Article

Three-Stage Estimation of the Mean and Variance of the Normal Distribution with Application to an Inverse Coefficient of Variation with Computer Simulation

Ali Yousef [1,*] **and Hosny Hamdy** [2]

1   Department of Mathematics, Kuwait College of Science and Technology, Kuwait City 27235, Kuwait
2   Faculty of Management Sciences, October University for Modern Sciences and Arts, 6th October City 12566, Egypt; gfarm84hamdy@gmail.com
*   Correspondence: A.yousef@kcst.edu.kw

check for updates

Abstract: This paper considers sequentially two main problems. First, we estimate both the mean and the variance of the normal distribution under a unified one decision framework using Hall's three-stage procedure. We consider a minimum risk point estimation problem for the variance considering a squared-error loss function with linear sampling cost. Then we construct a confidence interval for the mean with a preassigned width and coverage probability. Second, as an application, we develop Fortran codes that tackle both the point estimation and confidence interval problems for the inverse coefficient of variation using a Monte Carlo simulation. The simulation results show negative regret in the estimation of the inverse coefficient of variation, which indicates that the three-stage procedure provides better estimation than the optimal.

Keywords: asymptotic regret; loss function; normal distribution; three-stage sampling procedure

1. Introduction

Let $\{X_i,\ i \geq 1\}$ be a sequence of independent and identically distributed IID random variables from a normal distribution with mean $\mu \in \Re$ and variance $\sigma^2 \in \Re^+$, both μ and σ^2 are unknown. Assume further that a random sample of size $n(\geq 2)$ from the normal distribution becomes available then we propose to estimate μ and σ^2 by the corresponding sample measures $\overline{X}_n$ and S_n^2, respectively. It is a common practice, over the last decays, to treat each problem separately, where we consider one decision framework for each inference problem of the mean or the variance.

The objective in this paper is to combine the inference of both problems under one decision framework in order to achieve maximal use of the available sample information to handle these problems simultaneously. Given pre-defined α, $0 < \alpha < 1$ and $d(> 0)$, where $(1 - \alpha)$ is the confidence coefficient and $2d$ is the fixed-width of the interval, we want to construct a fixed-width $(= 2d)$ confidence interval for the mean μ whose confidence coefficient is at least the nominal value $100(1 - \alpha)\%$, where at the same time, we will be able to use the same available data to estimate the population variance σ^2 under squared-error loss function with linear sampling cost. Hence, we combine both optimal sample sizes in one decision rule to propose the three-stage sampling decision framework.

Therefore, the optimal sample size required to construct a fixed-width confidence interval for μ whose coverage probability is least the nominal value $100(1 - \alpha)\%$ must satisfy the following:

$$n_{conf}^* = (a/d)^2 \sigma^2 \tag{1}$$

where a is the upper $(\alpha/2)\%$ critical point of the standard normal distribution $N(0, 1)$. For more details about Equation (1), see Mukhopadhyay and de Silva ([1]; chapter 6, p. 97).

1.1. Minimum Risk Estimation

In the literature of sequential point estimation problems, one may consider several types of loss functions such as the squared-error loss function, the absolute-error loss function, the linex-loss function and others. It was shown that the commonly used one is the squared-error loss function due to its simplicity in mathematical computations; see, for example, Degroot [2]. Therefore, we write the loss incurred in estimating σ^2 by the corresponding sample measure S_n^2 as

$$L_n(A) = A\left(S_n^2 - \sigma^2\right)^2 + cn \tag{2}$$

where $A > 0$ is a known constant and c is the known cost per unit sample observation. We will elaborate on the determination of A in the following lines. Now, the risk corresponding to Equation (2) is

$$R_n(A) = 2A(n-1)^{-1}\sigma^4 + cn \approx 2A\sigma^4/n + cn, \tag{3}$$

Thus, the minimum value of n that minimizes the risk in Equation (3) is

$$n \geq n_{\text{point}}^* = \sqrt{2A/c}\,\sigma^2 \tag{4}$$

moreover, the associated minimum risk is

$$R_{n^*}(A) = 2cn^* \tag{5}$$

The value of n^* in Equation (4) is called the *optimal* sample size required to generate a point estimate for σ^2 under Equation (2) while Equation (5) is the minimum risk obtained if σ^2 is known.

1.2. A Unified One Decision Framework

If we want to combine both the confidence interval estimation and the point estimation in one decision framework, we have to have the constant $A = (1/2)\left(a^4/d^4\right)c$ to perform both confidence and point estimation in one decision rule. Careful investigation of the constant $A = \left(n^*/2\sigma^4\right)(cn^*)$ provided the statistical interpretation, that is cn^* is the cost of optimal sampling while $\left(n^*/2\sigma^4\right)$ represents the optimal information; in other words, it is the amount of information required to explore a unit of variance in order to achieve minimum risk. Thus A is the *cost of perfect information,* and it is contrary to what has been said in the literature—that it is the cost of estimation.

Therefore, we proceed to use the following optimal sample size, to perform the required inference,

$$n \geq n^* = \left(a^2/d^2\right)\sigma^2 = \xi\sigma^2, \xi = a^2/d^2 \tag{6}$$

Since σ^2 in Equation (6) is unknown, then no fixed sample size procedure can estimate the mean μ, independent of σ^2; see Dantzig [3]. Therefore, we resort to a triple sampling sequential procedure to achieve the previously required goals. Henceforth, we continue to use the asymptotic sample size defined in Equation (6) to propose the following triple sampling procedure to estimate the unknown population mean μ and the unknown population variance σ^2 via estimation of n^*.

2. Three-Stage Estimation of the Mean and Variance

In his seminal work, Hall [4] introduced the idea of sampling in three stages to tackle several problems in sequential estimation. He combined the asymptotic characteristics of one-by-one purely sequential sampling procedures of Anscombe [5], Robbins [6], and Chow and Robbins [7] and the operational saving made possible by Stein [8], and Cox [9] group sampling.

From 1965 until the early 1980s, the research in sequential estimation was mainly devoted to two types of sequential sampling procedures—the two-stage procedure, which satisfies the operational savings, and the one-by-one purely sequential procedure that satisfies the asymptotic efficiency. The objective was to use these methods under non-normal distributions. For brevity, see Mukhopadhyay [10], Mukhopadhyay and Hilton [11], Mukhopadhyay and Darmanto [12], Mukhopadhyay and Hamdy [13], Ghosh and Mukhopadhyay [14], Mukhopadhyay and Ekwo [15], Sinha and Mukhopadhyay [16], Zacks [17], and Khan [18]. For a complete list of references, see Ghosh, Mukhopadhyay, and Sen [19].

In the early 1980s, Hall [4,20] considered the normal distribution with an unknown finite mean and an unknown finite variance. His objective was to construct a confidence interval for the mean with a pre-assigned fixed-width and coverage probability. We will describe Hall's three-stage procedure in Section 2.1.

Since the publication of Hall's paper, research in multistage sampling has extended Halls results in several directions. Some have utilized the triple sampling technique to generate inference for other distributions, others have tried to improve the quality of inference such as protecting the inference against type II error probability, studying the characteristic operating curve, or/and discussing the sensitivity of triple sampling when the underlying distribution departs away from normality. For more details see Mukhopadhyay [21–23], Mukhopadhyay et al. [24], Mukhopadhyay and Mauromoustakos [25], Hamdy and Palotta [26], Hamdy et al. [27], Hamdy [28], Hamdy et al. [29], Lohr [30], Mukhopadhyay and Padmanabhan [31], Takada [32], Hamdy et al. [33], Hamdy [34], Al-Mahmeed and Hamdy [35], AlMahmeed et al. [36], Costanzo et al. [37], Yousef et. al. [38], Yousef [39], Hamdy et al. [40] and Yousef [41]. Liu [42] used Hall's results to tackle hypothesis-testing problems for the mean of the normal distribution while Son et al. [43] used the three-stage procedure to tackle the problem of testing hypotheses concerning shifts in the population normal mean with controlled Type II error probability.

