083.54
0.3	3	0	181	201.08	183.78	183.44	3653.66	1112.19	1107.19
		0.1	184	205.99	187.33	187.29	4214.73	1232.91	1227.82
		0.5	271	306.71	268.66	268.59	20663.67	4441.72	4483.53
0.3	5	0	109	118.80	111.95	111.96	756.60	339.36	337.82
		0.1	113	124.59	116.44	116.41	1025.78	425.69	420.68
		0.5	217	245.17	214.91	214.90	13497.12	3399.15	3361.60
0.5	2	0	271	307.77	273.15	273.12	12982.47	2055.09	2011.77
		0.1	274	310.85	275.83	275.83	13762.71	2182.54	2117.78
		0.5	339	384.59	340.87	340.82	33267.45	5054.56	4996.06
0.5	3	0	181	201.08	182.68	182.65	3653.66	810.60	811.36
		0.1	184	205.99	186.35	186.34	4214.73	912.86	908.43
		0.5	271	306.71	272.30	272.34	20663.67	3663.59	3619.87
0.5	5	0	109	118.80	110.83	110.64	756.60	260.76	253.56
		0.1	113	124.59	115.28	115.24	1025.78	326.71	325.10
		0.5	217	245.17	217.51	217.55	13497.12	2826.16	3268.98
0.7	2	0	271	307.77	284.86	284.83	12982.47	2914.55	2627.59
		0.1	274	310.85	288.04	288.05	13762.71	3141.04	2912.33
		0.5	339	384.59	362.42	362.62	33267.45	8418.81	7765.40
0.7	3	0	181	201.08	187.06	187.02	3653.66	894.12	836.69
		0.1	184	205.99	191.47	191.47	4214.73	1056.76	1010.24
		0.5	271	306.71	289.05	289.04	20663.67	5645.63	7501.32
0.7	5	0	109	118.80	111.65	111.65	756.60	240.19	241.27
		0.1	113	124.59	116.72	116.73	1025.78	326.37	326.79
		0.5	217	245.17	231.13	231.04	13497.12	4101.37	3785.33

All simulation results are based on 10^6 replications. For all comparisons in Table 1, expectation and variance of N is significantly smaller than that of N_2. Therefore, the modified three-stage sampling improves upon the two-stage sampling in the sense of substantially reducing the expected sampling cost of the experiment.

Significant reduction of variance also establishes superiority of the modified three-stage procedure over the two-stage procedure. This is not surprising since the distributions of N_2 usually have heavier tails than that of N as shown in Figures 2 and 3. Moreover, note that n^0 represents the minimum sample size required to achieve $(1-\alpha)$ coverage probability if all the unknown parameters were known. For an appropriate choice of p, Table 1 shows that the expected sample sizes required by the three-stage method are quite close to the values of minimum sample size n^0.

5. COMPUTATION OF COVERAGE PROBABILITY AND FUNCTIONALS OF $\hat{\sigma}_N^2$

The coverage probability of the interval estimator $\left(\hat{\sigma}_N^2 - \delta, \hat{\sigma}_N^2 + \delta\right)$ is given as $F_{\hat{\sigma}_N^2}\left[\sigma^2 + \delta\right] - F_{\hat{\sigma}_N^2}\left[\sigma^2 - \delta\right]$, where $F_{\hat{\sigma}_N^2}$ is the CDF of $\hat{\sigma}_N^2$. To compute the exact coverage probability, we derive the formula for $F_{\hat{\sigma}_N^2}$ first.

5.1. Computation of The CDF of $\hat{\sigma}_N^2$

For any $x > 0$, the CDF $F_{\hat{\sigma}_N^2}(x) = P(\hat{\sigma}_N^2 \leq x) = P_1 + P_2 + P_3$, where $P_1 = P(\hat{\sigma}_k^2 \leq x,\ K^* \leq k)$,

$$P_2 = P(\hat{\sigma}_{k+K_p}^2 \leq x,\ K^* > k,\ \widetilde{K} \leq k + K_p), \text{ and } P_3 = P(\hat{\sigma}_{\widetilde{K}}^2 \leq x,\ K^* > k,\ \widetilde{K} > k + K_p)$$

are the components of the CDF corresponding to stage I, II, and III respectively. Using (2.1), P_1 can be easily computed as

$$\begin{aligned} P_1 &= P\left\{V_{1k} + V_{2k} \leq mkx\,,\ (m-1)V_{1k}^2 + V_{2k}^2 \leq k\pi_k(\delta)\right\} \\ &= \int_0^{\min\left\{mkx\,,\,\sqrt{k\pi_k(\delta)}\right\}} F_{V_{1k}}\left[\min\left\{mkx - y\,,\ \sqrt{\frac{k\pi_k(\delta) - y^2}{m-1}}\right\}\right] dF_{V_{2k}}(y). \end{aligned} \tag{5.1}$$

The last equality is obtained by conditioning on V_{2k} and using independence of V_{1k} and V_{2k}. The limits of integration in (5.1) is obtained by choosing $y \geq 0$ such that $mkx - y \geq 0$ and $k\pi_k(\delta) - y^2 \geq 0$. The second component P_2 can be expressed using (4.2) as follows

$$
\begin{aligned}
P_2 =& P\left(V_{1k} + V^*_{1K_p} + V_{2k} + V^*_{2K_p} \le m(k+K_p)x,\ K^* > k,\right. \\
& \left.(m-1)(V_{1k} + V^*_{1K_p})^2 + (V_{2k} + V^*_{2K_p})^2 \le (k+K_p)\pi_{k+K_p}(\delta)\right) \\
=& \iint\limits_{R_1} P\left(V^*_{1j} + V^*_{2j} \le m(k+j)x - v_1 - v_2,\ (m-1)(v_1 + V^*_{1j})^2 + (v_2 + V^*_{2j})^2 < (k+j)\pi_{k+j}(\delta)\right) \\
& dF_{V_{2k}}(v_2) dF_{V_{1k}}(v_1).
\end{aligned}
$$

