TWO-STAGE ESTIMATION OF MEAN IN A NEGATIVE BINOMIAL DISTRIBUTION WITH APPLICATIONS TO MEXICAN BEAN BEETLE DATA

N. Mukhopadhyay[1,*] and B. M. de Silva[2]

[1]Department of Statistics, UBox 4120, University of Connecticut
Storrs CT 06269-4120 U.S.A.
E-Mail: mukhop@uconnvm.uconn.edu

[2]Department of Mathematics & Statistics, RMIT University
City Campus, GPO Box 2476V, Melbourne, Australia
E-Mail: desilva@rmit.edu.au

ABSTRACT

Working with insect counts, Anscombe (1949) emphasized negative binomial modeling by introducing a parameterization involving $\mu(>0)$ and $\kappa(>0)$. The parameters μ, κ stood for average "infestation" and "clumping" respectively. Assuming that κ was known, Willson and Folks (1983) adopted purely sequential sampling to estimate μ whereas Mukhopadhyay and Diaz (1985) developed a two-stage methodology because of its operational convenience. We first prove a new striking result (Theorem 2.1) that claims the *asymptotic second-order efficiency* property of the two-stage procedure.

In order to handle the case when κ is unknown, we develop a new approach (Section 3) for evaluating estimators of μ. We control a new criterion, namely the *integrated coefficient of variation* (ICV), by averaging the CV with respect to a weight function for κ. A two-stage methodology is proposed and both *first*- and *second-order* properties are highlighted (Theorems 3.1-3.3).

We summarize findings from extensive sets of simulations of the two-stage methodologies both when κ is known or unknown. When κ is unknown, the robustness of the proposed methodology with respect to choices of a weight function is critically examined. In the end, both methodologies are applied to four sets of Mexican bean beetle data with encouraging findings.

*Corresponding author

Key words and Phrases: Known clumping parameter; CV approach; Bounded CV; Unknown clumping parameter; Integrated CV approach; Bounded integrated CV; Weight function; Second-order properties; Simulations; Practical application.

AMS 2000 Subject Classifications: 62L12; 62P10; 62P12.

1. INTRODUCTION

The practical uses of a negative binomial distribution in modeling data from biological and agricultural studies are quite numerous to list them all. Early on, working with insect counts, Anscombe (1949) emphasized the role of negative binomial modeling. Bliss and Owen (1958) considered some related issues as well. Kuno (1969,1972) developed sequential sampling procedures for estimating the mean of a population whose variance is a quadratic function of the population mean. Binns (1975), on the other hand, considered sequential confidence interval procedures for the mean of a negative binomial model. The works of Willson (1981) and Willson and Folks (1983), however, constitute comprehensive studies of modern sequential sampling approaches for estimating the mean of a negative binomial model. The purely sequential nature of sampling, advocated in Willson (1981) and Willson and Folks (1983), can sometimes be operationally unattractive, and hence Mukhopadhyay and Diaz (1985) developed a two-stage estimation methodology in the case of the point estimation problem.

References to a negative binomial model or sequential and multi-stage sampling in agricultural experiments, including the area of monitoring pests, such as weeds and insects, are quite numerous. Other than the citations that have already been provided, we simply mention the following works in an alphabetical order for brevity: Allen et al. (1972), Anscombe (1949), Barrigossi (1997), Berti et al. (1992), Johnson et al. (1995), Marshall (1988), Mukhopadhyay (2002), Mulekar et al. (1993), Mulekar and Young (1991,2004), Nyrop and Binns (1991), Onsager (1976), Plant and Wilson (1985), Sterling (1976), Sylvester and Cox (1961), Waters (1955), Wiles et al. (1992), Wilson (1981), Young (1994,2004), and Zou (1998). Sequential and two-stage sampling that are most relevant to this present investigation can be found in Willson (1981), Willson and Folks (1983), and Mukhopadhyay and Diaz (1985) when the clumping parameter in a negative binomial model is assumed known.

Let us suppose that we can observe a sequence of responses $\{X_1, X_2, ...\}$ of independent random variables, each having the probability mass function

$$\mathcal{P}(X = x; \mu, \kappa) = \binom{\kappa + x - 1}{\kappa - 1} \left(\frac{\mu}{\mu + \kappa}\right)^x \left(\frac{\kappa}{\mu + \kappa}\right)^\kappa, \ x = 0, 1, 2, \ldots. \tag{1.1}$$

The parameters μ, κ are both assumed finite and positive. Briefly, we write that X has the NB(μ, κ) distribution.

Here, X may, for example, stand for the count of insects on a sampling unit of plants or it may be the count of some kind of weed on a sampling unit of agricultural plot. In these examples, for the distribution (1.1), the parameter μ stands for the average number of insects or average number of weeds per sampling unit, whereas the parameter κ points toward the degree of clumping of infestation per sampling unit. A small (large) κ indicates heavy (light) clumping. The parameterization laid down in (1.1) was introduced by Anscombe (1949). For the distribution (1.1), it turns out that the mean and variance are given by

$$\mathbf{E}_{\mu,\kappa}[X] = \mu \text{ and } \sigma^2 = V_{\mu,\kappa}[X] = \mu + \kappa^{-1}\mu^2. \tag{1.2}$$

Willson (1981) and Willson and Folks (1983) investigated purely sequential point and interval estimation of μ when κ was known. Mukhopadhyay and Diaz (1985) came up with a two-stage sampling strategy for the point estimation problem considered earlier by Willson (1981) and Willson and Folks (1983) because a two-stage sampling design is operationally simpler than full-blown purely sequential sampling. Mukhopadhyay and Diaz (1985) continued to assume that κ was known. With regard to the loss function, the main theme in these papers revolved around controlling the *coefficient of variation* (CV) associated with the mean estimator.

In Section 2 of this paper, we first discuss (Theorem 2.2) a very striking property of the two-stage sampling strategy of Mukhopadhyay and Diaz (1985) in which we claim that for large sample sizes, the difference between the average sample size and the optimal fixed-sample-size remains bounded! This shows that the two-stage procedure of Mukhopadhyay and Diaz (1985) is *asymptotically second-order efficient* in the sense of Ghosh and Mukhopadhyay (1981).

Then, we develop a suitable two-stage sampling strategy in Section 3 for the point estimation of μ assuming that κ is unknown. One may be tempted to plug in an appropriate estimate of κ in the two-stage procedure of Mukhopadhyay and Diaz (1985) and hope to proceed accordingly. But, that approach does not work well because estimation of κ when μ is also unknown runs into difficulties. The method of maximum likelihood or the method of moment estimation of κ crosses our mind, and yet either method fails to estimate κ well in a large segment of the parameter space! Willson (1981) investigated and reported some of the ramifications when κ had to be estimated.

When κ is unknown, we take a completely different approach with the realization that a practitioner may be able to combine the associated uncertainty in the form of a suitable weight function $g(k)$ where $k(> 0)$ is a typical value of κ. Then, while estimating the mean μ, instead of trying to control the coefficient of variation, we make it our goal to control a new criterion, namely the *integrated coefficient of variation* (ICV). Here, the averaging of the CV is carried out with the help of the weight function $g(.)$. It is interesting to note that this procedure enjoys a broad sense of robustness property with respect to possible choices of $g(.)$.

With the help of computer simulations, we evaluate the methodologies both when κ is known or unknown. The role of the weight function $g(.)$ is critically examined. We suggest guidelines for a practitioner to choose an appropriate weight function $g(.)$ that is to be used in the methodology.

We have also applied the original two-stage methodology (assuming that κ is known) and the new two-stage methodology (assuming that κ is unknown) for Mexican bean beetle datasets. Dr. Jose Barrigossi gathered these massive datasets when he was a Ph.D. student under Dr. Leon G. Higley in the Department of Entomology, University of Nebraska-Lincoln. Dr. Barrigossi and Dr. Higley, the sources for these datasets, and Professor Linda Young kindly made these datasets available to us. Some prior information about the parameters came along these lines: The parameter κ for these data is anticipated to be quite small, near 0.3 or 0.4, but it could be smaller. Incidentally, we have applied the original two-stage methodology for this data assuming first that κ is known to be 0.3 or 0.4. Then, we have proceeded to apply the new two-stage methodology with the mean ICV (MICV) approach assuming that κ is unknown. In this latter situation, the weight function $g(.)$ is chosen so that the weighted average of κ is close to 0.3 or 0.4 but then κ is also "allowed" to be smaller. The sense of uncertainty about κ is thus built within the weight function $g(.)$.

2. KNOWN CLUMPING PARAMETER: THE CV APPROACH

In this section, we assume that the mean $\mu(>0)$ is unknown but the clumping parameter $\kappa(>0)$ is known. Having recorded n observations $X_1, ..., X_n$, along the line of Willson (1981), we suppose that the loss incurred in estimating μ by the sample mean $\overline{X}_n (= n^{-1} \sum_{i=1}^{n} X_i)$ is given by

$$L_n = \mu^{-2}(\overline{X}_n - \mu)^2. \tag{2.1}$$

The risk function associated with (2.1) is then given by $\mathbf{E}_\mu[L_n] = \sigma^2(n\mu^2)^{-1}$. Observe that this risk function may be interpreted as the square of the CV and hence we refer to the associated methodology as *the CV approach*.

A very reasonable goal of an experimenter is to make $\mathbf{E}_\mu[L_n]$ "small". One may attempt to achieve this goal by first fixing a small preassigned number $c(>0)$ and then designing a sampling strategy that would enable one to claim $\mathbf{E}_\mu[L_n] \le c^2$. Now, the fixed-sample-size required to achieve the goal $\mathbf{E}_\mu[L_n] \le c^2$ turns out to be the smallest integer

$$n \ge \sigma^2(c\mu)^{-2} = c^{-2}(\mu^{-1} + \kappa^{-1}) = n^*, \text{ say.} \tag{2.2}$$

Here, n^* is referred to as the "optimal" fixed-sample-size, but its magnitude remains unknown! Hence, aiming for the implementation of the optimal fixed-sample-size design to collect data is definitely out of question. At this point, Willson (1981) and Willson and Folks (1983) developed an appropriate purely sequential estimation strategy. But, purely sequential estimation strategies may be time-consuming, costly, and operationally cumbersome in some situations. The two-stage estimation strategy of Mukhopadhyay and Diaz (1985) is operationally more convenient. For a general overview of the area of sequential estimation, one may refer to Ghosh et al. (1997), a comprehensive resource. First we summarize the two-stage sampling design (Section 2.1) and then establish some of its associated *second-order* properties (Section 2.2).

2.1. Sampling Design of Mukhopadhyay and Diaz

Recall the expression of n^* from (2.2) and observe that $n^* > (\kappa c^2)^{-1}$ whereas this lower bound is a known entity. Let us define the pilot sample size

$$m \equiv m(c) = \left\langle (\kappa c^2)^{-1} \right\rangle + 1 \tag{2.3}$$

where $< u >$ denotes the largest integer smaller than u. Let $X_1, ..., X_m$ denote the pilot observations. Next, we choose and fix a number $\gamma(>0)$ and define

$$N \equiv N(c) = \left\langle \left\{ (\overline{X}_m + m^{-\gamma})^{-1} + \kappa^{-1} \right\} c^{-2} \right\rangle + 1. \tag{2.4}$$

It should be clear that N is an estimator of n^* based on the pilot observations. The role of the term $m^{-\gamma}$ is to make sure that the pilot estimator $\overline{X}_m + m^{-\gamma}$ for μ remains positive with probability one.

Now, having determined N, we sample the difference $(N - m)$ in the second stage thereby obtaining additional observations $X_{m+1}, ..., X_N$. Then, based on the totality of all gathered observations $X_1, ..., X_m, X_{m+1}, ..., X_N$, Mukhopadhyay and Diaz (1985) proposed to estimate μ by the associated sample mean, $\overline{X}_N$.

Mukhopadhyay and Diaz (1985) proved the following *first-order* results among others for the two-stage estimation methodology (2.3)-(2.4): With all fixed $\gamma(>0)$, as $c \to 0$, one has

$$\begin{aligned} &\text{(i)} \quad N/n^* \to 1 \text{ w.p. } 1; \quad \text{(ii) } \mathbf{E}_\mu[N/n^*] \to 1; \\ &\text{(iii) } \mathbf{E}_\mu[L_N]/c^2 \to 1. \end{aligned} \tag{2.5}$$

These exact same properties were also proved by Willson (1981) and Willson and Folks (1983) for their purely sequential estimation strategy.

2.2. Second-Order Properties

There is some chance that the two-stage procedure (2.3)-(2.4) may not proceed beyond the pilot stage, but Theorem 2.1 clearly shows that the probability of that happening is very small indeed especially when c is small. In what follows, we also considerably strengthen the result given in part (ii) of (2.5) by deriving *second-order* bounds (Theorem 2.2) for $E_\mu[N] - n^*$ for small c. The Theorem 2.3 states some of the other specific characteristics which are important in their own right.

Theorem 2.1 *For the two-stage procedure* (2.3)-(2.4), *with* $\gamma > 0$, *we have as* $c \to 0$:

$$\mathcal{P}_\mu[N = m] = O(m^{-p}),$$

where p is any arbitrary positive number.

Theorem 2.2 *For the two-stage procedure* (2.3)-(2.4), *with* $\gamma > 1$, *we have as* $c \to 0$:

$$\eta + o(1) \le \mathbf{E}_\mu[N] - n^* \le \eta + 1 + o(1)$$

where $\eta = \kappa\sigma^2\mu^{-3}$.

Theorem 2.3 *For the two-stage procedure* (2.3) -(2.4), *we have*:

(i) $n^{*-1/2}(N - n^*) \xrightarrow{\pounds} N(0, \kappa)$ *as* $c \to 0$, *if* $\gamma > 1/2$;
(ii) $n^{*-1}(N - n^*)^2$ *is uniformly integrable for* $0 < c < c_0$ *with sufficiently small* c_0, *if* $\gamma > 1$;
(iii) $V_\mu[N] = \kappa n^* + o(n^*)$, *if* $\gamma > 1$.

The proofs are deferred to the Appendix. It is easy to see that the part (ii) in (2.5) follows from Theorem 2.2, that is the conclusion in Theorem 2.2 is indeed stronger if we choose $\gamma > 1$. Part (ii) in (2.5), however, holds for all $\gamma > 0$.