2.1. Three-Stage Sampling Procedure

As the name suggests, an inference in triple sampling is performed in three consecutive stages—the *pilot phase*, the *main study phase*, and the *fine-tuning phase*.

The Pilot Phase: In the pilot study phase, a random sample of size $m(\geq 2)$ from the population say, $(X_1, \ldots, X_m)$ to initiate sample measures, $\overline{X}_m$ for the population mean μ and S_m for the population standard deviation σ, where $\overline{X}_m = m^{-1} \sum_{i=1}^{m} X_i$ and $S_m^2 = (m-1)^{-1} \sum_{i=1}^{m} (X_i - \overline{X}_m)^2$.

The main Study Phase: In the main study phase, we only estimate a portion $\gamma \in (0, 1)$ of n^* to avoid possible oversampling. In literature, γ is known as the design factor. Let $[x]$ be the largest integer $\leq x$ and ξ as defined before we have

$$N_1 = \max\left\{m, \left[\gamma \xi S_m^2\right] + 1\right\} \tag{7}$$

If $m \geq N_1$ then we stop at this stage, otherwise we continue to sample an extra sample of size $N_1 - m$, say $X_{m+1}, X_{m+2}, \ldots, X_{N_1}$, then we update the sampling measures to $\overline{X}_{N_1}$ and S_{N_1} for the population's unknown parameters μ and σ, respectively. Hence, we proceed to define the fine-tuning phase.

The Fine-Tuning Phase: In the fine-tuning phase, the decision to stop or continue sampling is taken according to the following stopping rule

$$N = \max\left\{N_1, \left[\xi S_{N_1}^2\right] + 1\right\} \tag{8}$$

If $N_1 \geq N$ then sampling is terminated at this stage, or else we continue to sample an additional sample of size $N - N_1$, say $X_{N_1+1}, X_{N_1+2}, \ldots, X_N$. Hence, we augment the previously collected N_1 samples with the new $N - N_1$ to update the sample estimates to $\overline{X}_N$ and S_N for the unknown parameters μ and σ. Upon terminating the sampling process, we propose to estimate the unknown population

mean μ by the corresponding triple sampling confidence interval $I_N = \left(\overline{X}_N - d, \ \overline{X}_N + d\right)$ and the unknown population variance σ^2 by the corresponding triple sampling point estimate S_N^2.

The asymptotic results in this paper are developed under the Assumption (A) set forward by Hall [20] to develop a theory for the triple sampling procedure. That is,

Assumption (A) Let $\xi(> 0)$ such that $\xi(m) \to \infty$, $\limsup(m/\xi(m)) < \gamma\sigma^2$, and $\xi(m) = O\!\left(m^k\right)$, $k > 1$.

Preliminaries: Recall the sample variance $S_n^2 = (n-1)^{-1}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2$ for all $n \geq 2$, and consider the following Helmert's transformation to the original normal random variables $X_1, \ldots, X_n$, to write S_n^2 as an average of IID random variables for all $n \geq 2$. Now, let $Z_i = (X_i - \mu)/\sigma$ for $i = 1, 2, \ldots, n$, and write $W_i = (i(i+1))^{-1/2}\left\{\sum_{j=1}^{i} Z_j - i^2 Z_{i+1}\right\}$ for $i = 1, 2, \ldots, n-1$ and $W_n = n^{-1}\sum_{j=1}^{n} Z_j$. They $W_i's$ are IID $N(0, 1)$ for all $i = 1, 2, \ldots, n$. Let $V_i = \sigma^2 W_i^2$, then the random variables $V_1, V_2, \ldots, V_n$ are IID random variables each distributed as $\sigma^2\chi^2(1)$. Which means $\sum_{j=2}^{n} V_j \sim \sigma^2\chi^2(n-1)$. From Lemma 2 of Robbins [6], it follows that S_n^2 and $\overline{V}_n = (n-1)^{-1}\sum_{i=2}^{n} V_i$ are identically distributed for all $n \geq 2$. That is, $S_n^2 \overset{D}{=} \overline{V}_n$, for all $n \geq 2$.

We continue to use the representation of $\overline{V}_n$ instead of S_n^2 for all $n \geq 2$ to develop the asymptotic theory for both the main study phase and the fine-tuning phase.

2.2. The Asymptotic Characteristics of the Main Study Phase

Under Assumption (A), we have

As $\xi \to \infty$, $P\!\left(N = [\xi S_{N_1}^2] + 1\right) = P\!\left(N = [\xi \overline{V}_{N_1}] + 1\right) \to 1$ a.s and as $m \to \infty$, $N_1/\gamma n^* \to 1$, in probability likewise $m/\gamma n^* \to 1$ in probability. While, from the Anscombe [5] Central Limit Theorem, we have as $\xi \to \infty$, $\sqrt{N_1}\left(\overline{X}_{N_1} - \mu\right) \to N\!\left(0, \sigma^2\right)$ and $\sqrt{N_1}\left(S_{N_1}^2 - \sigma^2\right) \to N\!\left(0, 2\sigma^4\right)$ in distribution.

From Theorem 1 of Yousef et al. [38] as $\xi \to \infty$ we have

$$
\begin{aligned}
&\text{(i)} \quad E\!\left(\overline{X}_{N_1}\right) = \mu + o\!\left(\xi^{-1}\right) \\
&\text{(ii)} \quad E\left(\overline{X}_{N_1}^2\right) = \mu^2 + \sigma^2(\gamma n^*)^{-1} + o\!\left(\xi^{-2}\right) \\
&\text{(iii)} \quad Var\!\left(\overline{X}_{N_1}\right) = \sigma^2(\gamma n^*)^{-1} + o\!\left(\xi^{-2}\right)
\end{aligned}
\tag{9}
$$

Theorem 1. *Under Assumption (A) and using Equation (7), we can show for any real k, as $\xi \to \infty$*

$$
E\!\left(S_{N_1}^{2k}\right) = \sigma^{2k} + k(k-3)\sigma^{2k}(\gamma n^*)^{-1} + o\!\left(\xi^{-1}\right)
$$

Proof. Since $S_{N_1}^2$ and $\overline{V}_{N_1}$ are identically distributed, we write

$$
E\!\left(S_{N_1}^{2k}\right) = E\!\left(\overline{V}_{N_1}^{k}\right) = E\!\left((N_1 - 1)^{-1}\sum_{j=1}^{N_1-1} V_j\right)^{k}
$$

Conditioning on the $\sigma - field$ generated by $V_i(i = 1, 2, 3, \ldots, m-1)$ we have

$$
E\!\left(\overline{V}_{N_1}^{k}\right) = E\!\left((N_1 - 1)^{-k} E\!\left(\sum_{j=1}^{m-1} V_j + \sum_{j=m}^{N_1-1} V_j\right)^{k} | V_1, V_2, \ldots, V_{m-1}\right)
$$

By using binomial expansion, it follows

$$E\left(\overline{V}_{N_1}^k\right) = E\left((N_1-1)^{-k}\sum_{j=0}^{\infty}\lambda(k,\,j)\left(\sum_{j=1}^{m-1}V_j\right)^{k-j}E\left(\sum_{j=m}^{N_1-1}V_j\right)^j \Big| V_1,\,V_2,\ldots,V_{m-1}\right)$$

where $\lambda(k,j) = \dfrac{\prod\limits_{t=1}^{j}(k-t+1)}{j!}$ for $j = 1,\,2,\ldots$ and $\lambda(k,0) = 1$. Conditioning on the $\sigma-$

field generated by $V_i(i=1,\,2,\,3,\ldots,m-1)$ that is $\left(\sum\limits_{j=m}^{N_1-1}V_j/\sigma^2 | V_1,\,V_2,\ldots,V_{m-1}\right) \sim \chi^2(N_1-m)$.