The second equality is due to conditioning on $(V_{1k}, V_{2k}) = (v_1, v_2)$ and independence of (V_{1k}, V_{2k}) and (V^*_{1j}, V^*_{2j}), where the term $j = j(v_1, v_2)$ is defined in Theorem 4.1. The region of integration is

$$
R_1 = \left\{ \frac{(m-1)v_1^2 + v_2^2}{\pi_k(\delta)} > k \right\} = \{v_1 \ge 0,\ v_2 \ge l_2(v_1)\},
$$

where $l_2(v_1)$ is defined in Theorem 4.1. Finally, conditioning on $V^*_{2j} = y$ and using independence of V^*_{1j} and V^*_{2j}, we obtain

$$
\begin{aligned}
P_2 =& \iint\limits_{G_1} \int_0^{u_4} F_{V_{1j}} \left[\min\left\{ m(k+j)x - v_1 - v_2 - y,\ \sqrt{\frac{(k+j)\pi_{k+j}(\delta) - (v_2+y)^2}{m-1}} - v_1 \right\} \right] \\
& dF_{V_{2j}}(y) dF_{V_{2k}}(v_2) dF_{V_{1k}}(v_1), \qquad (5.2)
\end{aligned}
$$

where

$$
u_4 = u_4(v_1, v_2) = \min\left\{ m(k+j)x - v_1 - v_2,\ \sqrt{(k+j)\pi_{k+j}(\delta) - (m-1)v_1^2} - v_2 \right\}
$$

is obtained by considering y such that $\sqrt{\frac{(k+j)\pi_{k+j}(\delta) - (v_2+y)^2}{m-1}} - v_1 \ge 0$ and $m(k+j)x - v_1 - v_2 - y \ge 0$. The region of integration $G_1 = R_1 \cap \{(v_1, v_2) : u_4(v_1, v_2) \ge 0\}$.

Now, we need to compute the third component P_3 of the CDF of $\hat{\sigma}^2_N$. Using (4.2), we can write

$$
\begin{aligned}
P_3 =& P\left(V_{1k} + V^*_{1K_p} + \widetilde{V}_{1\widetilde{N}} + V_{2k} + V^*_{2K_p} + \widetilde{V}_{2\widetilde{N}} \le m\widetilde{K}x,\ K^* > k,\right.\\
&\left.(m-1)(V_{1k} + V^*_{1K_p})^2 + (V_{2k} + V^*_{2K_p})^2 > (k+K_p)\pi_{k+K_p}(\delta)\right)\\
=& \iint\limits_{R_1} \iint\limits_{G_2} P(\widetilde{V}_{1\widetilde{n}} + \widetilde{V}_{2\widetilde{n}} \le m\widetilde{j}x - v_1 - v_1^* - v_2 - v_2^*,\\
&(m-1)(v_1 + v_1^*)^2 + (v_2 + v_2^*)^2 > (k+j)\pi_{k+j}(\delta))dF_{V_{2j}}(v_2^*)dF_{V_{1j}}(v_1^*)dF_{V_{2k}}(v_2)dF_{V_{1k}}(v_1).
\end{aligned}
$$

The last equality is obtained by conditioning on $(V_{1k}, V_{2k}, V^*_{1j}, V^*_{2j})$ and using independence of $(\widetilde{V}_{1\widetilde{n}}, \widetilde{V}_{2\widetilde{n}})$ and $(V_{1k}, V_{2k}, V^*_{1j}, V^*_{2j})$. Given $(V_{1k}, V_{2k}, V^*_{1j}, V^*_{2j}) = (v_1, v_2, v_1^*, v_2^*)$,

$$
\widetilde{K} = \widetilde{j} = \left\lfloor \frac{(m-1)\,(v_1+v_1^*)^2 + (v_2+v_2^*)^2}{\pi_{k+j}(\delta)} \right\rfloor + 1, \quad \text{and} \quad \widetilde{N} = \widetilde{n} = \widetilde{j} - (k+j). \tag{5.3}
$$

For fixed (v_1, v_2), the region of integration

$$
\begin{aligned}
G_2 &= \left\{(v_1^*, v_2^*) : (m-1)(v_1+v_1^*)^2 + (v_2+v_2^*)^2 > (k+j)\pi_{k+j}(\delta)\right\}\\
&= \left\{v_1^* \ge 0,\ v_2^* \ge l_4(v_1, v_2, v_1^*)\right\},
\end{aligned}
$$

where

$$
l_4 = l_4(v_1, v_2, v_1^*) = \left[\sqrt{\{(k+j)\pi_{k+j}(\delta) - (m-1)(v_1+v_1^*)^2\}^+} - v_2\right]^+. \tag{5.4}
$$

Finally, conditioning on $\widetilde{V}_{2\widetilde{n}}$ and using independence of $\widetilde{V}_{1\widetilde{n}}$ and $\widetilde{V}_{2\widetilde{n}}$, we obtain

$$
\begin{aligned}
P_3 =& \iint\limits_{R_1} \iint\limits_{R_2} \int_0^{m\widetilde{j}x - v_1 - v_1^* - v_2 - v_2^*} F_{V_{1\widetilde{n}}}\left[m\widetilde{j}x - v_1 - v_1^* - v_2 - v_2^* - y\right]\\
&dF_{V_{2\widetilde{n}}}(y)dF_{V_{2j}}(v_2^*)dF_{V_{1j}}(v_1^*)dF_{V_{2k}}(v_2)dF_{V_{1k}}(v_1),
\end{aligned} \tag{5.5}
$$

where $R_2 = G_2 \cap \left\{m\widetilde{j}x - v_1 - v_1^* - v_2 - v_2^* \ge 0\right\}$. For fixed (v_1, v_2), the region

$$
\left\{m\widetilde{j}x - v_1 - v_1^* - v_2 - v_2^* \ge 0\right\} = \left\{v_1^* \ge \left[\frac{a}{2(\sqrt{m}-1)} - v_1\right]^+,\ v_2^* \ge \gamma_2 \ \text{ OR } \ v_2^* \le \gamma_1\right\},
$$

where $a = \frac{\pi_{k+j}(\delta)}{mx}$, $\gamma_1 = \gamma_1(v_1, v_2, v_1^*) = \frac{a}{2} - \sqrt{-(m-1)(v_1 + v_1^*)^2 + a(v_1 + v_1^*) + \frac{a^2}{4}} - v_2$, and $\gamma_2 = \gamma_2(v_1, v_2, v_1^*) = \frac{a}{2} + \sqrt{-(m-1)(v_1 + v_1^*)^2 + a(v_1 + v_1^*) + \frac{a^2}{4}} - v_2$. Therefore, the region

$$R_2 = \left\{ v_1^* \geq \left[\frac{a}{2(\sqrt{m} - 1)} - v_1 \right]^+, \ v_2^* \geq \max(l_4, \gamma_2) \right\} \bigcup \left\{ v_1^* \geq \left[\frac{a}{2(\sqrt{m} - 1)} - v_1 \right]^+, \ l_4 \leq v_2^* \leq \gamma_1 \right\}.$$