2.3. Simulation Results

We generated simulated data from negative binomial distributions with a wide range of values of μ and κ. Having fixed a set of values of c, μ, κ, and γ, we determined m and N (denoted by n) independently $15,000(= r$, say) times and obtained $n_{\min} = \min_{1\le i\le r} n_i$, $n_{\max} = \max_{1\le i\le r} n_i$, $\overline{n} = r^{-1}\sum_{i=1}^{r} n_i$, the minimum, maximum and average estimated sample size respectively. For every fixed pair of values of κ and c, however, it became clear to us that the performances of our simulated experiments generally showed unmistakably similar features whether we had fixed $\mu = 3, \mu < 3$, or $\mu > 3$. Hence, we have taken the liberty in summarizing our findings only from those simulations that were run with fixed $\mu = 3$ and known κ which is specified within Figures 1-4 that would follow. Plots of $n_{\min}$, $n_{\max}$, and $\overline{n}$ have been provided in Figure 1.

Figures 1.a-1.d respectively illustrate our results obtained in the case of small sample sizes ($n^* \le 100$), low-moderate sample sizes ($100 < n^* \le 200$), high-moderate sample sizes ($1000 < n^* \le 2000$), and large sample sizes ($n^* > 10,000$) respectively. These plots show the variation of $n_{\min}, n_{\max}$ and $\overline{n}$ with different choices of $\gamma \in (1, 3)$. These plots show clearly that the choice of γ that is used in (2.4) did not appreciably influence N, the estimator of the sample size n^*. Hence, in what follows, we fix only one value for γ, namely $\gamma = 2$.

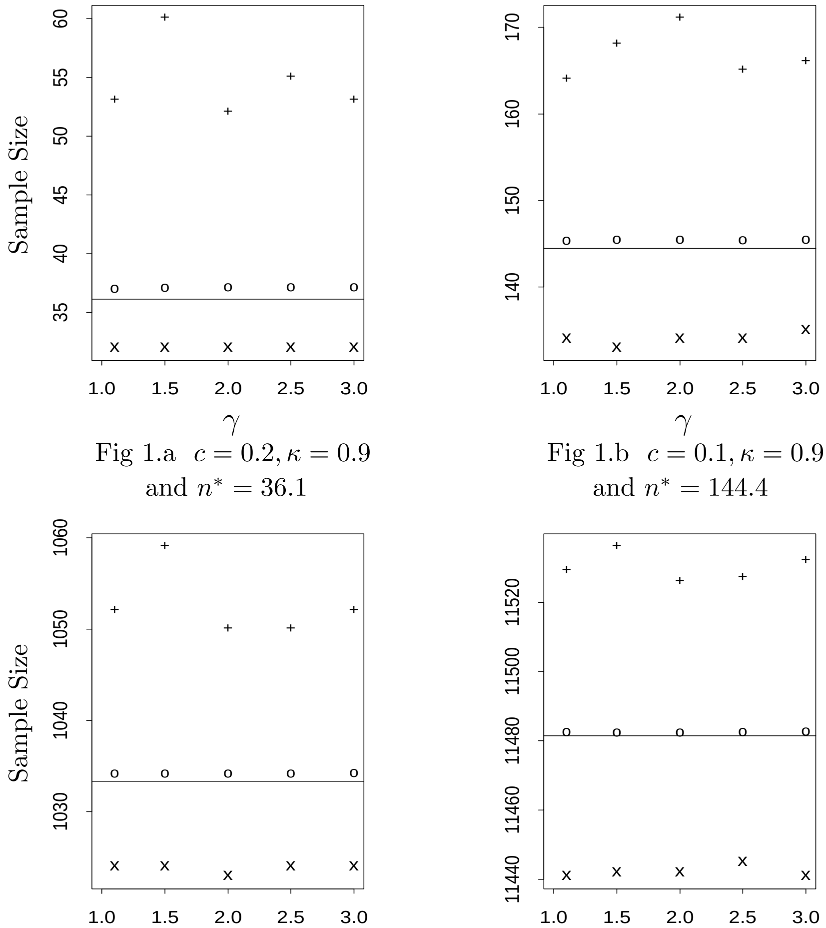

Fig 1.a $c = 0.2, \kappa = 0.9$ and $n^* = 36.1$

Fig 1.b $c = 0.1, \kappa = 0.9$ and $n^* = 144.4$

Fig 1.c $c = 0.1, \kappa = 0.1$ and $n^* = 1033.3$

Fig 1.d $c = 0.03, \kappa = 0.1$ and $n^* = 11481.5$

Figure 1. *Maximum, Minimum and Average Values of* N *Versus* γ (*Legend*: '+' $= n_{max}$, '×'$= n_{min}$, 'o' $= \overline{n}$ *and Horizontal Line at* n^*)

We do know that $(N-m)/m$ would converge to κ/μ in probability as $c \to 0$. Figure 2 plots $(\overline{n} - m)/m$, $(n_{\min} - m)/m$, and $(n_{\max} - m)/m$ for different values of c. We have also drawn a horizontal line that would correspond to the constant value κ/μ. For example, in Figure 2.a, we have $\mu = \kappa = 3$ and hence a horizontal line is drawn at 1 to highlight the limiting value. Clearly, the values of $(\overline{n} - m)/m$ always stayed very close to the limiting value κ/μ whatever be the choice for c within the range under consideration. Also, the values of $(n_{\min} - m)/m$ and $(n_{\max} - m)/m$ respectively increase and decrease to the same limiting value κ/μ as $c \to 0$.

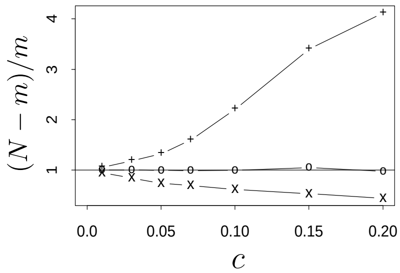

Fig 2.a $\kappa = 3$

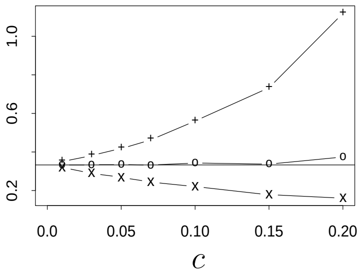

Fig 2.b $\kappa = 1$

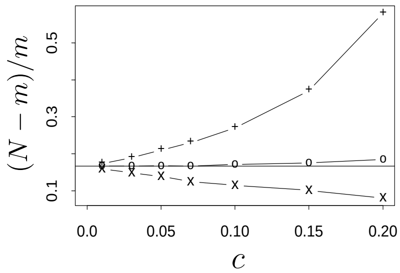

Fig 2.c $\kappa = 0.5$

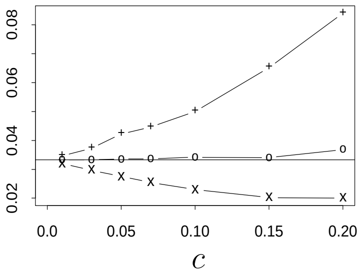

Fig 2.d $\kappa = 0.1$

Figure 2. $(N-m)/m$ *Versus* c *for* $\gamma = 2$
(*Legend*: '+' $= n_{max}$, '×'$= n_{min}$, 'o' $= \overline{n}$
and Horizontal Line at κ/μ)

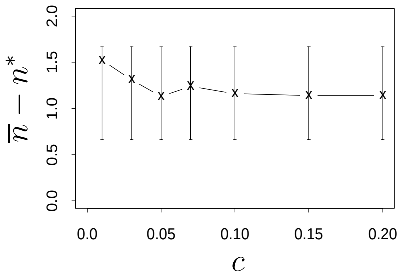

Fig 3.a $\kappa = 3$

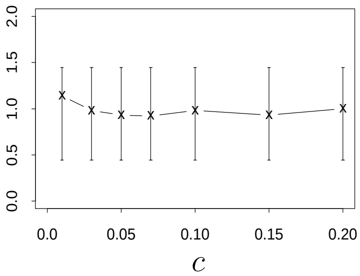

Fig 3.b $\kappa = 1$

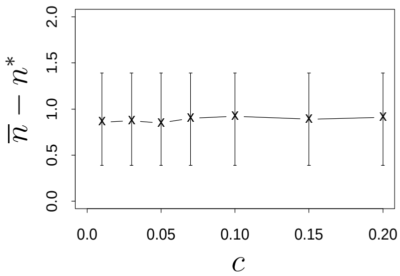

Fig 3.c $\kappa = 0.5$

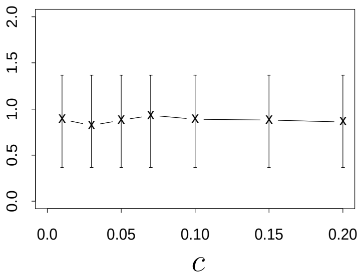

Fig 3.d $\kappa = 0.3$

Figure 3. $(\overline{n} - n^*)$ *Versus* c *for* $\gamma = 2$
Vertical Line Is Drawn From η *to* $\eta + 1$

In Figure 3, we see some "empirical justification" of the second-order result stated as Theorem 2.2. We should expect the values of $\overline{n} - n^*$ to lie between η and $\eta + 1$, especially for "large" values of n^*. As we did in the case of Figure 1, we had again chosen c and κ so that we could illustrate the behavior of $\overline{n} - n^*$ in the case of small, moderate as well as large values of n^*.

What we have found is that the values of $\overline{n} - n^*$ clearly lie between η and $\eta + 1$ for all choices of c under consideration, whatever be the known value of κ, large or small. In other words, the asymptotic bound for $\mathbf{E}_\mu[N - n^*]$ stated in Theorem 2.2 do guide us fairly accurately about what we may expect of $\mathbf{E}_\mu[N - n^*]$ for small and moderate values of n^*.

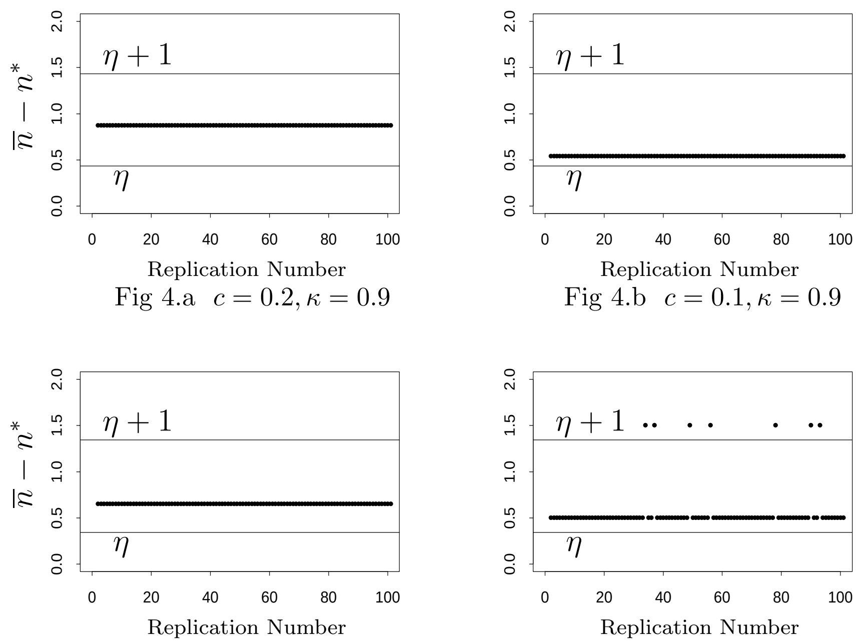

Figure 4. $(\overline{n} - n^*)$ *Versus Replication Number for* $\gamma = 2$

In Theorem 2.2, we concluded that $\mathbf{E}_\mu[N - n^*]$ was asymptotically bounded, but we could not provide the limiting value of $\mathbf{E}_\mu[N - n^*]$ as $c \to 0$. In order to investigate the status of the limiting value of $\mathbf{E}_\mu[N - n^*]$, we decided to proceed empirically as follows. For a fixed set of values of c, κ with $\mu = 3$, $\gamma = 2$, we found one $\overline{n}$ value based on $r = 10,000$ simulations. We record this as replication #1 and rewrite the corresponding $\overline{n}$ value as $\overline{n}_1$. Now, for the same fixed set of values of μ, c, κ, and γ, we ran 99 more independent replications, each based on $r = 10,000$ simulations that successively gave rise to 99 more $\overline{n}_i$ values, $i = 2, ..., 100$. This way, we obtained an entire set of 100 independent estimated values of $\mathbf{E}_\mu[N]$, each based on $r = 10,000$ simulations. Figure 4 plots $\overline{n}_i - n^*$, $i = 1, ..., 100$ for $\mu = 3$, $\gamma = 2$ and, four fixed sets of values of c and κ. We have included two small values, one medium value, and one very large value of n^* in this exercise. Only in Figure 4.d, we noticed a minuscule number (seven) of values of $\overline{n} - n^*$ out of 100 such values to stray outside of the interval $(\eta, \eta+1)$. In each situation, $\overline{n} - n^*$ values clearly appeared to converge to a limiting value that depended upon κ and μ. At this point, we conjecture that $\lim_{c\to 0} \mathbf{E}_\mu[N - n^*]$ is a constant $\xi \equiv \xi(\mu, \kappa)$. It will be an interesting mathematical exercise in the future to find an appropriate expression for the limiting value ξ.

3. UNKNOWN CLUMPING PARAMETER: THE ICV APPROACH

In this section, we assume that the mean $\mu(> 0)$ is unknown and the clumping parameter $\kappa(> 0)$ is also unknown. Having recorded n observations $X_1, ..., X_n$, we reconsider the loss function L_n from (2.1) incurred in estimating μ by the sample mean $\overline{X}_n (= n^{-1} \sum_{i=1}^n X_i)$. The risk function associated with L_n can be written as

$$\mathbf{E}_{\mu,\kappa}[L_n] = \sigma^2(n\mu^2)^{-1} = n^{-1}(\mu^{-1} + \kappa^{-1}). \tag{3.1}$$

Now, let us assume an appropriate weight function (or a density function) $g(k)$ for κ such that

$$\begin{aligned} &\text{(i)} \quad g(k) = 0 \text{ if } k \le 0, \quad \text{(ii)} \quad g(k) \ge 0 \text{ if } k > 0, \\ &\text{(iii)} \quad \textstyle\int_0^\infty g(k)dk = 1, \\ &\text{(iv)} \quad \textstyle\int_0^\infty k^{-s} g(k)dk \text{ is finite for all integers } s(> 0). \end{aligned} \tag{3.2}$$

Let us denote

$$\textstyle\int_0^\infty k^{-1} g(k)dk = a^{-1} \text{ with } a(> 0).$$

If $g(k)$ is chosen as a discrete weight function, then these and other integrals over k are to be replaced by the analogous finite (or infinite) sums over the appropriate domain of k. For brevity, however, we continue to work with the integrals with respect to $g(k)$ whenever needed.