It follows $E\left(\sum\limits_{j=m}^{N_1-1}V_j/\sigma^2 | V_1,\,V_2,\ldots,V_{m-1}\right)^j = 2^j\dfrac{\Gamma(j+(N_1-m)/2)}{\Gamma((N_1-m)/2)}$ where $\Gamma(x) = \int\limits_0^{\infty}t^{x-1}e^{-t}\,dt$ and hence

$$E\left(\sum_{j=m}^{N_1-1}V_j | V_1,\,V_2,\ldots,V_{m-1}\right)^j = (N_1-m)^j\sigma^{2j}\left(1+O\left(N_1^{-1}\right)\right). \quad \square$$

Further simplifications similar to those given in Hamdy [28], we get

$$E\left(S_{N_1}^{2k}\right) = \sigma^{2k}E\left(1+\frac{1}{N_1-1}\sum_{i=1}^{m-1}Y_i\right)^k + o\left(\xi^{-1}\right)$$

where $Y_i = \left(V_i-\sigma^2\right)/\sigma^2$ are IID with $E(Y_i) = 0$ and $Var(Y_i) = 2$.

By applying the first two terms of the infinite binomial series and taking the expectation, we get

$$E\left(S_{N_1}^{2k}\right) = \sigma^{2k} + \sigma^{2k}k\,E\left(\frac{1}{N_1-1}\sum_{i=1}^{m-1}Y_i\right) + \tfrac{1}{2}\sigma^{2k}k(k-1)E\left(\frac{1}{N_1-1}\sum_{i=1}^{m-1}Y_i\right)^2 + E(R(Y))$$

$$= \sigma^{2k} + I + II \tag{10}$$

$$E(R(Y)) = M\ E\left(\frac{1}{N_1-1}\sum_{i=1}^{m-1}Y_i\right)^3,\text{ where } M \text{ is a generic constant. Since } m-1\le N_1-1 \text{ we have}$$

$$E(R(Y)) = M\ E\left(\frac{1}{m-1}\sum_{i=1}^{m-1}Y_i\right)^3 = M(m-1)^3E\left(\overline{V}_m-\sigma^2\right)^3/(m-1)^3 = M\,E\left(\overline{V}_m-\sigma^2\right)^3 = 0.$$

Consider $I = \sigma^{2k}kE\left(\dfrac{1}{N_1-1}\sum\limits_{i=1}^{m-1}Y_i\right)$ and expand $(N_1-1)^{-1}$ around γn^*, and then take the expectation.

$$E\left\{(N_1-1)^{-1}\sum_{i=1}^{m-1}Y_i\right\} = 2(\gamma n^*)^{-1}E\left(\sum_{i=1}^{m-1}Y_i\right) - (\gamma n^*)^{-2}E\left(N_1\sum_{i=1}^{m-1}Y_i\right) + (1/2)E\left(\rho^{-3}\sum_{i=1}^{m-1}Y_i(N_1-\gamma n^*)^2\right),$$

where ρ is a random variable between N_1 and γn^*. It is not hard to show that $E\left\{\sum\limits_{i=1}^{m-1}Y_i(N_1-\gamma n^*)^2\rho^{-3}\right\} = o\left(\xi^{-1}\right)$, we have omitted the proof for brevity.

Substituting for $Y_i = \left(V_i-\sigma^2\right)/\sigma^2$, we have

$$I = \sigma^{2k}k\sigma(\gamma n^*)^{-1}(m-1)E\left(\frac{\overline{V}_m}{\sigma^2}-1\right) - \sigma^{2k}k\frac{\sigma}{2}(\gamma n^*)^{-2}(m-1)\gamma\xi\left(\frac{E\left(\overline{V}_m^2\right)}{\sigma^2}-E\left(\overline{V}_m\right)\right) + o\left(\xi^{-1}\right)$$

It follows that

$$I = -\frac{\sigma}{\gamma n^*}\sigma^{2k}k + o\left(\xi^{-1}\right) \tag{11}$$

Likewise, we recall the second term and expand $(N_1 - 1)^{-2}$ around γn^*, we get

$$II = \frac{1}{2}\sigma^{2k}k(k-1)E\left(\frac{1}{N_1 - 1}\sum_{i=1}^{m-1}Y_i\right)^2 = (\gamma n^*)^{-1}\sigma^{2k}k(k-1) \tag{12}$$

Substituting Equations (11) and (12) into Equation (10), we get the result. The proof is complete. As a particular case of Theorem 1, for $k = 1/2, \ 1, \ 2$ and $k = 3$ we have as $\xi \to \infty$,

$$
\begin{array}{ll}
\text{(i)} & E\left(S_{N_1}\right) = \sigma - \frac{5\sigma}{4\gamma n^*} + o\left(\xi^{-1}\right) \\[4pt]
\text{(ii)} & E\left(S_{N_1}^2\right) = \sigma^2 - \frac{2\sigma^2}{\gamma n^*} + o\left(\xi^{-1}\right) \\[4pt]
\text{(iii)} & E\left(S_{N_1}^4\right) = \sigma^4 - \frac{2\sigma^4}{\gamma n^*} + o\left(\xi^{-1}\right) \\[4pt]
\text{(iv)} & E\left(S_{N_1}^6\right) = \sigma^6 + o\left(\xi^{-1}\right)
\end{array}
\tag{13}
$$

while from the Equation (13) and the results of (ii) and (iii) we obtain

$$Var\left(S_{N_1}^2\right) = 2\sigma^4(\gamma n^*)^{-1} + o\left(\xi^{-1}\right) \tag{14}$$

The following Theorem 2 gives the second-order asymptotic expansion of the moments of a real-valued continuously differentiable function of $S_{N_1}^2$.

Theorem 2. *Under Assumption (A) and let $g(>0)$ be a real-valued continuously differentiable function in a neighborhood around σ^2 such that $\sup\limits_{n>m}|g'''(n)| = O(|g'''(n^*)|)$, then*

$$E\left\{g\left(S_{N_1}^2\right)\right\} = g\left(\sigma^2\right) - 2\sigma^2(\gamma n^*)^{-1}\left\{g'\left(\sigma^2\right) - (1/2)\sigma^2 g''\left(\sigma^2\right)\right\} + o\left(\xi^{-1}\right)$$

Proof. Taylor expansion of $g\left(S_{N_1}^2\right)$ around σ^2 provides,

$$g\left(S_{N_1}^2\right) = g\left(\sigma^2\right) + \left(S_{N_1}^2 - \sigma^2\right)g'\left(\sigma^2\right) + (1/2)\left(S_{N_1}^2 - \sigma^2\right)^2 g''\left(\sigma^2\right) + (1/6)\left(S_{N_1}^2 - \sigma^2\right)^3 g'''(\eta),$$

where η is a random variable between $S_{N_1}^2$ and σ^2. Now, taking the expectation all through we have,

$$E\left(g\left(S_{N_1}^2\right)\right) = g\left(\sigma^2\right) + E\left(S_{N_1}^2 - \sigma^2\right)g'\left(\sigma^2\right) + (1/2)E\left(S_{N_1}^2 - \sigma^2\right)^2 g''\left(\sigma^2\right) + (1/6)E\left(\left(S_{N_1}^2 - \sigma^2\right)^3 g'''(\eta)\right),$$

From Equation (13), parts (ii) and (14), we have

$$E\left(g\left(S_{N_1}^2\right)\right) = g\left(\sigma^2\right) - 2\sigma^2(\gamma n^*)^{-1}\left\{g'\left(\sigma^2\right) - \left(\sigma^2/2\right)g''\left(\sigma^2\right)\right\} + (1/6)E\left(\left(S_{N_1}^2 - \sigma^2\right)^3 g'''(\eta)\right),$$