Remark 5.1. Explicit limits of integration with respect to v_1 and v_2 in (5.2) and (5.5) are very difficult to determine. Therefore, we use the Monte Carlo integration technique to compute P_2 and P_3. First, we simulate a large number of pairs from the distribution of (V_{1k}, V_{2k}). Then, for P_2, we compute the integral with respect to y (using Gauss's quadrature method as in Rade and Westergren (1990)) for each pair of values of (V_{1k}, V_{2k}) that are inside G_1. The average of all these values of the integrals approximates P_2. Similarly, for P_3, one options is to compute the integrals successively with respect to y, v_2^*, and v_1^* for each pair of values of (V_{1k}, V_{2k}) that are inside R_1. Alternatively, we can use Monte Carlo integration to approximate the integrals with respect to v_2^*, v_1^*, v_2, and v_1 and use Gauss's quadrature method to compute the integral with respect to y.

5.2. Computation of Expectation, Variance, and MSE of $\hat{\sigma}_N^2$

Now, we derive the formula for expectation of $\hat{\sigma}_N^2$. Note that $E\left(\hat{\sigma}_N^2\right) = E_1 + E_2 + E_3$, where $E_1 = E\left[\hat{\sigma}_k^2 I\left(K^* \leq k\right)\right]$, $E_2 = E\left[\hat{\sigma}_{k+K_p}^2 I\left(K^* > k, \ \widetilde{K} \leq k + K_p\right)\right]$, and $E_3 = E\left[\hat{\sigma}_{\widetilde{K}}^2 I\left(K^* > k, \ \widetilde{K} > k + K_p\right)\right]$ are the components of the expected value corresponding to stage I, II, and III respectively. Since the first component E_1 is a function of (V_{1k}, V_{2k}) only (recall (2.1) and (4.2)), it can be easily computed as

$$E_1 = \int_0^{\sqrt{\frac{k\pi_k(\delta)}{m-1}}} \int_0^{\sqrt{k\pi_k(\delta) - (m-1)v_1^2}} \left[\frac{v_1 + v_2}{mk} \right] dF_{V_{2k}}(v_2) dF_{V_{1k}}(v_1). \tag{5.6}$$

The limits of integration are obtained by solving $\left\{ \frac{(m-1)v_1^2 + v_2^2}{\pi_k(\delta)} \leq k \right\}$.

To compute the second component E_2, first consider expectation over (V_{1k}, V_{2k}), i.e.,

$$E_2 = \int_0^\infty \int_{l_2(v_1)}^\infty E\left[\frac{v_1 + V_{1j}^* + v_2 + V_{2j}^*}{m(k+j)} I\left((m-1)(v_1 + V_{1j}^*)^2 + (v_2 + V_{2j}^*)^2 \leq (k+j)\pi_{k+j}(\delta)\right) \right]$$
$$dF_{V_{2k}}(v_2) dF_{V_{1k}}(v_1),$$

where $l_2(v_1)$ and $j = j(v_1, v_2)$ are defined in Theorem 4.1. Next, we integrate over the distribution of

(V_{1j}^*, V_{2j}^*) to obtain

$$E_2 = \int_0^\infty \int_{l_2}^\infty \int_0^{u_5} \int_0^{u_6} \left[\frac{v_1 + v_1^* + v_2 + v_2^*}{m(k+j)} \right] dF_{V_{2j}}(v_2^*) dF_{V_{1j}}(v_1^*) dF_{V_{2k}}(v_2) dF_{V_{1k}}(v_1), \tag{5.7}$$

where l_2 is defined in Theorem 4.1, $u_5 = u_5(v_1, v_2) = \left[\sqrt{\frac{(k+j)\pi_{k+j}(\delta)}{m-1}} - v_1 \right]^+$, and

$$u_6 = u_6(v_1, v_2, v_1^*) = \left[\sqrt{(k+j)\pi_{k+j}(\delta) - (m-1)(v_1 + v_1^*)^2} - v_2 \right]^+. \tag{5.8}$$

For fixed (v_1, v_2), we solve the inequality $\left\{ (m-1)(v_1 + v_1^*)^2 + (v_2 + v_2^*)^2 \le (k+j)\pi_{k+j}(\delta) \right\}$ and obtain the limits of integration for v_1^* and v_2^*.

Following similar steps as in the computation of E_2, we take expectation over $(V_{1k}, V_{2k}, V_{1K_p}^*, V_{2K_p}^*)$ to express the third component as

$$\begin{aligned} E_3 =& E\left[\frac{V_{1k} + V_{1K_p}^* + \widetilde{V}_{1\widetilde{N}} + V_{2k} + V_{2K_p}^* + \widetilde{V}_{2\widetilde{N}}}{m\widetilde{K}} \; I(K^* > k\,,\, \widetilde{K} > k + K_p) \right] \\ =& \int_0^\infty \int_{l_2}^\infty \int_0^\infty \int_{l_4}^\infty E\left[\frac{v_1 + v_1^* + \widetilde{V}_{1\widetilde{n}} + v_2 + v_2^* + \widetilde{V}_{2\widetilde{n}}}{m\widetilde{j}} \right] dF_{V_{2j}}(v_2^*) dF_{V_{1j}}(v_1^*) dF_{V_{2k}}(v_2) dF_{V_{1k}}(v_1), \end{aligned}$$

where l_4 and l_2 are defined in (5.4) and Theorem 4.1 respectively. $\widetilde{j}$ and $\widetilde{n}$ are defined in (5.3). The limits of integration for (v_1, v_2) and (v_1^*, v_2^*) are obtained from

$$\left\{ \frac{(m-1)v_1^2 + v_2^2}{\pi_k(\delta)} > k \right\} \quad \text{and} \quad \left\{ (m-1)(v_1 + v_1^*)^2 + (v_2 + v_2^*)^2 > (k+j)\pi_{k+j}(\delta) \right\}$$

respectively. Since $V_{1\widetilde{n}} \sim \sigma^2 (1 + (m-1)\rho) \chi^2_{\widetilde{n}}$ and $V_{2\widetilde{n}} \sim \sigma^2 (1-\rho) \chi^2_{\widetilde{n}(m-1)}$, $E\left[\widetilde{V}_{1\widetilde{n}} + \widetilde{V}_{2\widetilde{n}} \right] = m\widetilde{n}\sigma^2$. Therefore,

$$E_3 = \int_0^\infty \int_{l_2}^\infty \int_0^\infty \int_{l_4}^\infty E\left[\frac{v_1 + v_1^* + v_2 + v_2^* + m\widetilde{n}\sigma^2}{m\widetilde{j}} \right] dF_{V_{2j}}(v_2^*) dF_{V_{1j}}(v_1^*) dF_{V_{2k}}(v_2) dF_{V_{1k}}(v_1). \tag{5.9}$$