In Section 2, the risk function (3.1) was interpreted as the square of the CV. Now, the *integrated risk* (IR) function associated with (3.1) can be expressed as

$$IR_n = \int_0^\infty \mathbf{E}_{\mu,k}[L_n] g(k)dk = n^{-1}(\mu^{-1} + a^{-1}), \tag{3.3}$$

where IR_n is interpreted as the average of squared CV, averaged with respect to the chosen weight function $g(k)$. Hence, we refer to the methodology that we are about to present as *the integrated CV* (ICV) *approach*.

A very reasonable goal of an experimenter is to try and make IR_n "small". One may attempt to achieve this goal by first fixing a small preassigned number $c(>0)$ and then designing a sampling strategy that would enable one to claim that $IR_n \le c^2$. Now, the fixed sample size required to achieve this goal, namely to have $IR_n \le c^2$, turns out to be the smallest integer

$$n \ge c^{-2}(\mu^{-1} + a^{-1}) = n^{**}, \text{ say.} \tag{3.4}$$

Here, n^{**} may be referred to as the "optimal" fixed sample size, but its magnitude remains unknown because it involves μ! Hence, aiming to implement the "optimal fixed sample size design" to collect data is again definitely out of question.

First, we develop a new two-stage sampling design and then summarize some of its associated properties in Section 3.1 analogous to those discussed earlier in Sections 2.1-2.2.

3.1. A New Sampling Design and Its Properties

Recall the expression of n^{**} from (3.4) and observe that $n^{**} > (ac^2)^{-1}$ whereas this lower bound is a known entity with $a^{-1} = \int_0^\infty k^{-1}g(k)dk$. Let us define the pilot sample size

$$m \equiv m(c) = \left\langle (ac^2)^{-1} \right\rangle + 1, \tag{3.5}$$

and let $X_1, ..., X_m$ be the pilot observations. Next, we choose and fix a number $\gamma(>0)$ and define

$$N \equiv N(c) = \left\langle \left\{ (\overline{X}_m + m^{-\gamma})^{-1} + a^{-1} \right\} c^{-2} \right\rangle + 1. \tag{3.6}$$

It should be clear that N is an estimator of n^{**} based on the pilot observations. The role of the term $m^{-\gamma}$ is to make sure again that the pilot estimator $\overline{X}_m + m^{-\gamma}$ for μ remains positive with probability one.

Now, having determined N we sample the difference $(N - m)$ in the second stage thereby obtaining additional observations $X_{m+1}, ..., X_N$. Then, based on the totality of all gathered observations $X_1, ..., X_m, X_{m+1}, ..., X_N$, we propose to estimate μ by the associated sample mean, $\overline{X}_N$.

Two important entities we should carefully analyze are the following. First, we define the *integrated average sample number* ($IASN$)

$$IASN \equiv IASN_c = \int_0^\infty \mathbf{E}_{\mu,k}[N]g(k)dk, \tag{3.7}$$

and then the *integrated sequential risk* (ISR)

$$ISR \equiv ISR_c = \int_0^\infty \mathbf{E}_{\mu,k}[L_N]g(k)dk. \tag{3.8}$$

Obviously, both $IASN$ and ISR depend upon c. The following theorems summarize crucial asymptotic behavior of the $IASN$ and ISR criteria.

Theorem 3.1 *For the two-stage procedure (3.5)-(3.6), with $\gamma > 0$, we have as* $c \to 0$:

$$n^{**-1} IASN_c \to 1, thatis, \left[\int_0^\infty \mathbf{E}_{\mu,k}[N] g(k) dk \right] / n^{**} \to 1,$$

where $g(k)$ satisfies the conditions (i)-(v) *in* (3.2). *This property is referred to as the (first-order) integrated asymptotic efficiency.*

Theorem 3.2 *For the two-stage procedure* (3.5)-(3.6), *with $\gamma > 0$, we have as $c \to 0$:*

$$c^{-2} ISR_c \to 1, thatis, \left[\int_0^\infty \mathbf{E}_{\mu,k}[L_N] g(k) dk \right] / c^2 \to 1,$$

where $g(k)$ satisfies the conditions (i)-(v) *in* (3.2). *This property is referred to as the (first-order) integrated asymptotic risk efficiency.*

Theorem 3.1 shows that the two-stage procedure is "asymptotically efficient" in the sense that the integrated average sample size may be expected to be close to n^{**} whereas Theorem 3.2 shows that the integrated sequential risk may be expected to be near c^2, the preassigned target.

Theorem 3.3 *For the two-stage procedure* (3.5)-(3.6), *with $\gamma > 1$, we have as $c \to 0$:*

$$\eta^* + o(1) \le IASN_c - n^{**} \le \eta^* + 1 + o(1)$$

where $g(k)$ satisfies the conditions (i)-(v) *in* (3.2) *and $\eta^* = a\mu^{-3}(\mu + a^{-1}\mu^2)$. This property is referred to as the (second-order) integrated asymptotic efficiency.*

Theorem 3.3 provides the asymptotic second-order bounds for $IASN_c - n^{**}$. Obviously, this result is stronger than Theorem 3.1 when $\gamma > 1$. When we could assume that κ was known, Theorem 2.2 gave analogous asymptotic second-order bounds η and $\eta + 1$ with $\eta \equiv \eta(\mu, \kappa) = \kappa\mu^{-3}(\mu + \kappa^{-1}\mu^2)$. One may note that η^* is not merely given by $\int_0^\infty \eta(\mu, k) g(k) dk$, that is by averaging $\eta(\mu, k)$ with the assumed weight function $g(k)$ of κ.

We require complicated new techniques to prove these results. The details are delegated to the Appendix.

3.2. Some Guidelines to Choose A Weight Function

In order to implement the methodology (3.5)-(3.6), one only has to obtain "a" where $a^{-1} = \int_0^\infty k^{-1} g(k) dk$. If one feels comfortable with the characteristics cited in Theorems 3.1-3.2, then any choice of $\gamma > 0$ would work just fine. The boundedness of $IASN_c - n^{**}$ holds, however, when γ is chosen to exceed one. The condition regarding the finiteness of $\int_0^\infty k^{-s} g(k) dk$ for all positive integral s is used only in the proofs of Theorems 3.1-3.3. A practitioner, however, needs to specify $g(k)$ so that a can be found easily. Other specific elicitations of $g(k)$ beyond this are not essential for the methodology to work or its characteristics as stated in Theorems 3.1-3.3 to hold. That is, the proposed methodology is very "robust" with regard to many choices of the weight function $g(k)$.

Now, if the support of $g(k)$ is chosen compact, then conditions stated in (3.2) will be automatically satisfied. Again, in this situation there may be many possible choices of $g(k)$. Suppose, for example, that the unknown parameter κ is believed to be around 0.8 but there is considerable uncertainty around this value. In what follows, we give examples (Table 1) of six possible supports and some associated $g(k)$ functions. These choices tend to model the uncertainty of κ around the value 0.8 differently from one another, and yet for each $g(k)$, one may easily verify that $a^{-1} = 1.2987$.

Table 1. Examples of Weight Functions $g(k)$ for κ:
$a^{-1} = 1.2987$

k	$g_1(k)$	$g_2(k)$	$g_3(k)$	$g_4(k)$	$g_5(k)$	$g_6(k)$
0.7	0.3	0.25631	0.53343	0.4313	0.48961	0.50000
0.8	0.6	0.70000	0.06657	0.3000	0.20000	0.30000
0.9	0.1	0.04369	0.40000	0.2687	0.30000	0.15093
1.5					0.01039	0.04907

For these six and innumerable other choices of g functions with the same value of a^{-1}, the methodology (3.5) - (3.6) will coincide!

The g function's support may instead be $(0, \infty)$ or some particular subinterval of $(0, \infty)$. In reality, one may alternatively think of a specific g function first and then accordingly determine "a" that is to be used in (3.5)-(3.6). But, then there can be many other possible g functions with the matching value a. In this sense, we can again claim that the proposed estimation methodology (3.5)-(3.6) is "robust". This is indeed an attractive characteristic of the proposed approach.

From a purely methodological point of view, pondering about a specific choice of the g function may amount to more along mental conditioning rather than something that is an absolute necessity. When the support is $(0, \infty)$, we may want to focus on the density function corresponding to an inverse gamma distribution among a host of other choices. Again, a specific choice hardly matters as long as "a" remains unchanged. We may let

$$g(k) \equiv g(k; \alpha, \beta) = \begin{cases} 0 & \text{if } k \leq 0 \\ [\Gamma(\alpha)\beta^{\alpha}]^{-1} k^{-\alpha-1} \exp\left(-(\beta k)^{-1}\right) & \text{if } k > 0 \end{cases} \tag{3.9}$$

with known $\alpha > 0, \beta > 0$. Obviously, in this case one has $\int_0^{\infty} k^{-s} g(k) dk = \beta^s \Gamma(s + \alpha)/\Gamma(\alpha), s > 0$.

An experimenter may start by plotting the $g(k)$ function with several choices of α, β, examine which one captures best the sense of uncertainty about κ, and then proceed from there. On the other hand, even though κ is unknown, an experimenter may be able to specify what may be expected on an "average". That is, one may start with a conceived value of the reciprocal of $\beta(\alpha - 1)$ which immediately demands that α be chosen exceeding one. But, having fixed $\beta(\alpha - 1)$ with some $\alpha > 1$, the experimenter would still face many possible choices! Again, one may plot some of these g functions and fairly quickly zero in on one that captures best the sense of uncertainty in practice.

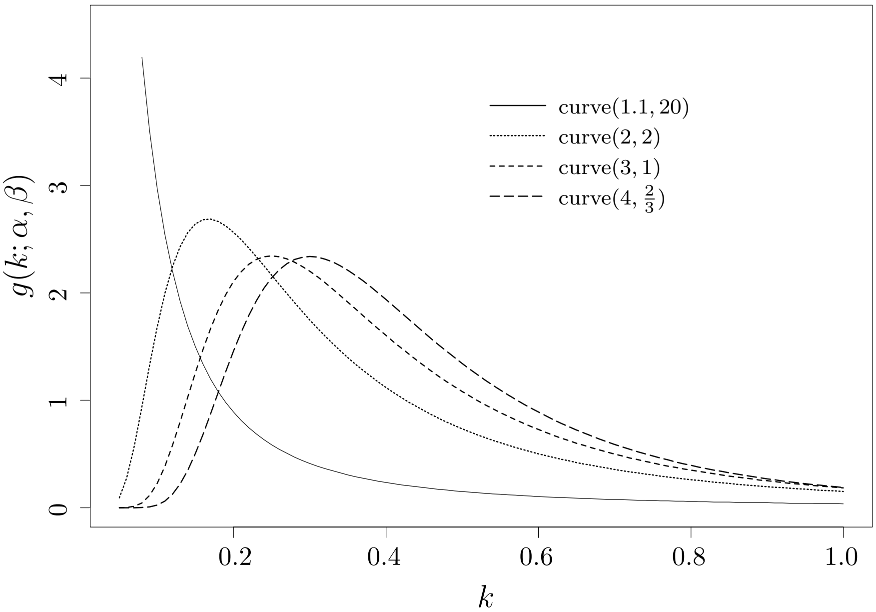

Figure 5. *Inverse Gamma Density (3.9) With* $(\alpha, \beta) = (1.1, 20), (2, 2),$ $(3, 1), (4, \frac{2}{3})$: *"Average"* κ *Is* $\frac{1}{2}$ *for Each Choice*

For example, suppose that g has to be chosen in such a way that the "average" of unknown κ with respect to g should be $\frac{1}{2}$, that is $\int_0^\infty kg(k;\alpha,\beta)dk = \beta^{-1}(\alpha - 1)^{-1} = \frac{1}{2}$. One can check, for example, that $(\alpha, \beta) = (1.1, 20)$, $(2, 2)$, $(3, 1)$, $(4, \frac{2}{3})$ are then some possible choices. One can clearly see from Figure 5 that these four g functions capture the sense of uncertainty about κ differently from one another. Once one zeroes in on a specific pair of $\alpha(> 1), \beta(> 0)$, then a^{-1} is simply $\alpha\beta$.

Alternatively, we may consider the density function associated with a lognormal distribution and let

$$g(k) \equiv g(k;\alpha,\beta) = \begin{cases} 0 & \text{if } k \le 0 \\ [\beta\sqrt{2\pi}]^{-1}k^{-1}\exp\left(-\frac{1}{2}(\log k - \alpha)^2\beta^{-2}\right) & \text{if } k > 0 \end{cases} \tag{3.10}$$

with known $-\infty < \alpha < \infty$, $\beta > 0$. Then, one obviously has

$$\int_0^\infty k^{-s} g(k) dk = \exp\left(-s\alpha + \tfrac{1}{2}s^2\beta^2\right).$$

An experimenter may again start by plotting the $g(k)$ function with several choices of α, β, examine which one captures best the sense of uncertainty about κ, and then proceed from there. We have $a^{-1} = \exp\left(-\alpha + \frac{1}{2}\beta^2\right)$ and we pick, for example, the pairs $(\alpha, \beta^2) = (2, 2), (3, 4), (\frac{5}{4}, \frac{1}{2}), (\frac{9}{8}, \frac{1}{4})$. In Figure 6, plots of these four g functions show how differently each captures the uncertainty about κ , and yet we have $a^{-1} = e^{-1} \approx 0.36788$ for each choice. That is, under any of these g functions, the statistical methodology (3.5)-(3.6) would coincide!