However, $(1/6)E\left(\left(S_{N_1}^2 - \sigma^2\right)^3 g'''(\eta)\right) \leq (1/6)E|S_{N_1}^2 - \sigma^2|^3 \sup\limits_{n\geq m}|g'''(n)| = O\left(|\xi^{-1}g'''(\eta)|\right)$ from Equation (13), part (iv), and the assumption that $g'''(\cdot)$ is a bounded function. The proof is complete. $\square$

Corollary 1. *Under Assumption (A) and let $g(>0)$, be a real-valued continuously differentiable function in a neighborhood around σ such that $\sup\limits_{n>m}|g'''(n)| = O(|g'''(n^*)|)$ then*

$$E\left\{g\left(S_{N_1}\right)\right\} = g(\sigma) + \sigma(4\gamma n^*)^{-1}\left\{\sigma g''(\sigma) - 5g'(\sigma)\right\} + o\left(\xi^{-1}\right)$$

Proof. First, by using Taylor series expansion of the function $g(\cdot)$ around σ, we get

$$g\left(S_{N_1}\right) = g(\sigma) - g'(\sigma)\left(S_{N_1} - \sigma\right) + \frac{1}{2}g''(\sigma)\left(S_{N_1} - \sigma\right)^2 + \frac{1}{6}g'''(\eta)\left(S_{N_1} - \sigma\right)^3$$

By taking the expectation all through we have

$$E\left(g\left(S_{N_1}\right)\right) = g(\sigma) - g'(\sigma)E\left(S_{N_1} - \sigma\right) + \frac{1}{2}g''(\sigma)E\left(S_{N_1} - \sigma\right)^2 + \frac{1}{6}E\left(g'''(\eta)\left(S_{N_1} - \sigma\right)^3\right)$$

by using Equation (13), parts (i), (ii) and (iii) and the fact, that $g'''(\cdot)$ is bounded. The proof is complete. As an especial case of Corollary 1, take $f(t) = t^{-1}$, and $f(t) = t^{-2}$ we obtain,

$$\begin{aligned}
&\text{(i)} \quad E\left(S_{N_1}^{-1}\right) = \sigma^{-1} + 7\sigma^{-1}(4\gamma n^*)^{-1} \\
&\text{(ii)} \quad E\left(S_{N_1}^{-2}\right) = \sigma^{-2} + 4\sigma^{-2}(\gamma n^*)^{-1} + o\left(\xi^{-1}\right) \\
&\text{(iii)} \quad \text{var}\left(S_{N_1}^{-1}\right) = \sigma^{-2}(2\gamma n^*)^{-1} + o\left(\xi^{-1}\right)
\end{aligned} \tag{15}$$

This completes our first assertion regarding the asymptotic characteristics of the main-study phase. In the following section, we find the asymptotic characteristics of the final random sample size. □

2.3. The Asymptotic Characteristics of the Fine-Tuning Phase

Asymptotic characteristics of the variable N are given in the following Theorem.

Theorem 3. *Under Assumption (A) and using Equation (8), let $h(> 0)$ be a real-valued continuously differentiable function in a neighborhood around n^* such that $\sup\limits_{n>m}|h'''(n)| = O|h'''(n^*)|$. Then as $\xi \to \infty$*

$$E\{h(N)\} = h(n^*) + \left(\frac{1}{2} - 2\gamma^{-1}\right)h'(n^*) + \gamma^{-1}n^*h''(n^*) + O\left(\xi^2 h'''(\xi)\right)$$

Proof. We write $N = \left[\xi S_{N_1}^2\right] + 1$, except possibly on a set $\phi = (N_1 < m) \cup \left(\xi\overline{V}_{N_1} < \gamma\xi\overline{V}_m + 1\right)$ of measure zero. Therefore, for real r, we have

$$\begin{aligned}
E(N^r) &= E\left(\left[\xi S_{N_1}^2\right] + 1\right)^r + \int_{\phi} N^r \, dp \\
&= E\left(\left(\left[\xi S_{N_1}^2\right] + 1\right) + \beta_{N_1}\right)^r + o\left(\xi^{r-1}\right)
\end{aligned}$$

provided that the *r*th moment exists, and $\beta_{N_1} = 1 - \left\{\left(\xi S_{N_1}^2\right) - \left[\xi S_{N_1}^2\right]\right\}$, where $[x]$ as defined before. From Hall [4], as $\xi \to \infty$, β_{N_1} is an asymptotically uniform distribution.

Now, for $r = 1$, we have,

$$\begin{aligned}
E(N) &= E\left(\left(\left[\xi S_{N_1}^2\right] + 1\right)\right) + E\left(\beta_{N_1}\right) + o(1) = \xi\left(\sigma^2 - \frac{2\sigma^2}{\gamma n^*} + o\left(\xi^{-1}\right) + \frac{1}{2} + o(1)\right) = \xi\sigma^2 - 2\gamma^{-1} + 1/2 + o(1) \\
&= n^* - 2\gamma^{-1} + 1/2 + o(1)
\end{aligned} \tag{16}$$

Likewise, for $r = 2$, we have

$$E(N - n^*)^2 = 2\gamma^{-1}n^* + O(\xi) \tag{17}$$

For $r = 3$, we have

$$E|N - n^*|^3 = O\left(\xi^2\right) \tag{18}$$

We turn to prove Theorem 3.

First, write $h(N)$ in Tayler series expansion as

$$E(h(N)) = h(n^*) + E(N - n^*)h'(n^*) + \frac{1}{2}E(N - n^*)^2 h''(n^*) + \frac{1}{6}E\left((N - n^*)^3 h'''(v)\right)$$

where v is a random variable between N and n^*. By using Equations (16)–(18) we have

$$E(h(N)) = h(n^*) + \left(\frac{1}{2} - 2\sigma^2 \gamma^{-1}\right)h'(n^*) + 2\sigma^2 \gamma^{-1} n^* h''(n^*) + \frac{1}{6}E\left((N - n^*)^3 h'''(v)\right)$$

However, $\frac{1}{6}E(N - n^*)^3 h'''(v) \le \frac{1}{6}E|N - n^*|^3 \sup_{n>m}|h'''(n)| = O|\xi^2 h'''(\xi)|$ from Equation (18) since $h(\cdot)$ and its derivatives are bounded. The proof is complete. □

Let N be defined as in Equation (8) and assume (A) holds, the asymptotic characteristics of the fine-tuning phase as $\xi \to \infty$ we have (see Yousef et al. [38])

$$E\left(\overline{X}_N\right) = \mu + o\left(\xi^{-1}\right)$$
$$E\left(\overline{X}_N^2\right) = \mu^2 + \sigma^2/n^* + o\left(\xi^{-1}\right)$$
$$Var\left(\overline{X}_N\right) = \sigma^2/n^* + o\left(\xi^{-2}\right)$$

Theorem 4. *Under Assumption (A) and using Equation (8), we can show that for any real k and as $\xi \to \infty$*

$$E\left(S_N^{2k}\right) = \sigma^{2k} + \frac{\sigma^{2k}}{n^*}k(\gamma k - \gamma - 2) + o\left(\xi^{-1}\right)$$

Proof. Write $E\left(S_N^{2k}\right) = E\left(\overline{V}_N^k\right)$, conditioning on the $\sigma - field$ generated $V_i(i = 1, 2, 3, \ldots, N_1 - 1)$, we have

$$E\left(\overline{V}_N^k\right) = E(N - 1)^{-1}E\left(\sum_{i=1}^{N_1 - 1} V_i + \sum_{N_1}^{N-1} V_i\right)^k |V_i(i = 1, 2, 3, \ldots, N_1 - 1)$$

Thus, we write the binomial expression as an infinite series and we get

$$E\left\{(N-1)^{-k}E\left(\sum_{i=1}^{N_1-1} V_i + \sum_{N_1}^{N-1} V_i\right)^k |V_i(i = 1, 2, 3, \ldots, N_1 - 1)\right\}$$
$$= E\left\{(N-1)^{-k}E\left(\sum_{i=1}^{\infty} \lambda(k, j)\left(\sum_{i=1}^{N_1-1} V_i\right)^{k-j}\left(\sum_{i=N_1}^{N-1} V_i\right)^j\right)|V_i(i = 1, 2, 3, \ldots, N_1 - 1)\right\},$$

where $\lambda(k, j) = \dfrac{\prod\limits_{t=1}^{j}(k-t+1)}{j!}$ for $j = 1, 2, \ldots$ and $\lambda(k, 0) = 1$.