Next, we provide the formula for the variance of $\hat{\sigma}_N^2$. Since the derivation of $V\left(\hat{\sigma}_N^2\right)$ and $E\left(\hat{\sigma}_N^2\right)$ involve similar steps, we only present the formula for $V\left(\hat{\sigma}_N^2\right)$. Let $\mu = E\left(\hat{\sigma}_N^2\right)$. Then $V\left(\hat{\sigma}_N^2\right) = V_1 + V_2 + V_3$, where

$$V_1 = \int_0^{\sqrt{\frac{k\pi_k(\delta)}{m-1}}} \int_0^{\sqrt{k\pi_k(\delta)-(m-1)v_1^2}} \left[\frac{v_1+v_2}{mk} - \mu\right]^2 dF_{V_{2k}}(v_2)dF_{V_{1k}}(v_1), \tag{5.10}$$

$$V_2 = \int_0^{\infty}\int_{l_2}^{\infty}\int_0^{u_5}\int_0^{u_6} \left[\frac{v_1+v_1^*+v_2+v_2^*}{m(k+j)} - \mu\right]^2 dF_{V_{2j}}(v_2^*)dF_{V_{1j}}(v_1^*)dF_{V_{2k}}(v_2)dF_{V_{1k}}(v_1), \tag{5.11}$$

and

$$V_3 = \int_0^{\infty}\int_{l_2}^{\infty}\int_0^{\infty}\int_{l_4}^{\infty} E\left[\frac{2\widetilde{n}\sigma^4\left\{m+2(m-1)\rho^2\right\}+\left\{v_1+v_1^*+v_1+v_1^*-m\widetilde{j}\mu+m\widetilde{n}\sigma^2\right\}^2}{m^2\widetilde{j}^2}\right]$$
$$dF_{V_{2j}}(v_2^*)dF_{V_{1j}}(v_1^*)dF_{V_{2k}}(v_2)dF_{V_{1k}}(v_1). \tag{5.12}$$

Finally, the mean square error of $\hat{\sigma}_N^2$ is given by $V\left(\hat{\sigma}_N^2\right)+\left\{E\left(\hat{\sigma}_N^2\right)-\sigma^2\right\}^2$.

5.3. Numerical Results

In Table 2, we compare the coverage probabilities obtained from the two-stage procedure of Haner and Zacks (2013) with the modified three-stage procedure for different values of p, m, and ρ. Here, CP2 represents coverage probability from the two-stage procedure, and CP and $\widehat{\text{CP}}$ represent *exact* values and *simulated* values of coverage probability obtained from the modified three-stage procedure respectively. All simulated values are based on 10^6 replications. To compute (5.2), (5.5), (5.7), (5.9), (5.11), and (5.12), large number of observations are drawn from the distribution of (V_{1k}, V_{2k}), and Monte Carlo integration method is used for the integrals with respect to v_1 and v_2. The integrals with respect to v_1^* and v_2^* in (5.5), (5.7), (5.9), (5.11), and (5.12) are also approximated using the Monte Carlo numerical integration. The number of replications used in all the Monte Carlo integrations is 20,000.

Table 2 illustrates that coverage probabilities achieved by the two-stage procedure and the modified procedure are nearly the same, and they are close to the desired level 0.9. Moreover, $\widehat{\text{CP}}$ and CP are very close to each other for all choices of p, m, and ρ. This validates numerical approximations of the exact coverage probabilities.

Table 3 shows that the expectation, variance, and MSE of estimated σ^2 at stopping for the two-stage

Table 2. Comparison of exact coverage probabilities (CP2) for the two-stage procedure, simulated ($\widehat{\text{CP}}$) and exact coverage probabilities (CP) for the modified three-stage procedure when $k = 30$, $\alpha = 0.1$, $\sigma = 1$, and $\delta = 0.1$ for different values of p, m, and ρ.

			$p = 0.3$		$p = 0.5$		$p = 0.7$	
m	ρ	CP2	$\widehat{\text{CP}}$	CP	$\widehat{\text{CP}}$	CP	$\widehat{\text{CP}}$	CP
2	0	0.8959	0.8907	0.8901	0.8923	0.8904	0.8983	0.9011
	0.1	0.8953	0.8901	0.8895	0.8918	0.8907	0.8988	0.8993
	0.5	0.8842	0.8834	0.8837	0.8886	0.8896	0.8968	0.8951
3	0	0.8994	0.8932	0.8940	0.8936	0.8937	0.8972	0.8992
	0.1	0.8989	0.8931	0.8936	0.8934	0.8930	0.8975	0.8977
	0.5	0.8838	0.8818	0.8819	0.8875	0.8836	0.8944	0.8926
5	0	0.9024	0.8968	0.8968	0.8958	0.8940	0.8975	0.8991
	0.1	0.9020	0.8961	0.8942	0.8958	0.8952	0.8975	0.8970
	0.5	0.8817	0.8784	0.8810	0.8826	0.8817	0.8913	0.8929

sampling are quite similar to that of the modified three-stage sampling for different choices of p, m, and ρ. For both the two-stage and the modified three-stage procedure, we can conclude that the estimators of σ^2 at stopping are *nearly unbiased*. Moreover, the variance and the MSE of the estimators of σ^2 at stopping (for both two-stage and modified three-stage) are nearly 0 which indicates that the estimators are quite precise.