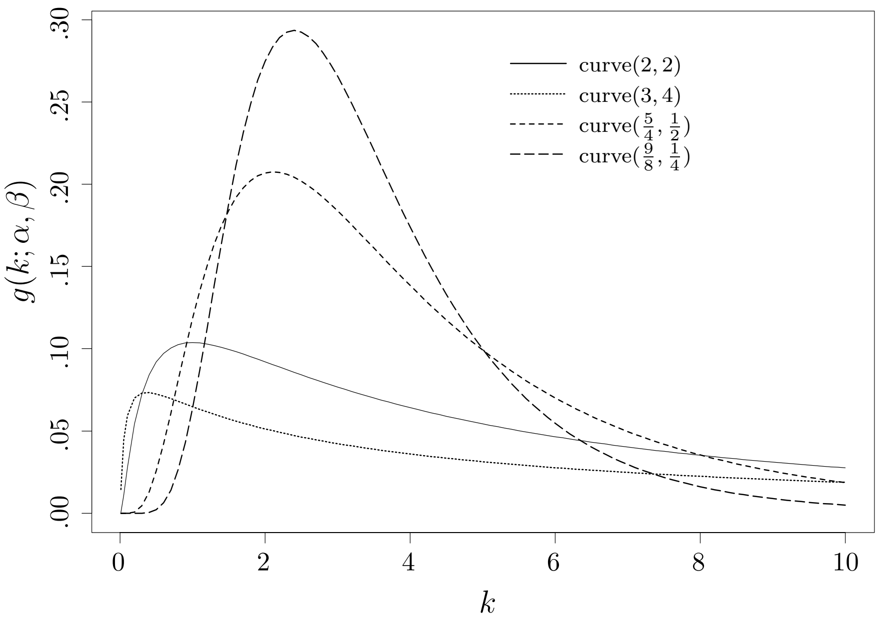

Figure 6. *Lognormal Density (3.10) With* $(\alpha, \beta^2) = (2, 2), (3, 4),$ $(\frac{5}{4}, \frac{1}{2}), (\frac{9}{8}, \frac{1}{4})$: $a^{-1} = e^{-1} \approx 0.36788$ *for Each Choice*

It should be reassuring to know that an explicit model for the uncertainty of κ would largely affect, for all practical purposes, *neither* data collection *nor* analysis of such data as long as one can zero in on the magnitude of "a".

Remark 3.1 A referee remarked that although an infinite number of $g(k)$'s result in any given value of a, the integrated risk for values of κ within the domain of $g(k)$ would likely depend on the shape of $g(k)$, and that it would be nice to know how the two might interrelate. From Figures 5 and 6, one notes that we have considered several kinds of shapes for the weight function $g(k)$. Admittedly, we have not experimented with all possible shapes for $g(k)$, but for those shapes that we have included here, we find that the actual shape of $g(k)$ has no impact on the performances of our proposed methodology (3.5)-(3.6) as long as one can zero in on the magnitude of "a".

3.3. Simulation Results

A statistical summary of simulation results obtained for the case where κ is assumed unknown is given in the following tables. As in Section 2.3, the simulation study was again conducted using a wide range of values of μ and κ. Each simulation run was replicated $15,000$ times for a fixed set of values of c, μ, κ and $\gamma = 2$. We considered $c = 0.2, 0.1$, and 0.03 in order to include examples corresponding to small, moderate, and large sample sizes n^{**}. But, for every fixed pair of values of κ and c, we again noted that the performances of our simulated experiments generally showed unmistakably similar features whether we had fixed $\mu = 3, \mu < 3$, or $\mu > 3$. Hence, we summarize our findings only from those simulations that were run with fixed $\mu = 3$.

Table 2. Simulated Values of Integrated Risk $IR_{n^{**}}$ with Fixed $\gamma = 2, a^{-1} = 1.2987$ and $g_1(k)$

κ	$\overline{IR}_{n^{**}}$	$\min(IR_{n^{**}})$	$\max(IR_{n^{**}})$
	$c = 0.20, \quad n^{**} = 40.80$		
0.7	0.03937	0.02911	0.04315
0.8	0.03938	0.02819	0.04234
0.9	0.03940	0.02852	0.04234
	$c = 0.10, \quad n^{**} = 163.20$		
0.7	0.00996	0.00922	0.01040
0.8	0.00996	0.00930	0.01027
0.9	0.00996	0.00941	0.01030
	$c = 0.03, \quad n^{**} = 1813.37$		
0.7	0.00090	0.00089	0.00091
0.8	0.00090	0.00089	0.00091
0.9	0.00090	0.00089	0.00091

In the first row of Table 2, having fixed $\mu = 3, c = 0.2, \kappa = 0.7$, and $\gamma = 2$, first we used the weight function $g_1(k)$ from Table 1 so that $a^{-1} = 1.2987$ which gave rise to $n^{**} = 40.801$. We generated a negative binomial distribution with $\mu = 3, \kappa = 0.7$, but pretending that they were unknown, we took a random sample of size $n = 41$, estimated IR_n given by (3.3), and replicated this process $15,000$ times independently. The corresponding estimated average ($\overline{IR}_{n^{**}}$), estimated minimum ($\min(IR_{n^{**}})$), and estimated maximum ($\max(IR_{n^{**}})$) values of the computed (integrated) risk are provided in the first row of Table 2. The other rows in Table 2 were constructed analogously by using the same weight function $g_1(k)$ from Table 1. Clearly, we found

$\overline{IR}_{n^{**}} \le c^2$ in every situation, however, in a very small number of simulations out of $15,000$ simulations, the computed value $IR_{n^{**}}$ exceeded c^2 ever so slightly, and hence $\max(IR_{n^{**}})$ provided in the last column turned out slightly larger than c^2.

Table 3. Estimated Average, Minimum, and Maximum Sample Size, and Average Estimated μ with Fixed $\gamma = 2, a^{-1} = 1.2987$ and $g_1(k)$

			κ	$\overline{n}$	$s_{\overline{n}}$	$n_{\min}$	$n_{\max}$	$\overline{x}$	$s_{\overline{x}}$
$c = 0.20$	$m = 33$	$n^{**} = 40.80$	0.7	41.75	0.0177	37	60	3.031	0.005
			0.8	41.74	0.0170	37	63	3.018	0.005
			0.9	41.67	0.0160	37	62	3.026	0.005
$c = 0.10$	$m = 130$	$n^{**} = 163.20$	0.7	164.17	0.0329	152	184	3.006	0.003
			0.8	164.14	0.0312	153	186	3.004	0.002
			0.9	164.05	0.0293	153	181	3.007	0.002
$c = 0.03$	$m = 1443$	$n^{**} = 1813.37$	0.7	1814.27	0.1069	1770	1867	3.001	0.001
			0.8	1814.14	0.0987	1773	1862	3.001	0.001
			0.9	1814.32	0.0942	1775	1861	3.000	0.001

In Table 3, we have supplied the average sample size $(\overline{n})$ and estimated population mean $(\overline{x})$ together with their estimated standard errors $s_{\overline{n}}, s_{\overline{x}}$ found in columns 3 and 7 respectively. We emphasize that the results presented in Table 3 were obtained assuming the weight function $g_1(k)$ for κ from Table 1 using $15,000$ simulations to produce each row. Further, $n_{\min}$ from column 4 clearly shows that simulations never terminated with pilot samples alone, $n_{\min} < n^{**} < n_{\max}$, and $(\overline{n} - n^{**})$ is positive as well as small for all values of c, κ under consideration.

The sets of simulations that provided the Tables 2-3 were then repeated assuming other g functions from Table 1. But, all these weight functions had the same associated a^{-1} value, namely 1.2987, and hence those additional sets of simulations produced results which were practically indistinguishable from what we found in Tables 2 and 3.

Next, in order to have some ideas about the performance of the proposed methodology for unknown but larger values of κ (> 1), we included the following two weight functions in our investigation:

$$g_7(k) = 0.25 \text{ for } k = 2, 4, 6, 8 \text{ and}$$
$$g_8(k) = 0.25, 0.50, 0.25 \text{ for } k = 10, 15, 20 \text{ respectively.}$$

Note that a^{-1} value associated with the weight functions $g_7(k)$ and $g_8(k)$ is 0.2604 and 0.0708 respectively. Thus, these two weight functions are associated with different n^{**} and η^* values.

Table 4. Estimated $IASN_c - n^{**}(= \Delta)$ Values Associated with $\gamma = 2$ and the Weight Function $g(k)$

	$g_1(k)$	$g_2(k)$	$g_3(k)$	$g_5(k)$	$g_7(k)$	$g_8(k)$
a^{-1}	1.2987	1.2987	1.2987	1.2987	0.2604	0.0708
η^*	0.42	0.42	0.42	0.42	0.76	1.90
			$c = 0.20$			
Δ	0.93	0.91	0.94	0.92	1.27	1.93
n^{**}	40.80	40.80	40.80	40.80	14.84	10.10
			$c = 0.10$			
Δ	0.94	0.92	0.94	0.94	1.32	2.36
n^{**}	163.20	163.20	163.20	163.20	59.37	40.41
			$c = 0.03$			
Δ	0.83	0.91	0.81	0.84	1.20	2.40
n^{**}	1813.40	1813.40	1813.40	1813.40	659.70	449.04

The integrated average sample number ($IASN_c$) from (3.7) was estimated in order to examine the validity of the result stated in Theorem 3.3 when the sample size n^{**} was small or moderate. These were estimated from $10,000$ simulations each, successively assuming the weight functions $g_i(k)$, $i = 1, \ldots, 8$. For the sake of brevity, however, we present Δ, estimated values of $IASN_c - n^{**}$, in Table 4 that are associated with the weight functions $g_i(k)$, $i = 1, 2, 3, 5, 7$, and 8. Clearly, the entries in Table 4 show that the estimated values of $IASN_c - n^{**}$ lie between η^* and $\eta^* + 1$. This feature was intuitively expected in view of Theorem 3.3 regardless of the specific nature of the weight function $g(k)$ that was used.

Remark 3.2 From Table 3, it may appear that the proposed estimator of μ has a slight positive bias, especially when $c = 0.20, m = 33$ and $n^{**} = 40.80$. But, in this particular case, we should note that both m and n^{**} are quite small whereas c appears fairly large. Hence, we would rather stay away from inferring too much from the three given simulated average $\overline{x}$ values! In the same table, however, when we consider the $\overline{x}$ and $s_{\overline{x}}$ values together for $c = 0.10, 0.03$, we note that the true $\mu(= 3)$ value is included in five out of the six two-standard deviation intervals around $\overline{x}$, namely $[\overline{x} - 2s_{\overline{x}}, \overline{x} + 2s_{\overline{x}}]$. Mathematically speaking, the magnitude of bias in $\overline{X}_N$ certainly depends upon μ, κ, m and the bias is of the order $O(m^{-2})$. That is, any bias in the estimator of μ is expected to be quite small in practice if c is small. This sentiment is clearly validated by Table 3.

4. APPLICATION TO MEXICAN BEAN BEETLE DATASETS

In this section, we handle four datasets consisting of beetle infestation of Mexican bean crop originally collected and investigated by Dr. Jose Barrigossi in his Ph.D. dissertation on *integrated pest management* (IPM) under his advisor, Dr. Leon G. Higley, in the Department of Entomology, University of Nebraska-Lincoln. For purposes of identification, we have named these datasets 'Dataset 1', 'Dataset 2', 'Dataset 3' and 'Dataset 4' respectively.

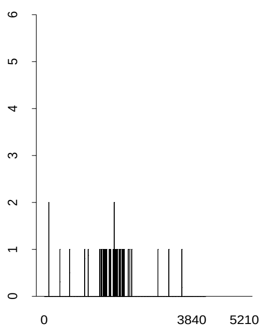

Fig 7.a Dataset 1

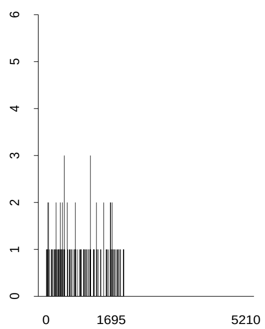

Fig 7.b Dataset 2

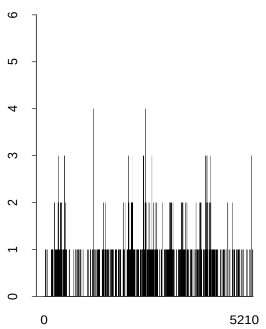

Fig 7.c Dataset 3

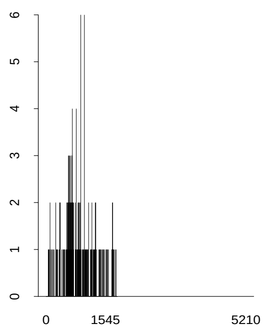

Fig 7.d Dataset 4

Figure 7. *Bar Graphs of the Four Datasets on Beetle Infestation of Mexican Bean Crop*

We aim at estimating the mean beetle infestation (μ) in Mexican bean crop with preassigned risk-bound c^2. Initially, we assumed that κ was known and hence used the CV approach (2.3)-(2.4), originally developed by Mukhopadhyay and Diaz (1985), to determine the required sample size for achieving the pre-fixed precision of the final estimator of μ. It was anticipated that κ was close to 0.3 or 0.4, and hence analysis was carried out using both these values. Next, the integrated coefficient of variation (ICV) approach developed in Section 3 was applied to these datasets for the situation assuming that κ was unknown.

Table 5. Descriptive Statistical Summaries of the Four Datasets on Beetle Infestation of Mexican Bean Crop

	Dataset 1	Dataset 2	Dataset 3	Dataset 4
No. of Obs.	3840	1695	5210	1545
Average	0.0148	0.0743	0.1451	0.1942
Variance	0.0157	0.0901	0.1690	0.2835
Std. Dev.	0.1252	0.3002	0.4111	0.5325

Figures 7.a through 7.d present the bar graphs of the Datasets 1-4. The numbers 3840, 1695, 5210 and 1545 in these figures respectively stand for the number of observations in the datasets. The gaps seen between bars correspond to 0 (zero) observed values. From Figure 7.a, we see that even though there are 3840 recorded observations in Dataset 1, many of them are 0's. The Datasets 3-4 look more appropriate for our analysis because these consist of large number of positive values and also include some high values. For example, the Dataset 3 show two 4's and Dataset 4 show two 6 's. In Table 5, we also present brief statistical summaries of these datasets. We assume that each dataset represents a random sample from a respective population and hence a random sample drawn with replacement from such a dataset would represent independent negative binomial random variables. In this analysis, we treat each dataset as a "population" and proceed to estimate the mean (μ) of each "population" by drawing random samples with replacement. That is, in each case, such observations drawn can be treated as independent random samples from a negative binomial population.