Conditioning on the $\sigma - field$ generated by $V_i(i = 1, 2, \ldots, N_1 - 1)$ the random sum $\left(\sum\limits_{i=N_1}^{N-1} V_i\right)|V_i(i = 1, 2, 3, \ldots, N_1 - 1)$ is distributed as $\sigma^2 \chi^2$-distribution with $N - N_1$ degrees of freedom. Therefore,

$$E\left(\sum_{i=N_1}^{N-1} V_i\right)^j |V_i(i = 1, 2, 3, \ldots, N_1 - 1) = \sigma^{2j}(N - N_1)^j\left(1 + O\left(N^{-1}\right)\right)$$

Consequently, this yields

$$E\left(S_N^{2k}\right) = \sigma^{2k} E\left(1 + \frac{1}{N-1}\sum_{i=1}^{N_1-1} Y_i\right)^k + o\left(\xi^{-1}\right) \tag{19}$$

Consider the first three terms in the expansion and the remainder term $R(\xi)$

$$E\left(S_N^{2k}\right) = \sigma^{2k} + \sigma^{2k}kE\left(\frac{1}{N-1}\sum_{i=1}^{N_1-1} Y_i\right) + \left(\frac{1}{2}\right)\sigma^{2k}k(k-1)E\left(\frac{1}{N-1}\sum_{i=1}^{N_1-1} Y_i\right)^2 + E(R(\xi)) \tag{20}$$

where $E(R(\xi)) = o\left(\xi^{-1}\right)$. Let us evaluate the second term $\sigma^{2k}kE\left(\frac{1}{N-1}\sum_{i=1}^{N_1-1} Y_i\right)$, first expand $(N-1)^{-1}$ around n^* using Taylor series $(N-1)^{-1} = n^{*-1} - (N-n^*)n^{*-2} + (1/2)(N-n^*)^2 v^{-3}$, where v is a random variable lies between N and n^*. Furthermore,

$$\begin{aligned}(N-1)^{-1} &= n^{*-1} - \xi\left(\overline{V}_{N_1} - \sigma^2\right)n^{*-2} + (1/2)\xi^2\left(\overline{V}_{N_1} - \sigma^2\right)^2 v^{-3},\\ &= n^{*-1} - \left(\frac{1}{N_1-1}\sum_{i=1}^{N_1-1} Y_i\right)n^{*-1} + (1/2)\left(\frac{1}{N_1-1}\sum_{i=1}^{N_1-1} Y_i\right)^2 n^{*2} v^{-3},\end{aligned}$$

where we have used the fact that $N \approx \xi\,\overline{V}_{N_1}$. Thus,

$$\sigma^{2k}kE\left\{\frac{1}{N-1}\sum_{i=1}^{N_1-1} Y_i\right\} = n^{*-1}\sigma^{2k}kE\left(\sum_{i=1}^{N_1-1} Y_i\right) - n^{*-1}\sigma^{2k}kE\left\{(N_1-1)^{-1}\left(\sum_{i=1}^{N_1-1} Y_i\right)^2\right\} + \frac{1}{2}n^{*2}\sigma^{2k}kE\left\{v^{-3}(N_1-1)^{-2}\left(\sum_{i=1}^{N_1-1} Y_i\right)^3\right\} \tag{21}$$

However, the first term in (21) $n^{*-1}\sigma^{2k}kE\left(\sum_{i=1}^{N_1-1} Y_i\right) = 0$, by Wald's [44] first equation.

For the second term in Equation (21), $n^{*-1}\sigma^{2k}kE\left\{(N_1-1)^{-1}\left(\sum_{i=1}^{N_1-1} Y_i\right)^2\right\}$, conditioning on the $\sigma-field$ generated by $V_i(i = 1, 2, \dots, m-1)$, we have

$$n^{*-1}\sigma^{2k}kE\left\{(N_1-1)^{-1}\left(\sum_{i=1}^{N_1-1} Y_i\right)^2\right\} = \sigma^{2k}kn^{*-1}E\left\{(N_1-1)^{-1}E\left(\left(\sum_{i=1}^{m-1} Y_i + \sum_{i=m}^{N_1-1} Y_i\right)^2\right)|V_i(i = 1, 2, 3, \dots, m-1)\right\}$$

Expanding the binomial term and taking the expectation conditional on $V_i(i = 1, 2, \dots, m-1)$ then expanding $(N_1-1)^{-1}$ in Taylor series we obtain

$$n^{*-1}\sigma^{2k}kE\left\{(N_1-1)^{-1}\left(\sum_{i=1}^{N_1-1} Y_i\right)^2\right\} = 2k\sigma^{2k}n^{*-1} + o\left(\xi^{-1}\right)$$

Now, recall the third term $\left(\frac{1}{2}\right)\sigma^{2k}k(k-1)E\left(\frac{1}{N-1}\sum_{i=1}^{N_1-1} Y_i\right)^2$ in Equation (21) and expand $(N_1-1)^{-2}$ in the Taylor series and applying Wald's [44] second equation, we get

$$\left(\frac{1}{2}\right)\sigma^{2k}k(k-1)E\left(\frac{1}{N_1-1}\sum_{i=1}^{N_1-1} Y_i\right)^2 = k(k-1)\sigma^{2k}(\gamma n^*)^{-1} + o\left(\xi^{-1}\right)$$

Finally, recall the remainder term in Equation (21), and consider the following two cases:

Case 1, if $v \geq n^*$ then $n^{*-1} \geq v^{-1}$ and

$$\frac{1}{2}n^{*2}\sigma^{2k}kE\left\{v^{-3}(N_1-1)^{-2}\left(\sum_{i=1}^{N_1-1}Y_i\right)^3\right\} \leq \frac{1}{2}n^{*-1}\sigma^{2k}m^{-2}5E(N_1-1) = o\left(\xi^{-1}\right),$$

as $m \to \infty$, since $m \leq N_1$ and Assumption (A) holds.

Case 2, if $v \leq n^*$ then $v \geq N \geq N_1 \geq m$ and $m^{-1} \geq v^{-1}$, it follows that

$$\frac{1}{2}n^{*2}\sigma^{2k}kE\left\{v^{-3}(N_1-1)^{-2}\left(\sum_{i=1}^{N_1-1}Y_i\right)^3\right\} \leq \frac{1}{2}n^{*2}\sigma^{2k}5k\,m^{-5}5E(N_1-1) = o\left(\xi^{-1}\right),$$

as $m \to \infty$ by Assumption (A). Therefore,

$$E\left(S_N^{2k}\right) = \sigma^{2k} - 2k\sigma^{2k}n^{*-1} + \gamma k(k-1)\sigma^{2k}n^{*-1} + o\left(\xi^{-1}\right)$$

The proof is complete. □

3. Three-Stage Coverage Probability of the Mean

$$\begin{aligned}
P(\mu \in I_N) &= P\left(|\overline{X}_N - \mu| \leq d\right) \\
&= \sum_{n=m}^{\infty} P\left(|\overline{X}_N - \mu| \leq d, N = n\right) = \sum_{n=m}^{\infty} P\left(|\overline{X}_N - \mu| \leq d|N = n\right)P(N = n)
\end{aligned}$$