APPENDIX

Table 4 presents a simulation study comparing the three-stage procedure and our modified three-stage procedure. Suppose N_3 denotes the stopping variable for the three-stage procedure. $\widehat{E}(N_3)$ and $\widehat{V}(N_3)$ denote the estimated value of the expectation and variance of N_3 respectively. $\widehat{\text{CP}_3}$ and $\widehat{E(\hat{\sigma}^2_{N_3})}$ denote the estimated coverage probability and expected value of the estimator of σ^2 at N_3. All the simulations in Table 4 are based on 100,000 replications. As we mentioned before, our modified three-stage sampling is only slightly different from the three-stage procedure. The difference is that we draw approximately $\lfloor k(1-p) \rfloor$ more samples in the second stage of modified three-stage procedure. Due to this extra sampling, the expected sample size $E(N_3)$ is always bigger than $E(N)$. However, this difference is not significant. Table 4 illustrates that the modified three-stage procedure slightly increases the coverage probability compared to the three-stage method in most cases. For the case of $p = 0.3$ and $m = 5$, the three-stage procedure stops early

Table 3. Comparison of expectation, variance, and mean square error of estimated σ^2 at stopping for the two-stage procedure and the modified three-stage procedure when $k = 30$, $\alpha = 0.1$, $\sigma = 1$, and $\delta = 0.1$

p	m	ρ	$E(\hat{\sigma}^2_{N_2})$	$\widehat{E(\hat{\sigma}^2_N)}$	$E(\hat{\sigma}^2_N)$	$V(\hat{\sigma}^2_{N_2})$	$\widehat{V(\hat{\sigma}^2_N)}$	$V(\hat{\sigma}^2_N)$	$MSE(\hat{\sigma}^2_{N_2})$	$\widehat{MSE(\hat{\sigma}^2_N)}$	$MSE(\hat{\sigma}^2_N)$
0.3	2	0	0.99257	0.99222	0.99218	0.003861	0.003948	0.003951	0.003917	0.004008	0.004012
		0.1	0.99249	0.99212	0.99230	0.003878	0.003963	0.003960	0.003934	0.004025	0.004020
		0.5	0.99126	0.99085	0.99087	0.004209	0.004151	0.004150	0.004285	0.004235	0.004233
0.3	3	0	0.99276	0.99243	0.99249	0.003748	0.003877	0.003878	0.003800	0.003934	0.003934
		0.1	0.99263	0.99229	0.99242	0.003760	0.003892	0.003898	0.003814	0.003951	0.003955
		0.5	0.99058	0.98998	0.98965	0.004201	0.004199	0.004206	0.004290	0.004300	0.004313
0.3	5	0	0.99293	0.99271	0.99277	0.003660	0.003788	0.003763	0.003710	0.003841	0.003816
		0.1	0.99263	0.99242	0.99246	0.003669	0.003801	0.003779	0.003723	0.003859	0.003836
		0.5	0.98939	0.98866	0.98818	0.004260	0.004305	0.004331	0.004373	0.004434	0.004471
0.5	2	0	0.99257	0.99233	0.99074	0.003861	0.003905	0.000257	0.003917	0.003963	0.000343
		0.1	0.99249	0.99222	0.99737	0.003878	0.003911	0.003959	0.003934	0.003972	0.003966
		0.5	0.99126	0.99106	0.99499	0.004209	0.004020	0.004379	0.004285	0.004100	0.004404
0.5	3	0	0.99276	0.99245	0.98870	0.003748	0.003865	0.007105	0.003800	0.003922	0.007233
		0.1	0.99263	0.99211	0.99284	0.003760	0.003873	0.002873	0.003814	0.003935	0.002924
		0.5	0.99058	0.99019	0.99090	0.004201	0.004063	0.002778	0.004290	0.004159	0.002861
0.5	5	0	0.99293	0.99268	0.99175	0.003660	0.003799	0.003808	0.003710	0.003853	0.003876
		0.1	0.99263	0.99232	0.99182	0.003669	0.003808	0.003803	0.003723	0.003867	0.003870
		0.5	0.98939	0.98882	0.98888	0.004260	0.004200	0.004167	0.004373	0.004325	0.004290
0.7	2	0	0.99257	0.99266	0.99469	0.003861	0.003758	0.003730	0.003917	0.003812	0.003758
		0.1	0.99249	0.99259	0.99308	0.003878	0.003752	0.003819	0.003934	0.003807	0.003867
		0.5	0.99126	0.99155	0.99042	0.004209	0.003803	0.003671	0.004285	0.003874	0.003763
0.7	3	0	0.99276	0.99268	0.99312	0.003748	0.003773	0.003704	0.003800	0.003827	0.003751
		0.1	0.99263	0.99252	0.99339	0.003760	0.003771	0.003797	0.003814	0.003827	0.003841
		0.5	0.99058	0.99074	0.99036	0.004201	0.003866	0.003812	0.004290	0.003952	0.003904
0.7	5	0	0.99293	0.99272	0.99181	0.003660	0.003762	0.003737	0.003710	0.003815	0.003804
		0.1	0.99263	0.99237	0.99206	0.003669	0.003763	0.003789	0.003723	0.003821	0.003852
		0.5	0.98939	0.98940	0.99054	0.004260	0.003976	0.003878	0.004373	0.004088	0.003968

and yields significantly low coverage probability whereas our modified method still provides the desired coverage probability. The modified procedure reduces the variability in the random sample size in most cases shown in Table 4. Moreover, the estimated bias of $\hat{\sigma}^2$ is slightly smaller for the modified three-stage sampling compared to the three-stage procedure.