4.1. Analysis with Known Clumping Parameter

As explained earlier, we started our analysis assuming that the clumping parameter κ was known. First, we proceeded with the assumption that $\kappa = 0.3$ and then the analysis was repeated assuming $\kappa = 0.4$. From the simulation study given in Section 2.3, we found that "results" were not going to be influenced by the choice of γ, and so we had fixed the value $\gamma = 2$ for all subsequent applications. We chose $c = 0.1$ in order to come up with moderately large sample sizes associated with the two-stage procedure (2.3)-(2.4). For a given value of κ, the initial sample size m was computed from (2.3). Table 6 shows $m = 334$ when $\kappa = 0.3$ and $m = 250$ when $\kappa = 0.4$. Clearly, a smaller value of κ is associated with a larger initial sample size m. The final sample size N was computed from (2.4). Here, $\overline{X}_m$ was obtained by taking a random sample of size m with replacement from the dataset under consideration. Next, by taking another random sample of size $N - m$ with replacement from the same dataset and then combining two sets of observations, we obtained the sample mean $\overline{X}_N$.

Since there are 3785 (that is, 98.6%) zero values in Dataset 1, it turns out that with $c = 0.1$ the required number of observations from this "population" unfortunately exceeds the number of observations in the dataset. Further, for the given parameter values, the optimal sample size, n^* was computed using (2.2) and replacing the population mean μ with the mean from the dataset. In the case of Dataset 1, we note that $n^* = 7070.2$ when $\kappa = 0.3$ and $n^* = 6986.8$ when $\kappa = 0.03$ (see Table 6). Clearly, these n^* values exceed the maximum number of observations available in the dataset. Both sample means obtained from Dataset 1 using $\kappa = 0.3$ and $\kappa = 0.4$ respectively under-estimate the mean of the dataset

Table 6. Application of Two-Stage Procedure (2.3)- (2.4) Using Datasets 1-4: $c = 0.1$ and $\gamma = 2$

		Dataset 1	Dataset 2	Dataset 3	Dataset 4
$\kappa = 0.3$	m	334	334	334	334
	N	5898	1618	1076	890
	$\overline{X}_N$	0.0129	0.0773	0.1292	0.1809
	n^*	7070.2	1678.6	1022.5	848.3
$\kappa = 0.4$	m	250	250	250	250
	N	5247	1337	926	794
	$\overline{X}_N$	0.0133	0.0733	0.1350	0.1877
	n^*	6986.8	1595.2	939.2	765.0

While using Datasets 1,3 and 4 with $\kappa = 0.3$ and $\kappa = 0.4$, we find that the sample mean under-estimates the respective mean of each dataset. In the case of Dataset 2, the sample mean over-estimates (under-estimates) the mean of this dataset when $\kappa = 0.3 (\kappa = 0.4)$. However, the margin of over- or under-estimation is rather small in all cases

4.2. Analysis with Unknown Clumping Parameter

We were told that the unknown clumping parameter κ could be somewhere around 0.3 and 0.4, and hence for simplicity we assumed the weight function associated with a uniform distribution on the interval $(0.3, 0.4)$. That is, we fixed

$$g(k) = 10I(0.3 \leq k \leq 0.4)$$

with $a^{-1} = 2.876821 \approx 2.877$. Here and elsewhere, $I(.)$ stands for the indicator function of (.).

Table 7 gives the results obtained by using the ICV approach (3.5)-(3.6) with $a^{-1} = 2.877$, $c = 0.1$ and $\gamma = 2$ successively on the Datasets 1-4. As expected, both the initial and final sample sizes in this situation fell between the corresponding values of m and N that were obtained (Table 6) when we had assumed a known value $\kappa = 0.3$ or $\kappa = 0.4$. Final estimator $\overline{X}_N$ under-estimated μ for all Datasets 1-4.

Table 7. Application of Two-Stage Procedure (3.5)- (3.6) Using Datasets 1-4: $c = 0.1$**,** $\gamma = 2$ **and** $a^{-1} = 2.877$

	Dataset 1	Dataset 2	Dataset 3	Dataset 4
m	288	288	288	288
N	5085	1488	1027	842
$\overline{X}_N$	0.0136	0.0726	0.1315	0.1876
n^{**}	7024.5	1632.9	976.9	802.7

Table 7 (and Table 6) shows that the accuracy of the estimated value of μ will depend on the calculated value of a^{-1}. To examine the sensitivity of the chosen value of a^{-1}, we obtained N and $\overline{X}_N$ for different choices of a^{-1}. Figures 8 and 9 show the nature of sensitivity of N and $\overline{X}_N$ over chosen values for a^{-1} while using Datasets 1, 2, 3 and 4. We may suppose, for example, that the unknown clumping parameter κ has its weight function $g(k)$ which is positive whenever $k \in (0.20, 1.0)$ and 0 elsewhere. Then, clearly we have $1 < a^{-1} < 5$. Therefore, in Figures 8 and 9, we present "sensitivity analysis" with $a^{-1} = 1.0, 1.2, 1.4, \ldots, 5.0$. From these figures, we note that although N increases with a^{-1}, the values of $\overline{X}_N$ become quite stable when $a^{-1} > 2.5$.

Remark 4.1 A referee has pointed out that in practice, the mean, and hence, variance would be related to the sampling unit size. We could not agree more. In order to take the effect of varying sampling unit size into account, one may start with a kind of stratification with respect to sampling unit size. Then, one might try to come up with some appropriate modifications of the proposed methodologies. The idea sounds simple enough, but it may be something else when one would proceed to check some of the associated theoretical properties. We hope to pursue this problem in a later communication.

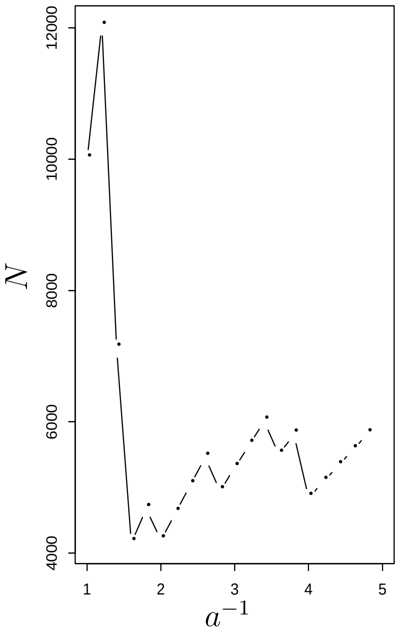

Fig 8.a N Versus a^{-1}

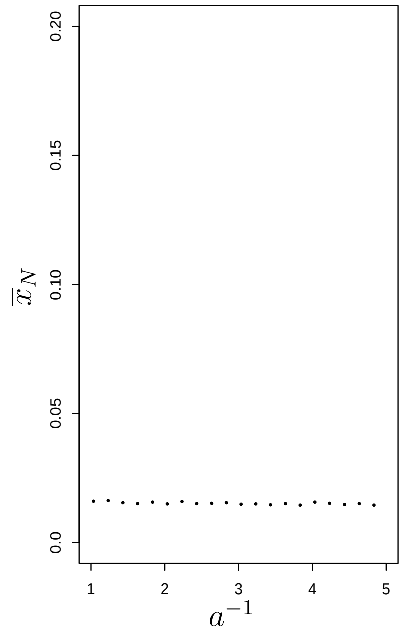

Fig 8.b $\overline{x}_N$ Versus a^{-1}

Figure 8. *Sample Size N and Sample Mean $\overline{x}_N$ Versus a^{-1} With $c = 0.1$, $\gamma = 2$ for Dataset 1*

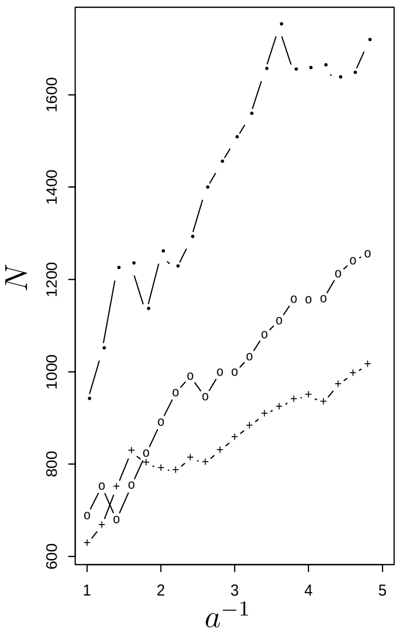

Fig 9.a N Versus a^{-1}

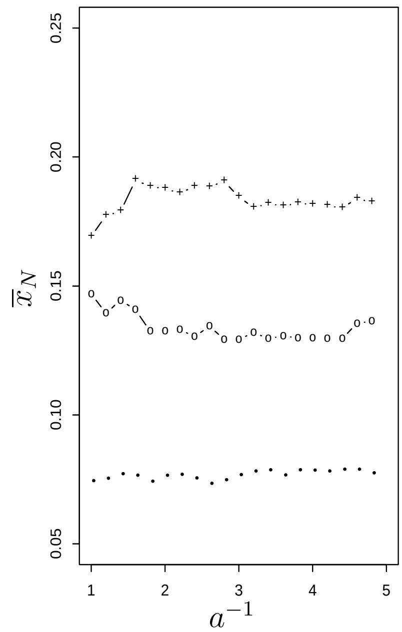

Fig 9.b $\overline{x}_N$ Versus a^{-1}

Figure 9. *Sample Size N and Sample Mean $\overline{x}_N$ Versus a^{-1} With $c = 0.1$,* $\gamma = 2$ ('$- \cdot -$'=*Dataset 2*,'$-$ o $-$'=*Dataset 3*, *and* '$- + -$'=*Dataset 4*)

APPENDIX A: DERIVATIONS AND PROOFS

We start with some preliminaries and then supply proofs of the Theorems 2.1-2.3 and Theorems 3.1-3.3. In order to prove Theorems 3.1-3.3, we need some lemmas which are proved here too.

In what follows, we frequently use a generic function $f(\overline{X}_m)$, and obtain the Taylor expansion of $f(x)$ with a remainder term by treating x as a positive *continuous* variable. We emphasize, however, that in the sequel an entity such as $\mathbf{E}_\mu[f(\overline{X}_m)]$ or $\mathbf{E}_{\mu,\kappa}[f(\overline{X}_m)]$ is evaluated using the true *discrete* distribution of the random variable $\overline{X}_m$ under consideration.

A.1. Some Preliminaries

We suppose that X has the distribution $\mathrm{NB}(\mu, \kappa)$. Let us temporarily drop the subscript for expectations, namely μ or μ, κ as the case may be. Recall that $\mathbf{E}[X] = \mu$ and $V[X] = \mu + \kappa^{-1}\mu^2$. The following factorial moments of X are easily found in Johnson and Kotz (1969, pp. 125-126):

$$\begin{aligned} \mathbf{E}[X(X-1)] &= (\kappa+1)\kappa^{-1}\mu^2; \\ \mathbf{E}[X(X-1)(X-2)] &= (\kappa+2)(\kappa+1)\kappa^{-2}\mu^3; \\ \mathbf{E}[X(X-1)(X-2)(X-3)] &= (\kappa+3)(\kappa+2)(\kappa+1)\kappa^{-3}\mu^4. \end{aligned} \tag{A.1}$$

Now, expressing X^3 and X^4 as

$$\begin{aligned} X^3 &= X(X-1)(X-2) + 3X(X-1) + X, \\ X^4 &= X(X-1)(X-2)(X-3) + 6X(X-1)(X-2) + 7X(X-1) + X, \end{aligned} \tag{A.2}$$

we obtain from (A.1)-(A.2):

$$\begin{aligned} \mathbf{E}[(X-\mu)^3] &= 2\kappa^{-2}\mu^3 + 3\kappa^{-1}\mu^2 + \mu; \\ \mathbf{E}[(X-\mu)^4] &= 3(\kappa+2)\kappa^{-3}\mu^4 + 6(\kappa+2)\kappa^{-2}\mu^3 + (3\kappa+7)\kappa^{-1}\mu^2 + \mu. \end{aligned} \tag{A.3}$$

Next, observing that $\sum_{i=1}^m X_i$ has the distribution $\mathrm{NB}(m\mu, m\kappa)$, from (A.3) we immediately obtain

$$\begin{aligned} \mathbf{E}[(\overline{X}_m-\mu)^3] &= (2\kappa^{-2}\mu^3 + 3\kappa^{-1}\mu^2 + \mu)m^{-2} = a_1 m^{-2}, \text{ say;} \\ \mathbf{E}[(\overline{X}_m-\mu)^4] &= 3\mu^2(1+\mu\kappa^{-1})^2 m^{-2} + (6\kappa^{-3}\mu^4 + 12\kappa^{-2}\mu^3 \\ &\quad + 7\kappa^{-1}\mu^2 + \mu)m^{-3} \\ &= a_2 m^{-2} + a_3 m^{-3}, \text{ say.} \end{aligned} \tag{A.4}$$

We remark in passing that we did not find the expressions given by (A.3)-(A.4) in a readily accessible source. So, these are included here for completeness.