Since $\overline{X}_N$ and the events $\{N = n\}, n = m, m+1, m+2,\ldots$ are independent because N is a function of $S_{N_1}^2$ also because $\overline{X}_N$ and $S_{N_1}^2$ are independent for all $n = m, m+1, m+2,\ldots$ for the normal distribution, it follows that,

$$P(\mu \in I_N) = \sum_{n=m}^{\infty} P\left(|\overline{X}_n - \mu| \leq d\right)P(N = n) = E_N\left(2\Phi\left(d\sqrt{N}/\sigma\right)\right) - 1$$

where $\Phi(u) = (2\pi)^{-1/2}\int_{-\infty}^{u} e^{-t^2/2}\, dt$. Recall Theorem 3, it follows

$$\begin{aligned}
E_N\left(2\Phi\left(d\sqrt{N}/\sigma\right)\right) - 1 &= (2\Phi(a) - 1) + \left(\tfrac{1}{2} - 2\gamma^{-1}\right)h'(n^*) + \gamma^{-1}n^*h''(n^*) + o\left(\xi^2 h'''(\xi)\right) \\
&= (1 - \alpha) + \tfrac{a}{n^*}\Phi(a)\left(\tfrac{1}{2} - 2\gamma^{-1}\right) - \tfrac{a}{2n^*}\Phi(a)\left(5 + a^2\right)\gamma^{-1} + O\left(d^2\right) \\
&= (1 - \alpha) - \tfrac{a}{2\gamma n^*}\Phi(a)\left(5 - \gamma + a^2\right) + o\left(d^2\right)
\end{aligned}$$

as $\xi \to \infty$ The quantity $(2\gamma)^{-1}\left(5 - \gamma + a^2\right)$ is known as the cost of ignorance or the cost of not knowing to σ^2 (see Simons [45] for details).

4. The Asymptotic Regret Incurred in Estimating σ^2

Theorem 5. *The risk associated with (2) as $m \to \infty$ is given by*

$$R_N(A) = cn^*(1 + \gamma) + c(\gamma - 4)(2\gamma)^{-1} + o(1)$$

Moreover, the asymptotic regret is

$$\omega(d) = cn^*(\gamma - 1) + c(\gamma - 4)(2\gamma)^{-1} + o(1)$$

Proof. Recall the squared-error loss function given in Equation (2) and take the expectation all through,

$$R_N(A) = E(L_N(A)) = (1/2)\left(cn^{*2}/\sigma^4\right)E\left(S_N^2 - \sigma^2\right)^2 + c\,E(N)$$

By using Equation (16) and Theorem 4 with $k = 1$ we have

$$R_N(d) = cn^*(1 + \gamma) + c(\gamma - 4)(2\gamma)^{-1} + o(1)$$

while the asymptotic regret of the triple sampling point estimation of σ^2 under (2) is

$$\omega(d) = E(L_N(d)) - E(L_{n^*}(d)) = cn^*(1 + \gamma) + c(\gamma - 4)(2\gamma)^{-1} - 2cn^* + o(1)$$
$$= cn^*(\gamma - 1) + c(\gamma - 4)(2\gamma)^{-1} + o(1)$$

The proof is complete. □

Clearly, for zero cost, we obtain zero regrets. While for a non zero cost, we obtain negative regret for all $0 < \gamma < 1$. This means that the triple sampling procedure provides better estimates than the optimal (see Martinsek [46]).

5. Simulation Results

Since the results are asymptotic, it is worth mentioning to record the performance of the estimates under a moderate sample size performance. Microsoft Developer Studio software was used to run FORTRAN codes using Equations (7) and (8). A series of 50,000 replications were generated from a normal distribution with different values of μ and σ^2. The optimal sample sizes were chosen to represent small, medium to large sample sizes $n^* = 24, 43, 61, 76, 96, 125, 171, 246$, and 500 with $\gamma = 0.5$ as recommended by Hall [4,20]. For brevity, we report the case at $m = 10$.

5.1. The Mean and the Variance of the Normal Distribution

We estimate the optimal final sample size and its standard error, the mean and its standard error, the coverage probability of the mean, the variance and its standard error, the asymptotic regret of using the sample variance instead of the population variance. For constructing a fixed-width confidence interval for the mean we take $\alpha = 0.05 \Rightarrow a = 1.96$. In each Table, we report $\overline{N}$ as an estimate of n^*, $S(\overline{N})$ as the standard error of $\overline{N}$, $\hat{\mu}$ as an estimate of μ with standard error $S(\hat{\mu})$. The estimated coverage probability is $1 - \hat{\alpha}$ while the estimated asymptotic regret is $\hat{\omega}$.

The simulation process is performed as follows: Fix γ, α and n^* as in Equation (6).

First: For the i-th sample generated from the normal distribution, take a pilot sample of size m, that is $(X_{1,i}, X_{2,i}, \ldots, X_{m,i})$.

Second: Compute the sample mean $\overline{X}_i$ and the sample variance S_i^2.

Third: Apply Equations (7) and (8) to determine the stopping sample size at this iteration, whether in the first stage or the second stage, say N_i^*.

The inverse coefficient of variation is the ratio of the population mean to the population standard deviation, that is $\theta = \mu/\sigma$, $\theta \in \mathfrak{R}$ (no singularity point can exist over the entire real line). Assume further a random sample of size $n(\geq 2)$ from the normal distribution becomes available, we propose to estimate θ by $\hat{\theta}_n = \overline{X}_n S_n^{-1}$. It is a dimensionless quantity that makes comparisons across several populations that have different units of measurements has useful meanings. In practical life, the inverse coefficient of variation is equal to the signal to noise ratio, which measures how much signal has been corrupted by noise (see McGibney and Smith [47]).

Fourth: Record the resultant sample size, the sample mean, the sample standard deviation, and the estimated inverse coefficient of variation $\left(N_i^*, \overline{X}_i^*, S_i^*, \hat{\theta}_i\right)$ for $i = 1, 2, \ldots, k$ where $k = 50,000$

Hence, for each experimental combination, we have four vectors of size k as follows:

$$\left(N_1^*, N_2^*, \ldots, N_k^*\right), \left(\overline{X}_1^*, \overline{X}_2^*, \ldots, \overline{X}_k^*\right), \left(S_1^{2*}, S_2^{2*}, \ldots, S_k^{2*}\right), \left(\hat{\theta}_1^*, \hat{\theta}_2^*, \ldots, \hat{\theta}_k^*\right)$$

Let $\overline{N} = k^{-1}\sum_{i=1}^{k} N_i^*$, $\hat{\mu} = \overline{\overline{X}} = k^{-1}\sum_{i=1}^{k} \overline{X}_i$, $\hat{\sigma}^2 = \overline{S}^2 = k^{-1}\sum_{i=1}^{k} S_i^{2*}$ and $\overline{\hat{\theta}} = k^{-1}\sum_{i=1}^{k}\theta_i^*$, where, $\overline{N}$, $\overline{\overline{X}}$, $\overline{S}^2$ and $\overline{\hat{\theta}}$ are respectively the estimated mean sample size, the estimated mean of the population mean, the estimated mean of the sample variance and the estimated mean of the inverse coefficient of variation across replicates. The standard errors are, respectively,

$$S(\overline{N}) = (k(k-1))^{-1/2}\sqrt{\sum_{i=1}^{k}\left(N_i^* - \overline{N}\right)^2}$$

$$S(\hat{\mu}) = (k(k-1))^{-1/2}\sqrt{\sum_{i=1}^{k}\left(\overline{X}_i^* - \overline{\overline{X}}\right)^2}$$

$$S(\overline{S}) = (k(k-1))^{-1/2}\sqrt{\sum_{i=1}^{k}\left(S_i^{2*} - \overline{S}^2\right)^2}$$

$$S(\overline{\theta}) = (k(k-1))^{-1/2}\sqrt{\sum_{i=1}^{k}\left(\hat{\theta}_i^* - \overline{\hat{\theta}}\right)^2}$$

Fifth: The simulated regret $\hat{\omega}(A) = Ak^{-1}\sum_{i=1}^{k}\left(S_i^{2*} - \sigma^2\right)^2 + c\overline{N} - R_{n^*}$ with $A = (a/d)^4 c$.