The R Code for Exact Computation of The CDF of N, Functionals of N, The Coverage Probability, and The Functionals of $\hat{\sigma}^2_N$

```
sig<-1; rho<-0             # parameters
p<-0.3; m<- 5; k<- 30      # p=proportion, m=dimension, k=pilot sample size
alpha<- 0.1; del<- 0.1     # 1-alpha = confidence level, 2*del=fixed-width
F.V1 <- function(x, r) {                        # CDF of V1
  pchisq(x/((sig^2)*( 1+(m-1)*rho)),r) }
F.V2 <- function(x, r) {                        # CDF of V2
  pchisq(x/((sig^2)*( 1-rho)), r*(m-1) )}
f.V1 <- function(x, r) {                        # PDF of V1
  (1/( (sig^2)*( 1 + (m-1)*rho) ))*dchisq(   x/((sig^2)*( 1 + (m-1)*rho )),  r
      )}
f.V2 <- function(x, r) {                         # PDF of V2
  (1/( (sig^2)*(1-rho) ))*dchisq(   x/((sig^2)*( 1-rho)), r*(m-1) ) }
pi <- function(del, r){ (m*(m-1)*(r^2) )/( (2*qt(alpha/2, r)*qt(alpha/2, r))/(
    m*del^2) ) }
kp<-function(v1k,v2k){ floor( ( ( ((m-1)*v1k^2 +v2k^2)/ pi(del,k)  ) -k)*p ) +
     1 } # Additional sample in the 2nd stage

##==== CDF of N as in Theorem 4.1. using "integrate" function of R =====

F1<-integrate( function(y){ # Computing the integral in (4.3)
F.V1( sqrt( (k*pi(del,k)-y^2)/(m-1) ) ,k)*f.V2(y,k)},0,sqrt(k*pi(del,k))  )$
    value
F.N<-function(n){    # Computing the integral in (4.4)
c1<-k*pi(del,k); c2<- (k+ (n-k)/p)*pi(del,k)
F3.int<-function(y,v2k,v1k){
```

Table 4. Comparison of expectation and variance of sample sizes, coverage probabilities, and expected value of estimated σ^2 at stopping for the three-stage procedure with that of our modified three-stage procedure for $k = 30$, $\alpha = 0.1$, $\sigma = 1$, and $\delta = 0.1$