A.2. Proof of Theorem 2.1

We can obviously claim that

$$m \leq \kappa^{-1}c^{-2} + 1 \text{ and } N \geq c^{-2}\left\{(\overline{X}_m + m^{-\gamma})^{-1} + \kappa^{-1}\right\}.$$

Thus, we can express

$$\begin{aligned}
&\mathcal{P}_\mu(N = m) \\
&\quad \leq \mathcal{P}_\mu\left[m \geq c^{-2}\left\{(\overline{X}_m + m^{-\gamma})^{-1} + \kappa^{-1}\right\}\right] \\
&\quad \leq \mathcal{P}_\mu\left[\kappa^{-1}c^{-2} + 1 \geq c^{-2}\left\{(\overline{X}_m + m^{-\gamma})^{-1} + \kappa^{-1}\right\}\right] \\
&\quad = \mathcal{P}_\mu\left[\overline{X}_m - \mu \geq c^{-2} - \mu - m^{-\gamma}\right].
\end{aligned} \tag{A.5}$$

It is clear that $\mathbf{E}_\mu[\left|\overline{X}_m - \mu\right|^{2p}] = O(m^{-p})$ for any fixed $p(>0)$. Now, we pick c sufficiently small $(\leq c_0(>0))$ such that $c^{-2} - \mu - m^{-\gamma}$ is positive. Then, for $c \leq c_0$, using Markov inequality, we can rewrite (A.5) as

$$\begin{aligned}
&\mathcal{P}_\mu(N = m) \\
&\quad = \mathcal{P}_\mu\left[\left|\overline{X}_m - \mu\right| \geq c^{-2} - \mu - m^{-\gamma}\right] \\
&\quad \leq (c^{-2} - \mu - m^{-\gamma})^{-2p}\mathbf{E}_\mu\left[\left|\overline{X}_m - \mu\right|^{2p}\right] \\
&\quad = (c^{-2} - \mu - m^{-\gamma})^{-2p}O(m^{-p}) \\
&\quad = O(m^{-3p}).
\end{aligned} \tag{A.6}$$

But, $p(>0)$ is arbitrary! ■

A.3. Proof of Theorem 2.2

Let us assume that $\gamma > 1$. Now, using (A.4) and Taylor expansion, with some random variable U between $\overline{X}_m$ and μ, we can write

$$\begin{aligned}
&\mathbf{E}_\mu[(\overline{X}_m + m^{-\gamma})^{-1}] \\
&\quad = (\mu + m^{-\gamma})^{-1} + (\mu + m^{-\gamma})^{-3}\mathbf{E}_\mu[(\overline{X}_m - \mu)^2] \\
&\qquad - (\mu + m^{-\gamma})^{-4}\mathbf{E}_\mu[(\overline{X}_m - \mu)^3] + (\mu + m^{-\gamma})^{-5}\mathbf{E}_\mu[(\overline{X}_m - \mu)^4] \\
&\qquad - \mathbf{E}_\mu[(\overline{X}_m - \mu)^5(U + m^{-\gamma})^{-6}] \\
&\quad = \mu^{-1} + O(m^{-\gamma}) + \{\mu^{-3} + O(m^{-\gamma})\}\sigma^2 m^{-1} + O(m^{-2}) \\
&\qquad - E_\mu[(\overline{X}_m - \mu)^5(U + m^{-\gamma})^{-6}].
\end{aligned} \tag{A.7}$$

Next, we proceed to prove that

$$\mathbf{E}_\mu[(\overline{X}_m - \mu)^5(U + m^{-\gamma})^{-6}] = O(m^{-2}). \tag{A.8}$$

Now, we first note that $U > \frac{1}{2}\mu$ on the set $[\overline{X}_m \geq \frac{1}{2}\mu]$ and hence

$$\begin{aligned}
&\mathbf{E}_\mu\left[\left|(\overline{X}_m - \mu)^5 (U + m^{-\gamma})^{-6} I(\overline{X}_m \geq \tfrac{1}{2}\mu)\right|\right] \\
&\quad \leq (\tfrac{1}{2}\mu)^{-6}\mathbf{E}_\mu\left[\left|\overline{X}_m - \mu\right|^5\right] \\
&\quad = O(m^{-5/2}),
\end{aligned} \tag{A.9}$$

which is $O(m^{-2})$. Here and elsewhere, we write $I(.)$ for the indicator function of $(.)$.

Next, we first use Holder's inequality and then Markov inequality to obtain

$$\begin{aligned}
&\mathbf{E}_\mu\left[\left|(\overline{X}_m - \mu)^5 (U + m^{-\gamma})^{-6} I(\overline{X}_m < \tfrac{1}{2}\mu)\right|\right] \\
&\quad \leq m^{6\gamma}\mathbf{E}_\mu[\left|\overline{X}_m - \mu\right|^5 I(\overline{X}_m < \tfrac{1}{2}\mu)] \\
&\quad \leq m^{6\gamma}\mathbf{E}_\mu^{5/6}[\left|\overline{X}_m - \mu\right|^6]\mathcal{P}_\mu^{1/6}(\overline{X}_m < \tfrac{1}{2}\mu) \\
&\quad \leq O(m^{6\gamma - \frac{5}{2}})\mathcal{P}_\mu^{1/6}(\left|\overline{X}_m - \mu\right| > \tfrac{1}{2}\mu) \\
&\quad = O(m^{6\gamma - \frac{5}{2}})O(m^{-\frac{p}{6}}),
\end{aligned} \tag{A.10}$$

which is $O(m^{-2})$, if $p(>0)$ is chosen sufficiently large. Thus, (A.8) follows from (A.9) and (A.10). (A.7) is now rewritten as

$$\mathbf{E}_\mu[(\overline{X}_m + m^{-\gamma})^{-1}] = \mu^{-1} + O(m^{-\gamma}) + \{\mu^{-3} + O(m^{-\gamma})\}\sigma^2 m^{-1} + O(m^{-2}). \tag{A.11}$$

Now, from (2.4), we obtain

$$\left\{\kappa^{-1} + (\overline{X}_m + m^{-\gamma})^{-1}\right\} c^{-2} \leq N \leq \left\{\kappa^{-1} + (\overline{X}_m + m^{-\gamma})^{-1}\right\} c^{-2} + 1, \tag{A.12}$$

and hence utilizing (2.2) and (A.11), we can write

$$\begin{aligned}
&\left\{\mathbf{E}_\mu[(\overline{X}_m + m^{-\gamma})^{-1}] - \mu^{-1}\right\} c^{-2} \leq \mathbf{E}_\mu[N - n^*] \\
&\quad \leq \left\{\mathbf{E}_\mu[(\overline{X}_m + m^{-\gamma})^{-1}] - \mu^{-1}\right\} + 1.
\end{aligned} \tag{A.13}$$

But, the left hand side of (A.13) reduces to $\kappa\sigma^2\mu^{-3}$ once we exploit the fact that $c^{-2} \approx m\kappa$. ■

A.4. Proof of Theorem 2.3

Part (i): With $\gamma > \frac{1}{2}$, observe that

$$\begin{aligned}
&m^{1/2}[(\overline{X}_m + m^{-\gamma}) - \mu] \xrightarrow{\pounds} N(0, \sigma^2) \text{ as } c \to 0 \\
\Rightarrow\ & Y_m \equiv m^{1/2}[(\overline{X}_m + m^{-\gamma})^{-1} - \mu^{-1}] \xrightarrow{\pounds} N(0, \sigma^2\mu^{-2}) \text{ as } c \to 0,
\end{aligned} \tag{A.14}$$

via Taylor expansion. Next, from (A.12), note that $n^{*-1/2}(N - n^*)$ and $n^{*-1/2}c^{-2}[(\overline{X}_m + m^{-\gamma})^{-1} - \mu^{-1}]$ must have the same asymptotic distribution. But, we can rewrite

$$n^{*-1/2}c^{-2}[(\overline{X}_m + m^{-\gamma})^{-1} - \mu^{-1}] = n^{*-1/2}c^{-2}m^{-1/2}Y_m.$$

That is, in view of (A.14), we can conclude that

$$n^{*-1/2}(N-n^*) \xrightarrow{\pounds} N(0,\kappa) \text{ as } c \to 0,$$

since $\lim_{c\to 0} n^{*-1/2}c^{-2}m^{-1/2} = \sigma^{-1}\mu\kappa^{1/2}$.

Part (ii) Let us denote $Q = \left\{\kappa^{-1} + (\overline{X}_m + m^{-\gamma})^{-1}\right\} c^{-2}$ so that we obviously have $Q \le N \le Q+1$, and hence we can write

$$|N-n^*| \le |Q-n^*| + 1. \tag{A.15}$$

Thus, part (ii) would follow if we verify that the following assertion holds:

$$(Q-n^*)^2/n^* \text{ is uniformly integrable for sufficiently small } c, \text{ if } \gamma > \tfrac{1}{2}. \tag{A.16}$$

Now, with $\rho = \min(2,\gamma)$, we can rewrite (A.11) as

$$\mathbf{E}_\mu[(\overline{X}_m + m^{-\gamma})^{-1}] = \mu^{-1} + \mu^{-3}\sigma^2 m^{-1} + o(m^{-\rho}). \tag{A.17}$$

Using a similar technique we can also have

$$\mathbf{E}_\mu[(\overline{X}_m + m^{-\gamma})^{-2}] = \mu^{-2} + (\mu^{-2} + 2\mu^{-4})\sigma^2 m^{-1} + o(m^{-\rho}). \tag{A.18}$$

Then, combining (A.17)-(A.18) one obtains

$$\mathbf{E}_\mu\left[\left\{(\overline{X}_m + m^{-\gamma})^{-1} - \mu^{-1}\right\}^2\right] = \mu^{-2}\sigma^2 m^{-1} + o(m^{-\rho}), \tag{A.19}$$

so that

$$\mathbf{E}_\mu\left[m\left\{(\overline{X}_m + m^{-\gamma})^{-1} - \mu^{-1}\right\}^2\right] \to \mu^{-2}\sigma^2$$

when $\gamma > 1$. In other words, $\mathbf{E}_\mu[mc^4(Q-n^*)^2] \to \mu^{-2}\sigma^2$, and hence $\mathbf{E}_\mu[(Q-n^*)^2/n^*] \to \kappa$ when $\gamma > 1$. Also, along the lines of part (i), we can claim that $n^{*-1/2}(Q-n^*) \xrightarrow{\pounds} N(0,\kappa)$. This distributional convergence along with the immediately preceding moment convergence imply the validity of the assertion made in (A.16).

Part (iii) We start by writing

$$V_\mu[N] = \mathbf{E}_\mu[(N - \mathbf{E}_\mu[N])^2] = \mathbf{E}_\mu[(N-n^*)^2] - (\mathbf{E}_\mu[N] - n^*)^2. \tag{A.20}$$

Next, in view of parts (i) and (ii), we have $\mathbf{E}_\mu[(N-n^*)^2] = \kappa n^* + o(n^*)$ if $\gamma > 1$. From Theorem 2.2, we immediately conclude that

$$n^{*-1/2}(\mathbf{E}_\mu[N] - n^*) = o(1) \Rightarrow (\mathbf{E}_\mu[N] - n^*)^2 = o(n^*) \tag{A.21}$$

when $\gamma > 1$. Now, part (iii) follows immediately by combining (A.20) and (A.21). ■

A.5. Proof of Theorem 3.1

Along the line of (A.7), with arbitrary $\gamma > 0$ and some random variable U between $\overline{X}_m$ and μ, we can write

$$\begin{aligned}\mathbf{E}_{\mu,\kappa}[(\overline{X}_m + m^{-\gamma})^{-1}] &= (\mu + m^{-\gamma})^{-1} + \mathbf{E}_{\mu,\kappa}[(\overline{X}_m - \mu)^2 (U + m^{-\gamma})^{-3}] \\ &= (\mu + m^{-\gamma})^{-1} + \mathbf{E}_{\mu,\kappa}[W], \text{ say.}\end{aligned} \tag{A.22}$$

Thus, along the line of (A.13), we can obtain

$$\begin{aligned}&\left\{a^{-1} + \mathbf{E}_{\mu,\kappa}[(\overline{X}_m + m^{-\gamma})^{-1}]\right\} c^{-2} \le \mathbf{E}_{\mu,\kappa}[N] \\ &\quad \le \left\{a^{-1} + \mathbf{E}_{\mu,\kappa}[(\overline{X}_m + m^{-\gamma})^{-1}]\right\} c^{-2} + 1,\end{aligned}$$

so that we have

$$\begin{aligned}&\Big\{a^{-1} + (\mu + m^{-\gamma})^{-1} + \int_0^\infty \mathbf{E}_{\mu,k}[W] g(k) dk \Big\} c^{-2} \le IASN_c \\ &\quad \le \Big\{a^{-1} (\mu + m^{-\gamma})^{-1} + \int_0^\infty \mathbf{E}_{\mu,k}[W] g(k) dk \Big\} c^{-2} + 1 \\ &\Rightarrow (a^{-1} + \mu^{-1})^{-1} \Big[\big\{a^{-1} + (\mu + m^{-\gamma})^{-1}\big\} + \int_0^\infty \mathbf{E}_{\mu,k}[W] g(k) dk \Big] \\ &\quad \le IASN_c / n^{**} \le (a^{-1} + \mu^{-1})^{-1} \Big[\big\{a^{-1} + (\mu + m^{-\gamma})^{-1}\big\} \\ &\qquad\qquad + \int_0^\infty \mathbf{E}_{\mu,k}[W] g(k) dk \Big] + n^{**-1}.\end{aligned} \tag{A.23}$$

Now, the desired result clearly follows from (A.23) if we prove the following:

$$\int_0^\infty \mathbf{E}_{\mu,k}[W] g(k) dk \to 0 \text{ as } c \to 0. \tag{A.24}$$

In doing so, let us upgrade the argument that led to (2.6) in Mukhopadhyay and Diaz (1985) to obtain

$$\begin{aligned}&\mathbf{E}_{\mu,\kappa}[W I(\overline{X}_m \ge \tfrac{1}{2}\mu)] \le 8(\mu + \kappa^{-1}\mu^2)\mu^{-3} m^{-1} \\ &\Rightarrow 0 \le \int_0^\infty \mathbf{E}_{\mu,k}[W I(\overline{X}_m \ge \tfrac{1}{2}\mu)] g(k) dk \le 8(\mu + a^{-1}\mu^2)\mu^{-3} m^{-1} \\ &\qquad \Rightarrow \lim_{c\to 0} \int_0^\infty \mathbf{E}_{\mu,k}[W I(\overline{X}_m \ge \tfrac{1}{2}\mu)] g(k) dk = 0.\end{aligned} \tag{A.25}$$

Similarly, upgrading the argument that led to (2.7) in Mukhopadhyay and Diaz (1985), with $Z = \left|\overline{X}_m - \mu\right|$ and arbitrary positive integer p, we claim

$$\mathbf{E}_{\mu,\kappa}[W I(\overline{X}_m < \tfrac{1}{2}\mu)] \le (2/\mu)^{2p} m^{3\gamma} \mathbf{E}_{\mu,\kappa}^{1/2}[Z^4] \mathbf{E}_{\mu,\kappa}^{1/2}[Z^{2p}]. \tag{A.26}$$

At this point, referring to the expression of the j^{th} factorial moment of X given in Johnson and Kotz (1969, p. 126), it follows that for sufficiently small c, with some positive integer $q \equiv q(p)$, we can express

$$\mathbf{E}_{\mu,\kappa}[Z^4]\mathbf{E}_{\mu,\kappa}[Z^{2p}] \leq m^{-2-p}\sum_{s=0}^{q} b_s\kappa^{-s}$$

where b_s's are non-negative, b_s's may involve μ and other constants, but not κ. Then, we obtain

$$\begin{aligned} &0 \leq \int_0^\infty \mathbf{E}_{\mu,k}[WI(\overline{X}_m < \tfrac{1}{2}\mu)]g(k)dk \leq (2/\mu)^p m^\gamma m^{-1-\frac{1}{2}q}\Delta^*, \\ &\text{where } \Delta^* = \int_0^\infty \sqrt{\sum_{s=0}^{q} b_s k^{-s}} g(k)dk, \Delta^* > 0, \end{aligned} \tag{A.27}$$