Tables 1 and 2 below show the performance of the estimates under $m = 10$ and $\gamma = 0.5$.

Table 1. Three-stage estimation of the mean and variance of the normal distribution under a unified stopping rule with $m = 10$, $\gamma = 0.5$, $\mu = 10$, $\sigma = 5$, $\alpha = 0.05$.

n^*	$\overline{N}$	$S(\overline{N})$	$\hat{\mu}$	$S(\hat{\mu})$	$1-\hat{\alpha}$	$\hat{\sigma}$	$S(\hat{\sigma})$	ω
24	20.4	0.045	9.989	0.006	0.892	4.660	0.035	−27.58
43	38.4	0.068	9.997	0.004	0.905	4.767	0.031	−47.55
61	56.2	0.084	10.002	0.003	0.918	4.848	0.026	−65.80
76	71.4	0.095	9.997	0.003	0.925	4.892	0.022	−80.64
96	91.5	0.107	10.000	0.002	0.933	4.925	0.019	−100.50
125	120.4	0.121	10.000	0.002	0.936	4.943	0.016	−129.56
171	166.8	0.142	9.999	0.002	0.940	4.964	0.013	−175.21
246	242.6	0.170	10.000	0.001	0.945	4.978	0.010	−249.37
500	498.0	0.245	10.001	0.001	0.947	4.989	0.007	−501.99

Table 2. Three-stage estimation of the mean and variance of the normal distribution under a unified stopping rule with $m = 10$, $\gamma = 0.5$, $\mu = 5$, $\sigma = 10$, $\alpha = 0.05$.

n^*	$\overline{N}$	$S(\overline{N})$	$\hat{\mu}$	$S(\hat{\mu})$	$1-\hat{\alpha}$	$\hat{\sigma}$	$S(\hat{\sigma})$	ω
24	20.4	0.044	5.002	0.011	0.890	86.798	0.141	−27.65
43	38.5	0.068	5.007	0.008	0.905	90.952	0.125	−47.52
61	56.2	0.084	5.011	0.007	0.916	93.866	0.104	−65.84
76	71.1	0.095	5.006	0.006	0.924	95.392	0.090	−80.86
96	91.3	0.107	5.007	0.005	0.931	96.975	0.077	−100.68
125	120.4	0.121	5.002	0.004	0.935	97.756	0.064	−129.57
171	167.1	0.141	4.998	0.004	0.941	98.603	0.052	−174.91
246	242.4	0.169	4.999	0.003	0.946	99.088	0.042	−249.62
500	497.7	0.249	4.997	0.002	0.947	99.597	0.029	−502.29

Regarding the final random sample size N, we noticed that as n^* increases, $\overline{N}$ is always less than n^* (early stopping) with standard error increases, $\overline{N}/n^* \approx 1$. While as n^* increases $\hat{\mu} \approx \mu$ and $\hat{\sigma} \approx \sigma$ with

standard error decreases. Regarding the coverage probability, the three-stage procedure under the rules in Equations (7) and (8) provides coverage probabilities that are always less than the desired nominal value while it attains it only asymptotically. Regarding the estimated asymptotic regret, $\hat{\omega}$ we obtain negative regret, which agrees with the result of Theorem 5.

5.2. The Inverse Coefficient of Variation

As an application, we invest the three-stage estimation of both the mean and the variance to estimate the inverse coefficient of variation θ, and its standard error $S(\hat{\theta})$, the coverage probability of θ and the asymptotic regret. To estimate θ we perform the previous steps in addition to the simulated regret $\hat{\omega}(A)$ using a squared-error loss function with linear sampling cost is

$$\hat{\omega}(A) \approx Ak^{-1}\sum_{i=1}^{k}\left(\hat{\theta}_i - \overline{\hat{\theta}}\right)^2 + c\overline{N} - R_{n^*}$$

Table 3 below shows the performance of the procedure for estimating θ. Obviously, as n^* increases $\hat{\theta}/\theta \approx 1$ with standard errors decrease. Regarding the coverage probability of θ we noticed that $P\left(|\hat{\theta}_N - \theta| \leq d\right) \geq 0.95$ for all $\theta \in \mathfrak{R}$. This means that the procedure attains exact consistency. Regarding the asymptotic regret, we noticed that as n^* increases, the regret decreases with negative values. This means that the three-stage procedure does better than the optimal.

Table 3. Three-stage estimation of the inverse coefficient of variation under a unified stopping rule.

	$\mu = 10$, $\sigma = 5$, $\theta = 2$					$\mu = 5$, $\sigma = 10$, $\theta = 0.5$			
n^*	$\hat{\theta}$	$S(\hat{\theta})$	$\hat{w}$	$1 - \hat{\alpha}$	n^*	$\hat{\theta}$	$S(\hat{\theta})$	$\hat{w}$	$1 - \hat{\alpha}$
24	2.280	0.003	−27.47	0.979	24	0.571	0.002	−27.64	1.000
43	2.213	0.003	−47.49	0.956	43	0.554	0.001	−47.51	1.000
61	2.143	0.002	−65.76	0.967	61	0.537	0.001	−65.84	1.000
76	2.098	0.002	−80.61	0.981	76	0.526	0.001	−80.86	1.000
96	2.067	0.001	−100.47	0.990	96	0.518	0.001	−100.68	1.000
125	2.047	0.001	−129.54	0.995	125	0.512	0.001	−129.57	0.999
171	2.027	0.001	−175.20	0.998	171	0.506	0.000	−174.91	1.000
246	2.016	0.001	−249.36	0.999	246	0.504	0.000	−249.62	1.000
500	2.008	0.000	−501.99	1.000	500	0.501	0.000	−502.29	1.000

6. Conclusions

We use a three-stage procedure to tackle the point estimation problem for the variance while estimating the mean by a confidence interval with preassigned width and coverage probability. We use one unified stopping rule for this estimation and use the results in developing both point and interval estimations for the inverse coefficient of variation. Monte Carlo simulations were performed to investigate the performance of all estimators. We conclude that the estimation of the inverse coefficient of variation through the mean and variance obtained better results with negative regret. As an application in engineering reliability see Ghosh, Mukhopadhyay and Sen ([19]; chapter 1, p. 11). For applications in real-world problems see Mukhopadhyay, Datta, and Chattopadhyay [48].

Author Contributions: Conceptualization, A.Y.; methodology, A.Y and H.H.; software, A.Y.; validation, A.Y and H.H.; formal analysis, A.Y.; investigation, A.Y and H.H.; resources, A.Y.; data curation, A.Y and H.H.; writing—original draft preparation, A.Y Yousef.; writing—review and editing, A.Y.; visualization, A.Y.