p	m	ρ	$\widehat{E}(N_3)$	$\widehat{E}(N)$	$\widehat{V}(N_3)$	$\widehat{V}(N)$	$\widehat{\text{CP}_3}$	$\widehat{\text{CP}}$	$\widehat{E(\hat{\sigma}^2_{N_3})}$	$\widehat{E(\hat{\sigma}^2_{N})}$
0.3	2	0	270.17	272.66	3716.1	2779.93	0.8862	0.8918	0.99049	0.99206
		0.1	272.55	275.11	3881.76	2922.09	0.8848	0.8912	0.99038	0.99204
		0.5	330.64	335.24	7804.18	6077.08	0.8761	0.8834	0.98823	0.99075
0.3	3	0	181.29	183.73	1804.16	1106.59	0.8824	0.8947	0.98819	0.99239
		0.1	184.76	187.32	1965.39	1227.89	0.8809	0.8931	0.98809	0.99222
		0.5	262.94	268.58	6107.13	4433.74	0.8686	0.8834	0.98546	0.98992
0.3	5	0	97.70	111.94	1891.49	339.55	0.7570	0.8968	0.96950	0.99255
		0.1	103.98	116.42	1949.60	424.65	0.7632	0.8966	0.97121	0.99228
		0.5	207.18	214.80	5179.63	3381.98	0.8505	0.8785	0.97969	0.98864
0.5	2	0	271.77	273.17	2274.65	2054.24	0.8896	0.8921	0.99197	0.99231
		0.1	274.34	275.91	2406.31	2173.33	0.8905	0.8915	0.99179	0.99213
		0.5	337.67	340.46	5396.04	5010.41	0.8862	0.8890	0.99026	0.99083
0.5	3	0	181.87	182.61	955.87	813.63	0.8918	0.8942	0.99209	0.99226
		0.1	185.46	186.35	1058.48	912.86	0.8927	0.8934	0.99193	0.99210
		0.5	269.58	272.17	4007.03	3639.23	0.8832	0.8875	0.98962	0.99019
0.5	5	0	110.63	110.78	342.01	259.69	0.8941	0.8963	0.99192	0.99256
		0.1	114.85	115.21	424.97	325.22	0.8938	0.8963	0.99139	0.99209
		0.5	214.52	217.60	3191.90	2836.62	0.8773	0.8828	0.98765	0.98878
0.7	2	0	282.59	284.93	2829.83	2904.65	0.8981	0.9006	0.99235	0.99261
		0.1	285.88	288.16	3035.61	3146.50	0.8992	0.9001	0.99246	0.99243
		0.5	359.61	362.29	8189.91	8321.97	0.8949	0.8958	0.99108	0.99130
0.7	3	0	185.43	187.03	878.02	891.05	0.8958	0.8984	0.99248	0.99256
		0.1	189.61	191.51	1018.86	1048.40	0.8957	0.8993	0.99212	0.99242
		0.5	286.21	289.10	5544.93	5583.15	0.8924	0.8950	0.99015	0.99058
0.7	5	0	110.68	111.69	249.85	239.44	0.8963	0.8987	0.99236	0.99274
		0.1	115.53	116.74	328.37	323.62	0.8973	0.8977	0.99207	0.99229
		0.5	227.87	231.13	4019.3	4114.08	0.8878	0.8905	0.98876	0.98931

```
j<-  floor( ( ( ((m-1)*v1k^2 +v2k^2)/ pi(del,k)  ) -k)*p ) + 1;
F.V1( sqrt( ( n*pi(del,k+j)-(v2k+y)^2 )/(m-1) ) -v1k , j )*f.V2(y ,j)*f.V1(v1k
   ,k)*f.V2(v2k,k)  }
F3<-integrate(
    function(v1k) {
      f2<-function(v1k)
   { integrate( function(v2k){
     sapply(v2k, function(v2k){integrate( function(y) F3.int(y,v2k,v1k),
     0, sqrt( n*pi(del,k+kp(v1k,v2k)) - (m-1)*v1k^2)-v2k  )$value})  },
     sqrt(max(0,  c1 - (m-1)*v1k^2)), sqrt( c2-(m-1)*v1k^2)  )$value }
     sapply(v1k, f2)
    }, 0, sqrt(c2/(m-1)), rel.tol=0.001 )$value
return(F1+F3)  }
M<-160   # M is chosen such that FN[M] is very close to 1; M is different for
   different cases
FN<-rep(0, M+1);     FN[1]<-F1
system.time( for(i in 1: M)
{ FN[i+1]<-F.N(k+i)
  if(FN[i]-FN[i+1]>0.0001) stop("CDF decreasing")
  if(FN[i+1]>1) stop("CDF greater than one")
} )
EN<- k+ sum(1-FN)
PN<-rep(0,length(FN))
PN[1]<-F1
for(i in 1: (length(FN)-1) ){ PN[i+1]<- FN[i+1]-FN[i]  } # The PMF of N
E.Nsq<- sum( (c(k: (k+length(FN)-1) )^2)*PN )
VarN<- E.Nsq - EN^2
EN; VarN      # Expectation and variance of stopping variable N

##======= CDF of MLE of sigma^2 at stopping N ==========
# Z and W are nodes and weights for Gauss's quadrature method given in Rade
   and Westergren (1990)
Z<-c(-0.1488743390,-0.4333953941,-0.6794095683,-0.8650633667,-0.9739065285,
```

```
     0.9739065285,0.8650633667,0.6794095683,0.4333953941,0.1488743390)
W<-c(0.2955242247,0.2692667193,0.2190863625,0.1494513492,0.0666713443,
     0.0666713443,0.1494513492,0.2190863625,0.2692667193,0.2955242247)
F.sigsq.N<-function(x)  {
#----- Computing P1 as in the equation (5.1) --------
int1<-function(y){F.V1( min( m*k*x -y, sqrt( (k*pi(del,k) - y^2)/(m-1) ) ) , k
   )*f.V2(y,k) }
P1<- integrate(int1, 0, min(m*k*x, sqrt(k*pi(del,k)) ) )$value
#----- Computing P2 as in (5.2) using Monte Carlo and Gauss's quadrature -----
M1<-20000; t.P2<-rep(0,M1); t.P3<-rep(0,M1); set.seed(123) # M1= no. of
   replications for Monte Carlo
for(i in 1:M1){
v1<- (sig^2)*(1+(m-1)*rho)*rchisq(1,k)
v2<- (sig^2)*(1-rho)*rchisq(1, k*(m-1))
j<-floor( ( ( ((m-1)*v1^2 +v2^2)/ pi(del,k)  ) -k)*p ) + 1
f2<-function(y){
F.V1( min( m*(k+j)*x -v1-v2-y, sqrt( ( (k+j)*pi(del,k+j) - (v2+y)^2 )/(m-1) )-
   v1 ) , j)*f.V2(y, j)    }
u2<- min( m*(k+j)*x -v1-v2 , sqrt((k+j)*pi(del,k+j) -(m-1)*v1^2 ) -v2 )
s2<-seq(0,u2,0.5); fs2<-sapply(s2,f2)
eps<-0.0000001
if(sum(fs2>eps)>0 ){ l<-s2[min(which(fs2>eps))];   u<- s2[max(which(fs2>eps))]
     }
if(sum(fs2>eps)==0 ) { l<-0; u<-0}
t.P2[i]<-  0.5*(u-l)*sum( sapply(0.5*(l+u) + 0.5*(u-l)*Z, f2)*W )
#---- Computing P3 as in (5.5) using Monte Carlo and Gauss's quadrature ----
v1s<-(sig^2)*(1+(m-1)*rho)*rchisq(1,j)     # v1s= 1 draw from distn of V1j.
   star
v2s<-(sig^2)*(1-rho)*rchisq(1, j*(m-1))    # v2s= 1 draw from distn of V2j.
   star
j.til<- floor( ((m-1)*(v1+v1s)^2 + (v2+v2s)^2 )/pi(del, k+j) ) +1
lv<- sqrt( max(0, (k+j)*pi(del, k+j) -(m-1)*(v1+v1s)^2) )-v2
f3<-function(y){
```

```
if( v2s >= lv ) { return( F.V1(m*j.til*x -v1-v2-v1s-v2s-y, j.til-k-j)*f.V2(y,
   j.til-k-j) ) } else return(0)  }
vf3<- Vectorize(f3,"y")
u3<- m*j.til*x-v1-v2-v1s-v2s
if( u3>0 )
{     s3<-seq(0, u3 ,0.5); fs3<-sapply(s3,f3)
      eps<-0.0000001
      if(sum(fs3>eps)>0 ){ l<-s3[min(which(fs3>eps))];   u<- s3[max(which(fs3>
         eps))]  }
      if(sum(fs3>eps)==0 ) { l<-0; u<-0}
      t.P3[i]<-  0.5*(u-l)*sum( sapply(0.5*(l+u) + 0.5*(u-l)*Z, f3)*W )
} else t.P3[i]<- 0     }
P1+mean(t.P2)+mean(t.P3)   }
system.time(p2<-F.sigsq.N(sig^2 + del))
system.time(p1<-F.sigsq.N(sig^2 - del))
(CP<- p2-p1)  # The exact coverage probability

##======== Computing E( MLE of sigma^2 at N) ===============

f1<-function(v1,v2){ ((v1+v2)/(m*k))*f.V1(v1,k)*f.V2(v2,k) }
E1<- integrate(function(v1) { # Computing E1 as in (5.6)
      sapply(v1, function(v1){
      integrate(function(v2) f1(v1,v2), 0, sqrt(k*pi(del,k) - (m-1)*v1^2) )$
         value }  )
      }, 0, sqrt((k*pi(del,k))/(m-1))  )$value
M2<-20000; t.E2<-rep(0,M2); t.E3<-rep(0,M2); set.seed(123) # M2= No. of
   replications for Monte Carlo
system.time( for(i in 1:M2){
v1<- (sig^2)*(1+(m-1)*rho)*rchisq(1,k)
v2<- (sig^2)*(1-rho)*rchisq(1, k*(m-1))
j<-floor( ( ( ((m-1)*v1^2 +v2^2)/ pi(del,k)  ) -k)*p ) + 1
#------- Computing E2 and E3 as in (5.7) and (5.9) using Monte Carlo and
   integrate function of R ------
```

```
f2<-function(v1s,v2s){ ( (v1+v1s+v2+v2s)/(m*(k+j)) )*f.V1(v1s,j)*f.V2(v2s,j) }
t.E2[i]<- integrate(function(v1s) {
   sapply(v1s, function(v1s){
   integrate(function(v2s) f2(v1s,v2s), 0, sqrt( (k+j)*pi(del,k+j) - (m-1)*(v1
       +v1s)^2 ) -v2 )$value }  )
  }, 0, sqrt( ( (k+j)*pi(del,k+j) )/(m-1) )-v1  )$value
v1s<- (sig^2)*(1+(m-1)*rho)*rchisq(1,j)   # v1s= 1 draw from distn of V1j.star
v2s<- (sig^2)*(1-rho)*rchisq(1, j*(m-1))  # v2s= 1 draw from distn of V2j.star
j.til<- floor( ((m-1)*(v1+v1s)^2 + (v2+v2s)^2 )/(pi(del,k+j)) ) +1
  if( v2s> sqrt( max( 0, (k+j)*pi(del,k+j)-(m-1)*(v1+v1s)^2 ) )-v2 )
t.E3[i]<- (v1+v1s+v2+v2s+m*(j.til-k-j)*sig^2 )/(m*j.til) else t.E3[i]<-0
  } )
(E.sigsq<- E1+  mean(t.E2 + t.E3)) # Expectation of MLE of sigma^2 at stopping

##======= Computing the Var( MLE of sigma^2 at N) and MSE =========

mu<-E.sigsq; t.V2<-rep(0,M2); t.V3<-rep(0,M2);
Vf1<-function(v1,v2){ f.V1(v1,k)*f.V2(v2,k)*((v1+v2)/(m*k) -mu )^2 }
var1<- integrate(function(v1) {# Computing V1 as in (5.10)
   sapply(v1, function(v1){
   integrate(function(v2) Vf1(v1,v2), 0, sqrt(k*pi(del,k) - (m-1)*v1^2) )$
       value }  )
}, 0, sqrt((k*pi(del,k))/(m-1))  )$value
system.time( for(i in 1:M2){
v1<- (sig^2)*(1+(m-1)*rho)*rchisq(1,k)
v2<- (sig^2)*(1-rho)*rchisq(1, k*(m-1))
j<-floor( ( ( ((m-1)*v1^2 +v2^2)/ pi(del,k)  ) -k)*p ) + 1
#--- Computing V2 and V3 as in (5.11) and (5.12) using Monte Carlo and Gauss's
     quadrature ---
Vf2<-function(v1s,v2s){ ( (  (v1+v1s+v2+v2s)/(m*(k+j)) -mu  )^2 )*f.V1(v1s,j)*
   f.V2(v2s,j) }
Vg2<-function(v1s){l<- 0; u<-sqrt( (k+j)*pi(del,k+j) - (m-1)*(v1+v1s)^2 ) -v2
```

```
0.5*(u-l)*sum( sapply(0.5*(u+l) + 0.5*(u-l)*Z, function(v2s) Vf2(v1s,v2s)  )*W
     ) }
l<-0; u<- sqrt( ( (k+j)*pi(del,k+j) )/(m-1) )-v1;
t.V2[i]<-  0.5*(u-l)*sum( sapply(0.5*(u+l) + 0.5*(u-l)*Z, Vg2 )*W )
v1s<- (sig^2)*(1+(m-1)*rho)*rchisq(1,j) # v1s= 1 draw from distn of V1j.star
v2s<- (sig^2)*(1-rho)*rchisq(1, j*(m-1))# v2s= 1 draw from distn of V2j.star
j.til<- floor( ((m-1)*(v1+v1s)^2 + (v2+v2s)^2 )/(pi(del,k+j)) ) +1
if( v2s > sqrt( max( 0, (k+j)*pi(del,k+j)-(m-1)*(v1+v1s)^2 ) )-v2  )
t.V3[i]<- ( 2*(j.til-k-j)*(sig^4)*( (1+(m-1)*rho)^2 + (m-1)*(1-rho)^2 ) +
      (  v1+v1s+v2+v2s-m*j.til*mu +m*(j.til-k-j)*sig^2   )^2 )/(m*j.til)^2
         else t.V3[i]<- 0 } )
var.sigsq <- var1+ mean(t.V2 + t.V3)  # variance of MLE of sigma^2 at stopping
MSE.sigsq<- var.sigsq + (mu-sig^2)^2  # MSE of MLE of sigma^2 at stopping
E.sigsq; var.sigsq; MSE.sigsq;        # expectation, variance, and MSE
```