But, by Jensen's inequality, one observes that

$$\begin{aligned} &0 < \Delta^* \leq \sqrt{\sum_{s=0}^{q}\int_0^\infty b_s k^{-s} g(k)dk} \quad \text{which is finite, and does not involve } c \\ &\Rightarrow \lim_{c\to 0}\int_0^\infty \mathbf{E}_{\mu,k}[WI(\overline{X}_m < \tfrac{1}{2}\mu)]g(k)dk = 0, \end{aligned} \tag{A.28}$$

once p is chosen sufficiently large. Now combining the last steps from (A.25) and (A.28), the validity of (A.24) is immediate. ■

A.6. Proof of Theorem 3.2

Let us first denote $\boldsymbol{X}_m = (X_1, ..., X_m)$, $S_m = \sum_{i=1}^m X_i$, $S_N = \sum_{i=1}^N X_i$, $S_N^* = \sum_{i=1}^{N-m} X_i$ and express

$$\begin{aligned} &\mathbf{E}_{\mu,\kappa}[(\overline{X}_N - \mu)^2] \\ &\quad = \mathbf{E}_{\mu,\kappa}\left[\tfrac{(S_m - m\mu)^2}{N^2} + \tfrac{(S_N^* - (N-m)\mu)^2}{N^2} + 2\tfrac{(S_m - m\mu)(S_N^* - (N-m)\mu)}{N^2}\right] \\ &\quad = \mathbf{E}_{\mu,\kappa}\left\{\mathbf{E}\left[\tfrac{(S_m - m\mu)^2}{N^2} + \tfrac{(S_N^* - (N-m)\mu)^2}{N^2} + 2\tfrac{(S_m - m\mu)(S_N^* - (N-m)\mu)}{N^2}\middle| \boldsymbol{X}_m\right]\right\}. \end{aligned} \tag{A.29}$$

Now, while taking the conditional expectation, we observe that given $\boldsymbol{X}_m$ we can claim (i) both S_m and N are fixed and (ii) S_N^* is the sum of $N - m$ independent and identically distributed random variables. Thus, (A.29) simplifies to

$$\begin{aligned} \mathbf{E}_{\mu,\kappa}[(\overline{X}_N - \mu)^2] &= \mathbf{E}_{\mu,\kappa}\left[\tfrac{(S_m - m\mu)^2}{N^2}\right] + \mathbf{E}_{\mu,\kappa}\left[\tfrac{(N-m)\sigma^2}{N^2}\right] \\ &= J_{\mu,\kappa}^{(1)} + J_{\mu,\kappa}^{(2)}, \text{ say.} \end{aligned} \tag{A.30}$$

Lemma A.1 *For the two-stage procedure (3.4)-(3.5) , with* $\gamma > 0$, *we have*:

$$c^{-2}\mu^{-2}\int_0^\infty \sigma^2 \mathbf{E}_{\mu,k}\left[\frac{1}{N}\right] g(k)dk \to 1 \text{ as } c \to 0.$$

Lemma A.2 *For the two-stage procedure (3.4)-(3.5), with* $\gamma > 0$, *we have*:

$$mc^{-2}\mu^{-2}\int_0^\infty \sigma^2 \mathbf{E}_{\mu,k}\left[\frac{1}{N^2}\right] g(k)dk \to \frac{\mu}{\mu+a} \text{ as } c \to 0.$$

Lemma A.3 *For the two-stage procedure (3.4)-(3.5), with* $\gamma > 0$, *we have*:

$$c^{-2}\mu^{-2}\int_0^\infty J^{(1)}_{\mu,k} g(k)dk \to \frac{\mu}{\mu+a} \text{ as } c \to 0$$

where $J^{(1)}_{\mu,\kappa} = \mathbf{E}_{\mu,\kappa}\left[\frac{1}{N^2}(S_m - m\mu)^2\right]$.

Subsequently, we prove these three results. Recall from (3.7) that $ISR_c/c^2 = \mu^{-2}c^{-2}\int_0^\infty E_{\mu,k}[(\overline{X}_N - \mu)^2]g(k)dk$. Hence, assuming the validity of these three lemmas, the desired result follows immediately from (A.30). ■

A.6.1. Proof of Lemma A.1

Let us first define

$$M_d = c^{-2}\left\{a^{-1} + (\overline{X}_m + m^{-\gamma})^{-1}\right\} + d \text{ with } d = 0, 1$$

and investigate the behavior of $c^{-2}\mu^{-2}\sigma^2\mathbf{E}_{\mu,\kappa}\left[\frac{1}{M_d}\right]$. Observe that

$$M_d^{-1} = \frac{ac^2(\overline{X}_m + m^{-\gamma})}{(\overline{X}_m + m^{-\gamma})(1 + dac^2) + a} = f(\overline{X}_m), \text{ say}, \tag{A.31}$$

where $f(x) = \frac{ac^2(x + m^{-\gamma})}{(x + m^{-\gamma})(1 + dac^2) + a}$, $x > 0$. Now, we note that

$$\begin{aligned} f'(x) &= a^2c^2[(x + m^{-\gamma})(1 + dac^2) + a]^{-2} \text{ and} \\ f''(x) &= -2a^2c^2(1 + dac^2)[(x + m^{-\gamma})(1 + dac^2) + a]^{-3}. \end{aligned}$$

Then, by Taylor expansion, with some random variable U between $\overline{X}_m$ and μ, we can write

$$\begin{aligned} M_d^{-1} &= f(\mu) + (\overline{X}_m - \mu)f'(\mu) + \tfrac{1}{2}(\overline{X}_m - \mu)^2 f'(U) \\ \Rightarrow \mathbf{E}_{\mu,\kappa}[M_d^{-1}] &= \tfrac{ac^2(\mu + m^{-\gamma})}{(\mu + m^{-\gamma})(1 + dac^2) + a} - a^2c^2(1 + dac^2) \\ &\quad \times \mathbf{E}_{\mu,\kappa}[(\overline{X}_m - \mu)^2\{(U + m^{-\gamma})(1 + dac^2) + a\}^{-3}] \\ &= J^{(3)}_{\mu,\kappa} - a^2c^2(1 + dac^2)J^{(4)}_{\mu,\kappa}, \text{ say}. \end{aligned} \tag{A.32}$$

First, it is easy to see that

$$\begin{aligned}
&\lim_{c\to 0}\int_0^\infty c^{-2}\mu^{-2}\sigma^2 J^{(3)}_{\mu,k} g(k)dk \\
&\quad = \lim_{c\to 0}\tfrac{a(\mu+m^{-\gamma})}{(\mu+m^{-\gamma})(1+dac^2)+a}\mu^{-2}\int_0^\infty \sigma^2 g(k)dk \\
&\quad = \lim_{c\to 0}\tfrac{a(\mu+m^{-\gamma})}{(\mu+m^{-\gamma})(1+dac^2)+a}\mu^{-2}(\mu+a^{-1}\mu^2)=1.
\end{aligned} \tag{A.33}$$

Next, we estimate $J^{(4)}_{\mu,\kappa}$ as follows: We write

$$\begin{aligned}
J^{(4)}_{\mu,\kappa} &\le \{m^{-\gamma}(1+dac^2)+a\}^{-3}\mathbf{E}_{\mu,\kappa}\left[(\overline{X}_m-\mu)^2\right] \\
&\le \{m^{-\gamma}(1+dac^2)+a\}^{-3}\sigma^2 m^{-1}.
\end{aligned}$$

Now, we consider the second term in the last step of (A.32). With $\Delta^* \equiv \int_0^\infty \sigma^2 g(k)dk = \mu + a^{-1}\mu^2$, we have

$$\begin{aligned}
&c^{-2}\mu^{-2}a^2c^2(1+dac^2)\{m^{-\gamma}(1+dac^2)+a\}^{-3}m^{-1}\int_0^\infty \sigma^2 g(k)dk \\
&\quad \le \Delta^*\mu^{-2}a^2(1+dac^2)\{(\tfrac{1}{2}\mu+m^{-\gamma})(1+dac^2)+a\}^{-3}m^{-1},
\end{aligned} \tag{A.34}$$

which converges to zero as $c\to 0$. A combination of (A.31)-(A.34) proves the following result for all $\gamma > 0$:

$$\lim_{c\to 0} c^{-2}\mu^{-2}\int_0^\infty \sigma^2 \mathbf{E}_{\mu,k}\left[\frac{1}{M_d}\right] g(k)dk = 1 \text{ with } d=0,1. \tag{A.35}$$

Along the line of (A.12), from (3.5) we can claim that

$$M_1^{-1} \le N^{-1} \le M_0^{-1},$$

and hence, in view of (A.35), the lemma follows. ■

A.6.2. Proof of Lemma A.2

We reconsider M_d from Section A.6.1 and investigate the behavior of

$$mc^{-2}\mu^{-2}\sigma^2\mathbf{E}_{\mu,\kappa}\left[\frac{1}{M_d^2}\right].$$

Observe that

$$M_d^{-2} = \tfrac{a^2c^4(\overline{X}_m+m^{-\gamma})^2}{[(\overline{X}_m+m^{-\gamma})(1+dac^2)+a]^2} = f(\overline{X}_m), \text{ say,} \tag{A.36}$$

where $f(x) = \tfrac{a^2c^4(x+m^{-\gamma})^2}{[(x+m^{-\gamma})(1+dac^2)+a]^2}, x>0$. Observe that

$$\begin{aligned}
f'(x) &= 2a^3c^4(x+m^{-\gamma})[(x+m^{-\gamma})(1+dac^2)+a]^{-3} \text{ and} \\
f''(x) &= 2a^3c^4\{-2(x+m^{-\gamma})(1+dac^2)+a\}[(x+m^{-\gamma})] \\
&\quad \times (1+dac^2)+a]^{-4}.
\end{aligned}$$

Then, by Taylor expansion, with some random variable W between $\overline{X}_m$ and μ, we can write

$$M_d^{-2} = f(\mu) + (\overline{X}_m - \mu)f^{'}(\mu) + \tfrac{1}{2}(\overline{X}_m - \mu)^2 f^{''}(W),$$

which implies that

$$\begin{aligned}\mathbf{E}_{\mu,\kappa}[M_d^{-2}] &= \tfrac{a^2c^4(\mu+m^{-\gamma})^2}{[(\mu+m^{-\gamma})(1+dac^2)+a]^2} + a^3c^4\mathbf{E}_{\mu,\kappa}\left[(\overline{X}_m - \mu)^2 \tfrac{\{-2(W+m^{-\gamma})(1+dac^2)+a\}}{[(W+m^{-\gamma})(1+dac^2)+a]^4}\right] \\ &= J_{\mu,\kappa}^{(5)} + a^3c^4\mathbf{E}_{\mu,\kappa}\left[J_{\mu,\kappa}^{(6)}\right], \text{ say.}\end{aligned} \tag{A.37}$$

First, it is easy to see that

$$\begin{aligned}\lim_{c\to 0} c^{-2}\mu^{-2}\int_0^\infty m\sigma^2 J_{\mu,k}^{(5)} g(k)dk &= \lim_{c\to 0}\mu^{-2}m\tfrac{a^2c^2(\mu+m^{-\gamma})^2}{[(\mu+m^{-\gamma})(1+dac^2)+a]^2}\int_0^\infty \sigma^2 g(k)dk \\ = \lim_{c\to 0}\mu^{-2}m\tfrac{a^2c^2(\mu+m^{-\gamma})^2}{[(\mu+m^{-\gamma})(1+dac^2)+a]^2}(\mu + a^{-1}\mu^2) &= \frac{\mu}{\mu+a}.\end{aligned} \tag{A.38}$$

Next, we first write

$$J_{\mu,\kappa}^{(6)} I(\overline{X}_m \geq \tfrac{1}{2}\mu) \leq \{-(\mu + 2m^{-\gamma})(1+dac^2)+a\}a^{-4}\sigma^2 m^{-1}, \tag{A.39}$$

and with $\Delta^* \equiv \int_0^\infty \sigma^4 g(k)dk$, we can immediately express

$$\begin{aligned}&\lim_{c\to 0} c^{-2}\mu^{-2}a^3c^4\int_0^\infty m\sigma^2\mathbf{E}_{\mu,\kappa}\left[J_{\mu,k}^{(6)} I(\overline{X}_m \geq \tfrac{1}{2}\mu)\right] g(k)dk \\ &\quad \leq \mu^{-2}a^{-1}\Delta^*\lim_{c\to 0} c^2\{-(\mu+2m^{-\gamma})(1+dac^2)+a\},\end{aligned} \tag{A.40}$$

which converges to zero as $c \to 0$.