References

1. Mukhopadhyay, N.; de Silva, B. *Sequential Methods and Their Applications*; CRC: New York, NY, USA, 2009.
2. Degroot, M.H. *Optimal Statistical Decisions*; McGraw-Hill: New York, NY, USA, 1970.
3. Dantzig, G.B. On the nonexistence of tests of 'students' hypothesis having power functions independent of σ. *Ann. Math. Stat.* **1940**, *11*, 186–192. [CrossRef]
4. Hall, P. Asymptotic theory and triple sampling of sequential estimation of a mean. *Ann. Stat.* **1981**, *9*, 1229–1238. [CrossRef]
5. Anscombe, F.J. Sequential estimation. *J. Roy. Stat. Soc.* **1953**, *15*, 1–21. [CrossRef]
6. Robbins, H. *Sequential Estimation of the Mean of a Normal Population. Probability and Statistics (Harald Cramer Volume)*; Almquist and Wiksell: Uppsala, Sweden, 1959; pp. 235–245.
7. Chow, Y.S.; Robbins, H. On the asymptotic theory of fixed width sequential confidence intervals for the mean. *Ann. Math. Stat.* **1965**, *36*, 1203–1212. [CrossRef]
8. Stein, C. A two-sample test for a linear hypothesis whose power is independent of the variance. *Ann. Math. Stat.* **1945**, *16*, 243–258. [CrossRef]
9. Cox, D.R. Estimation by double sampling. *Biometrika* **1952**, *39*, 217–227. [CrossRef]
10. Mukhopadhyay, N. Sequential estimation of a location parameter in exponential distributions. *Calcutta Stat. Assoc. Bull.* **1974**, *23*, 85–95. [CrossRef]
11. Mukhopadhyay, N.; Hilton, G.F. Two-stage and sequential procedures for estimating the location parameter of a negative exponential distribution. *S. Afr. Stat. J.* **1986**, *20*, 117–136.
12. Mukhopadhyay, N.; Darmanto, S. Sequential estimation of the difference of means of two negative exponential populations. *Seq. Anal.* **1988**, *7*, 165–190. [CrossRef]
13. Mukhopadhyay, N.; Hamdy, H.I. On estimating the difference of location parameters of two negative exponential distributions. *Can. J. Stat.* **1984**, *12*, 67–76. [CrossRef]
14. Ghosh, M.; Mukhopadhyay, N. Sequential point estimation of the parameter of a rectangular distribution. *Calcutta Stat. Assoc. Bull.* **1975**, *24*, 117–122. [CrossRef]
15. Mukhopadhyay, N.; Ekwo, M.E. Sequential estimation problems for the scale parameter of a Pareto distribution. *Scand. Actuar. J.* **1987**, 83–103. [CrossRef]
16. Sinha, B.K.; Mukhopadhyay, N. Sequential estimation of a bivariate normal mean vector. *Sankhya Ser. B* **1976**, *38*, 219–230.
17. Zacks, S. Sequential estimation of the mean of a lognormal distribution having a prescribed proportional closeness. *Ann. Math. Stat.* **1966**, *37*, 1688–1696. [CrossRef]
18. Khan, R.A. Sequential estimation of the mean vector of a multivariate normal distribution. *Indian Stat. Inst.* **1968**, *30*, 331–334.
19. Ghosh, M.; Mukhopadhyay, N.; Sen, P.K. *Sequential Estimation*; Wiley: New York, NY, USA, 1997.
20. Hall, P. Sequential estimation saving sampling operations. *J. Roy. Stat. Soc. B* **1983**, *45*, 1229–1238. [CrossRef]
21. Mukhopadhyay, N. A note on three-stage and sequential point estimation procedures for a normal mean. *Seq. Anal.* **1985**, *4*, 311–319. [CrossRef]
22. Mukhopadhyay, N. Sequential estimation problems for negative exponential populations. *Commun. Stat. Theory Methods A* **1988**, *17*, 2471–2506. [CrossRef]
23. Mukhopadhyay, N. Some properties of a three-stage procedure with applications in sequential analysis. *Indian J. Stat Ser. A* **1990**, *52*, 218–231.
24. Mukhopadhyay, N.; Hamdy, H.I.; Al Mahmeed, M.; Costanza, M.C. Three-stage point estimation procedures for a normal mean. *Seq. Anal.* **1987**, *6*, 21–36. [CrossRef]
25. Mukhopadhyay, N.; Mauromoustakos, A. Three-stage estimation procedures of the negative exponential distribution. *Metrika* **1987**, *34*, 83–93. [CrossRef]
26. Hamdy, H.I.; Pallotta, W.J. Triple sampling procedure for estimating the scale parameter of Pareto distribution. *Commun. Stat. Theory Methods* **1987**, *16*, 2155–2164. [CrossRef]
27. Hamdy, H.I.; Mukhopadhyay, N.; Costanza, M.C.; Son, M.S. Triple stage point estimation for the exponential location parameter. *Ann. Inst. Stat. Math.* **1988**, *40*, 785–797. [CrossRef]
28. Hamdy, H.I. Remarks on the asymptotic theory of triple stage estimation of the normal mean. *Scand. J. Stat.* **1988**, *15*, 303–310.

29. Hamdy, H.I.; AlMahmeed, M.; Nigm, A.; Son, M.S. Three-stage estimation for the exponential location parameters. *Metron* **1989**, *47*, 279–294.

30. Lohr, S.L. Accurate multivariate estimation using triple sampling. *Ann. Stat.* **1990**, *18*, 1615–1633. [CrossRef]

31. Mukhopadhyay, N.; Padmanabhan, A.R. A note on three-stage confidence intervals for the difference of locations: The exponential case. *Metrika* **1993**, *40*, 121–128. [CrossRef]

32. Takada, Y. Three-stage estimation procedure of the multivariate normal mean. *Indian J. Stat. Ser. B* **1993**, *55*, 124–129.

33. Hamdy, H.I.; Costanza, M.C.; Ashikaga, T. On the Behrens-Fisher problem: An integrated triple sampling approach. **1995**, in press.

34. Hamdy, H.I. Performance of fixed width confidence intervals under Type II errors: The exponential case. *South. African Stat. J.* **1997**, *31*, 259–269.

35. AlMahmeed, M.; Hamdy, H.I. Sequential estimation of linear models in three stages. *Metrika* **1990**, *37*, 19–36. [CrossRef]

36. AlMahmeed, M.; AlHessainan, A.; Son, M.S.; Hamdy, H.I. Three-stage estimation for the mean of a one-parameter exponential family. *Korean Commun. Stat.* **1998**, *5*, 539–557.

37. Costanza, M.C.; Hamdy, H.I.; Haugh, L.D.; Son, M.S. Type II error performance of triple sampling fixed precision confidence intervals for the normal mean. *Metron* **1995**, *53*, 69–82.

38. Yousef, A.; Kimber, A.; Hamdy, H.I. Sensitivity of Normal-Based Triple Sampling Sequential Point Estimation to the Normality Assumption. *J. Stat. Plan. Inference* **2013**, *143*, 1606–1618. [CrossRef]

39. Yousef, A. *Construction a Three-Stage Asymptotic Coverage Probability for the Mean Using Edgeworth Second-Order Approximation. Selected Papers on the International Conference on Mathematical Sciences and Statistics 2013*; Springer: Singapore, 2014; pp. 53–67.

40. Hamdy, H.I.; Son, S.M.; Yousef, S.A. Sensitivity Analysis of Multi-Stage Sampling to Departure of an underlying Distribution from Normality with Computer Simulations. *J. Seq. Anal.* **2015**, *34*, 532–558. [CrossRef]

41. Yousef, A. A Note on a Three-Stage Sequential Confidence Interval for the Mean When the Underlying Distribution Departs away from Normality. *Int. J. Appl. Math. Stat.* **2018**, *57*, 57–69.

42. Liu, W. A k-stage sequential sampling procedure for estimation of a normal mean. *J. Stat. Plan. Inf.* **1995**, *65*, 109–127. [CrossRef]

43. Son, M.S.; Haugh, H.I.; Hamdy, H.I.; Costanza, M.C. Controlling type II error while constructing triple sampling fixed precision confidence intervals for the normal mean. *Ann. Inst. Stat. Math.* **1997**, *49*, 681–692. [CrossRef]

44. Wald, A. *Sequential Analysis*; Wiley: New York, NY, USA, 1947.

45. Simon, G. On the cost of not knowing the variance when making a fixed-width confidence interval for the mean. *Ann. Math. Stat.* **1968**, *39*, 1946–1952. [CrossRef]

46. Martinsek, A.T. Negative regret, optimal stopping, and the elimination of outliers. *J. Amer. Stat. Assoc.* **1988**, *83*, 160–163. [CrossRef]

47. McGibney, G.; Smith, M.R. An unbiased signal to noise ratio measure for magnetic resonance images. *Med. Phys.* **1993**, *20*, 1077–1079. [CrossRef]

48. Mukhopadhyay, N.; Datta, S.; Chattopadhyay, S. *Applied Sequential Methodologies: Real-World Examples with Data Analysis*; CRC Press: Boca Raton, FL, USA, 2004. [CrossRef]