ACKNOWLEDGEMENTS

The author is eternally grateful to Professor Shelemyahu Zacks for suggesting this problem and for his constant support and guidance throughout this work. The author is thankful to the Editor, Professor Nitis Mukhopadhyay, for his careful review and suggestions that helped improve this manuscript.

REFERENCES

Abramowitz, M. and Stegun, I. A. (1965). *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables*. 1^{st} ed. New York: Dover Publications.

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

Dantzig, G. B. (1940). On The Non-existance of Tests of Student's Hypothesis Having Power Functions Independent of σ, *Annals of Mathematical Statistics* 11: 186–192.

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. (1990). *Handbook of Sequential Analysis*, New York: Dekker.

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

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

Hall, P. (1981). Asymptotic Theory of Triple Sampling for Estimation of a Mean, *Annals of Statistics*, 9: 1229–1238.

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.

Haner, D. M. and Zacks, S. (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.

Lehmann, E. L. (1997). *Testing Statistical Hypotheses*. 2^{nd} ed. New York: Springer.

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

Mukhopadhyay, N. (1976). Fixed-Width Confidence Intervals for The Mean Using a Three-Stage Procedure. Unpublished report.

Mukhopadhyay, N. (1990). Some Properties of a Three-Stage Procedure with Applications in Sequential Analysis, *Sankhya, Series A* 52: 218–231.

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

Mukhopadhyay, N. and Solanky, T. K. S. (1994). *Multistage Selection and Ranking Procedures: Second-Order Asymptotics*. New York: Marcel Dekker.

R-codes. The exact formulas are computed in this paper using the R code given in `https://sites.google.com/site/shyamalkd/contact/R%20code%20for%20modified%20three-stage%20paper.rtf`.

Rade, L. and Westergren, B. (1990). *Beta Mathematics Handbook: Concepts, Theorems, Methods, Algorithms, Formulas, Graphs, Tables*. 2^{nd} ed. Boca Raton: CRC Press.

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.

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

Zacks, S., and Mukhopadhyay, N. (2007). Distributions of Sequential and Two-Stage Stopping Times for Fixed-Width Confidence Intervals in Bernoulli Trials: Application in Reliability, *Sequential Analysis*, 26: 425–441.

Zacks, S., and Ramig, P. (1987). Confidence Intervals for The Common Variance of Equicorrelated Normal Random Variables, *Contributions to The Theory and Applications of Statistics. A Volume in Honor of Herbert Soloman*. New York: Academic Press, Inc.