Now, along the lines of (A.26)-(A.28), with a positive integer p and an appropriate positive integer $q \equiv q(p)$, and with a positive number Δ^* not involving c, we can claim:

$$\begin{aligned}&\lim_{c\to 0} c^{-2}\mu^{-2}a^3c^4\int_0^\infty m\sigma^2\mathbf{E}_{\mu,\kappa}\left[J_{\mu,k}^{(6)} I(\overline{X}_m < \tfrac{1}{2}\mu)\right] g(k)dk \\ &\quad \leq \Delta^*\mu^{-2}(2/\mu)^p\lim_{c\to 0} c^2 m^{1-\frac{1}{2}q} = 0,\end{aligned} \tag{A.41}$$

since q can be made arbitrarily large by choosing p large. A combination of (A.40)-(A.41) proves the following result for all $\gamma > 0$:

$$\lim_{c\to 0} c^{-2}\mu^{-2}a^3c^4\int_0^\infty m\sigma^2\mathbf{E}_{\mu,\kappa}\left[J_{\mu,k}^{(6)}\right] g(k)dk = 0 \text{ with } d = 0,1. \tag{A.42}$$

Now, a combination of (A.37),(A.38), and (A.42) proves the following result for all $\gamma > 0$:

$$\lim_{c\to 0} c^{-2}\mu^{-2}m\int_0^\infty \sigma^2\mathbf{E}_{\mu,k}\left[\frac{1}{M_d^2}\right] g(k)dk = \frac{\mu}{\mu+a} \text{ with } d = 0,1. \tag{A.43}$$

Again, since we have $M_1^{-1} \leq N^{-1} \leq M_0^{-1}$, we can claim that the lemma follows. ■

A.6.3. Proof of Lemma A.3

In order to handle $c^{-2}\mu^{-2}J^{(1)}_{\mu,\kappa}$, we reconsider M_d yet one more time from Section A.6.1 and investigate the behavior of $c^{-2}\mu^{-2}\mathbf{E}_{\mu,\kappa}\left[(S_m - m\mu)^2 \frac{1}{M_d^2}\right]$. Using techniques similar to those in (A.36)-(A.37), with some random variable W between $\overline{X}_m$ and μ, we note that

$$\begin{aligned}
&\mathbf{E}_{\mu,\kappa}[M_d^{-2}(S_m - m\mu)^2] \\
&\quad = \tfrac{a^2c^4(\mu+m^{-\gamma})^2 m\sigma^2}{[(\mu+m^{-\gamma})(1+dac^2)+a]^2} + 2a^3c^4m^2 \\
&\qquad \times \mathbf{E}_{\mu,\kappa}\left[(\overline{X}_m - \mu)^3 \frac{(W+m^{-\gamma})}{[(W+m^{-\gamma})(1+dac^2)+a]^3}\right] \\
&\quad = J^{(7)}_{\mu,\kappa} + 2a^3c^4 J^{(8)}_{\mu,\kappa}, \text{ say.}
\end{aligned} \tag{A.44}$$

First, it is easy to see that

$$\begin{aligned}
&\lim_{c\to 0} c^{-2}\mu^{-2}\int_0^\infty J^{(7)}_{\mu,k} g(k)dk \\
&\quad = \lim_{c\to 0}\mu^{-2}m\tfrac{a^2c^2(\mu+m^{-\gamma})^2}{[(\mu+m^{-\gamma})(1+dac^2)+a]^2}\int_0^\infty \sigma^2 g(k)dk \\
&\quad = \lim_{c\to 0}\mu^{-2}m\tfrac{a^2c^2(\mu+m^{-\gamma})^2}{[(\mu+m^{-\gamma})(1+dac^2)+a]^2}(\mu + a^{-1}\mu^2) = \frac{\mu}{\mu+a}.
\end{aligned} \tag{A.45}$$

Next, we can obviously claim:

$$\begin{aligned}
\left|J^{(8)}_{\mu,\kappa}\right| &\leq \mathbf{E}_{\mu,\kappa}\left[\left|\overline{X}_m - \mu\right|^3 \frac{1}{[(W+m^{-\gamma})(1+dac^2)+a]^2}\right] \\
&\leq a^{-2}\mathbf{E}_{\mu,\kappa}\left[\left|\overline{X}_m - \mu\right|^3\right] \leq \mathbf{E}^{3/4}_{\mu,\kappa}\left[\left|\overline{X}_m - \mu\right|^4\right] \\
&= [a_2 + a_3]^{3/4} m^{-3/2}, \text{ using (A.4).}
\end{aligned} \tag{A.46}$$

Thus, with $\Delta^* \equiv \int_0^\infty [a_2 + a_3]^{3/4} g(k)dk$ which does not involve c, from (A.46) we can immediately write

$$\lim_{c\to 0} c^{-2}\mu^{-2}a^3c^4\int_0^\infty \left|J^{(8)}_{\mu,k}\right| g(k)dk \leq \mu^{-2}a^3\Delta^* \lim_{c\to 0} c^2 m^{-3/2} = 0. \tag{A.47}$$

In other words, by combining (A.44), (A.45), and (A.47), we have shown that

$$\lim_{c\to 0} c^{-2}\mu^{-2}\int_0^\infty \mathbf{E}_{\mu,k}\left[(S_m - m\mu)^2 \frac{1}{M_d^2}\right] g(k)dk = \frac{\mu}{\mu+a}.$$

Again, we claim that the lemma follows since $M_1^{-1} \leq N^{-1} \leq M_0^{-1}$. ■

A.7. Proof of Theorem 3.3

Let us assume that $\gamma > 1$. In what follows, when we include terms such as $O(m^{-s})$, it is our understanding that these terms do not involve the unknown parameter κ. Now, using (A.4), (A.7) and some random variable U between $\overline{X}_m$ and μ, we can write

$$\begin{aligned} &\mathbf{E}_{\mu,\kappa}\left[(\overline{X}_m + m^{-\gamma})^{-1}\right] \\ &\quad = (\mu + m^{-\gamma})^{-1} + (\mu + m^{-\gamma})^{-3}\sigma^2 m^{-1} - (\mu + m^{-\gamma})^{-4} a_1 m^{-2} \\ &\qquad + (\mu + m^{-\gamma})^{-5}\{a_2 m^{-2} + a_3 m^{-3}\} - \mathbf{E}_{\mu,\kappa}\left[(\overline{X}_m - \mu)^5 (U + m^{-\gamma})^{-6}\right]. \end{aligned} \tag{A.48}$$

where the expressions of $a_i \equiv a_i(\mu, \kappa), i = 1, 2, 3$ are explicitly available. Let us denote

$$a_i^* \equiv a_i^*(\mu) = \int_0^\infty a_i(\mu, k) g(k) dk, i = 1, 2, 3.$$

Next, employing arguments similar to those used previously, one can prove that

$$\lim_{c \to 0} m^2 \int_0^\infty \mathbf{E}_{\mu,k}\left[\left|\overline{X}_m - \mu\right|^5 (U + m^{-\gamma})^{-6}\right] g(k) dk = 0. \tag{A.49}$$

Now, combining (A.48)-(A.49), we can express

$$\begin{aligned} &\int_0^\infty \left\{\kappa^{-1} + \mathbf{E}_{\mu,k}\left[(\overline{X}_m + m^{-\gamma})^{-1}\right]\right\} c^{-2} g(k) dk \\ &\quad = c^{-2} a^{-1} + c^{-2}\left[(\mu + m^{-\gamma})^{-1} + (\mu + m^{-\gamma})^{-3} m^{-1}(\mu + a^{-1}\mu^2)\right. \\ &\quad \left. -(\mu + m^{-\gamma})^{-4} a_1^* m^{-2} + (\mu + m^{-\gamma})^{-5}\{a_2^* m^{-2} + a_3^* m^{-3}\} + O(m^{-2})\right], \end{aligned} \tag{A.50}$$

so that using the lower bound in (A.12), we have

$$\begin{aligned} &\int_0^\infty \mathbf{E}_{\mu,k}[N] g(k) dk - n^{**} \\ &\quad \geq \int_0^\infty \left\{\kappa^{-1} + \mathbf{E}_{\mu,k}\left[(\overline{X}_m + m^{-\gamma})^{-1}\right]\right\} c^{-2} g(k) dk - c^{-2}(a^{-1} + \mu^{-1}) \\ &\quad = c^{-2}\left[O(m^{-\gamma}) + (\mu^{-3} + O(m^{-\gamma})) m^{-1}(\mu + a^{-1}\mu^2) - (\mu^{-4} + O(m^{-\gamma}))\right. \\ &\quad \left. \times a_1^* m^{-2} + (\mu^{-5} + O(m^{-\gamma})) m\{a_2^* m^{-2} + a_3^* m^{-3}\} + O(m^{-2})\right]. \end{aligned} \tag{A.51}$$

Now, noting that $\lim_{c \to 0} m^{-1} c^{-2} = a$ and simplifying (A.51), we can write

$$\lim_{c \to 0}\left[\int_0^\infty \mathbf{E}_{\mu,k}[N] g(k) dk - n^{**}\right] \geq a\mu^{-3}(\mu + a^{-1}\mu^2). \tag{A.52}$$

The upper bound in Theorem 3.3 follows similarly from (A.50) after using the upper bound from (A.12). ■

ACKNOWLEDGMENT

Dr. Jose Barrigossi gathered the Mexican bean beetle datasets when he was a Ph.D. student in the Department of Entomology, University of Nebraska-Lincoln. We express our sincerest gratitude to Dr. Jose Barrigossi, Dr. Leon Higley, and Professor Linda Young for kindly making these datasets available to us. We are also immensely grateful to three referees for sharing with us a number of insightful comments on an earlier draft. Some of the thoughtful concerns and queries raised on an earlier draft have led us to rethink our position and add the Remarks 3.1-3.2 and Remarks 4.1-4.3 for clarifications. We take this opportunity to thank the associate editor and the referees for sharing their enthusiasm as well as for giving helpful pointers and positive feedback.

REFERENCES

1. Allen, J.D.; Gonzalez, D.; Gokhale, D.V. Sequential Sampling Plans for the Bollworm. *Helliothis zea Environmental Entomology* **1972**, *1*, 771-780.

2. Anscombe, F.J. The Statistical Analysis of Insect Counts Based on the Negative Binomial Distribution. *Biometrics* **1949**, *5*, 165-173.

3. Anscombe, F.J. Sampling Theory of the Negative Binomial and Logarithmic Series Distribution. *Biometrika* **1950**, *37*, 358-382.

4. Barrigossi, J.A.F. Development of an IPM System for the Mexican Bean Beetle (Epilachna Varivestis Mulsant) As a Pest of Dry Bean (Phaseolus Vulgaris L.). Ph.D. thesis **1997**, Department of Entomology, University of Nebraska, Lincoln.

5. Berti, A.; Zanin, G.; Baldoni, G.; Grignani, C.; Mazzoncini, M.; Montemurro, P.; Tei, F.; Vazzana, C.; Viggiani, P. Frequency Distribution of Weed Counts and Applicability of a Sequential Sampling Method to Integrated Weed Management. *Weed Research* **1992**, *32*, 39-44.

6. Binns, D. Sequential Estimation of the Mean of a Negative Binomial Distribution. *Biometrika* **1975**, *62*, 433-440.

7. Bliss, C.L.; Owen, A.R.C. Negative Binomial Distributions with a Common k. *Biometrika* **1958**, *45*, 37-38.

8. Ghosh, M.; Mukhopadhyay, N. Consistency and Asymptotic Efficiency of Two-Stage and Sequential Estimation Procedures. *Sankhya, Ser. A* **1981**, *43*, 220-227.

9. Ghosh, M.; Mukhopadhyay, N.; Sen, P.K. *Sequential Estimation*; Wiley: New York, **1997**.

10. Johnson, G.; Mortensen, D.A.; Young, L.J.; Martin, A.R. The Stability of Weed Seedling Population Models and Parameters in Eastern Nebraska Corn (Zea mays) and Soybean (Glycine max) Fields. *Weed Science* **1995**, *43*, 604-611.

11. Johnson, N.L.; Kotz, S. *Distributions in Statistics*: *Discrete Distributions*; Wiley: New York, **1969**.

12. Kuno, E. A New Method of Sequential Sampling to Obtain Population Estimates with a Fixed Level of Accuracy. *Res. Pop. Ecol.* **1969**, *11*, 127-136.

13. Kuno, E. Some Notes on Population Estimation by Sequential Sampling. *Res. Pop. Ecol.* **1972**, *14*, 58-73.

14. Marshall, E.J.P. Field-Scale Estimates of Grass Weed Population in Arable Land. *Weed Research* **1988**, *28*, 191-198.

15. Mukhopadhyay, N. Sequential Sampling. In *The Encyclopedia of Environmetrics*, Volume 4; Shaarawi, A. H., Piegorsch, W. W., eds.; Wiley: Chichester, **2002**; pp. 1983-1988.

16. Mukhopadhyay, N.; Diaz, J. Two-Stage Sampling for Estimating the Mean of a Negative Binomial Distribution. *Sequential Anal.* **1985**, *4*, 1-18.

17. Mulekar, M.S.; Young, L.J. Approximations for a Fixed Sample Size Selection Procedure for Negative Binomial Populations. *Commun. Statist.* **1991**, *20*, 1767-1776.

18. Mulekar, M.S.; Young, L.J. Sequential Estimation in the Agricultural Sciences. In *Applied Sequential Methodologies*; Mukhopadhyay, N., Datta, S. and Chattopadhyay, S., eds.; Marcel Dekker: New York, 2004; pp. 293-318.

19. Mulekar, M.S.; Young, L.J.; Young, J.H. Introduction to 2-SPRT for Testing Insect Population Densities. *Environ. Entomol.* **1993**, *22*, 346-351.

20. Nyrop, J.P.; Binns, M. Quantitative Methods for Designing and Analyzing Sampling Program for Use in Pest Management. In Handbook of Pest Management, Vol.2; Pimental, D. ed.; CRC Press: Boca Raton, **1991**; pp. 67-132.

21. Onsager, J.A. The Rationale of Sequential Sampling, with Emphasis on Its Use in Pest Management. Technical Bulletin 1526, Agricultural Research Service, USDA, Washington, D.C., **1976**.

22. Plant, R.E., and Wilson, L.T. A Bayesian Method for Sequential Sampling and Forecasting in Agricultural Pest Management. *Biometrics* **1985**, *41*, 203-214.

23. Sterling, W.L. Sequential Decision Plans for the Management of Cotton Anthropods in Southeast Queensland. *Austr. Jour. Ecol.* **1976**, *1*, 265-274.

24. Sylvester, E.S.; Cox, E.I. Sequential Plans for Sampling Aphids on Sugar Beets in Kern County, California. *Jour. Econ. Entomol.* **1961**, *54*, 1080-1085.

25. Waters, W.E. Sequential Analysis of Forest Insect Surveys. *Forest Sci.* **1955**, *1*, 68-79.

26. Wiles, L.J.; Oliver, G.W.; York, A.C.; Gold, H.J.; Wilkerson, G.G. Spatial Distribution of Broadleaf Weeds in North Carolina Soybean (Glycine max). *Weed Sci.* **1992**, *40*, 554-557.

27. Willson, L.J. Estimation and Testing Procedures for the Parameters of the Negative Binomial Distribution. Ph.D. thesis **1981**, Department of Statistics, Oklahoma State University, Stillwater.

28. Willson, L.J.; Folks, J.L. Sequential Estimation of the Mean of the Negative Binomial Distribution. *Sequential Anal.* **1983**, *2*, 55-70.

29. Wilson, L.T. Development of an Optimal Monitoring Program in Cotton: Emphasis on Spider Mites and Heliothis spp. *Entomophaga* **1982**, *27*, 45-50.

30. Young, L.J. Computations of Some Exact Properties of Wald's SPRT When Sampling From a Class of Discrete Distributions. *Biometr. Jour.* **1994**, *36*, 627-637.

31. Young, L.J. Sequential Testing in the Agricultural Sciences. In *Applied Sequential Methodologies*; Mukhopadhyay, N., Datta, S. and Chattopadhyay, S., eds.; Marcel Dekker: New York, 2004; pp. 381-410.

32. Zou, G. Weed Population Sequential Sampling Plan and Weed Seedling Emergence Pattern Prediction. Ph.D. thesis **1998**, Department of Plant Science, University of Connecticut, Storrs.