Ann Inst Stat Math (2018) 70:1049–1075
https://doi.org/10.1007/s10463-017-0620-2

Purely sequential bounded-risk point estimation of the negative binomial mean under various loss functions: one-sample problem

Nitis Mukhopadhyay[1] · **Sudeep R. Bapat**[1,2]

Received: 13 February 2017 / Revised: 19 July 2017 / Published online: 13 October 2017

Abstract A negative binomial (NB) distribution is useful to model over-dispersed count data arising from agriculture, health, and pest control. We design purely sequential bounded-risk methodologies to estimate an unknown NB mean $\mu(>0)$ under different forms of loss functions including customary and modified Linex loss as well as squared error loss. We handle situations when the thatch parameter $\tau(>0)$ may be assumed known or unknown. Our proposed methodologies are shown to satisfy properties including first-order asymptotic efficiency and first-order asymptotic risk efficiency. Summaries are provided from extensive sets of simulations showing encouraging performances of the proposed methodologies for small and moderate sample sizes. We follow with illustrations obtained by implementing estimation strategies using real data from statistical ecology: (1) weed count data of different species from a field in Netherlands and (2) count data of migrating woodlarks at the Hanko bird sanctuary in Finland.

Keywords Linex loss · CV approach · First-order asymptotic efficiency · First-order asymptotic risk efficiency · Migrating woodlarks data · Over-dispersed count data · Squared error loss · Statistical ecology · Weed count data · Bird sanctuary data

✉ Nitis Mukhopadhyay
nitis.mukhopadhyay@uconn.edu

Sudeep R. Bapat
sudeep.bapat@uconn.edu

1 Department of Statistics, University of Connecticut, AUST Building, U-4120, 215 Glenbrook Road, Storrs, CT 06269-4120, USA

2 Department of Statistics and Applied Probability, University of California, Santa Barbara, CA 93106-3110, USA

1 Introduction

In this paper, we develop purely sequential methodologies for estimating the mean μ of a *negative binomial* (NB) distribution under various forms of the loss function. A NB distribution has been used widely to model over-dispersed count data in ecology and agriculture. Anscombe (1949, 1950) emphasized the role of NB modeling in the case of insect count data in entomological research.

We will work with a NB parametrized by Anscombe (1949) where one assumes the following *probability mass function* (p.m.f.):

$$f(x;\mu,\tau) \equiv P_{\mu,\tau}(X=x) = \left(1+\frac{\mu}{\tau}\right)^{-\tau} \frac{\Gamma(\tau+x)}{x!\Gamma(\tau)} \left(\frac{\mu}{\mu+\tau}\right)^{x}, x = 0, 1, 2, \ldots \tag{1}$$

Here, the response variable X may stand for the count of insects on plants or count of a particular variety of weed in an agricultural plot.

The probability model (1) is referred to as a NB distribution involving two parameters abbreviated as NB(μ,τ) where $0 < \mu, \tau < \infty$. This notation pretends that μ, τ are both unknown. If τ is known, then we will interpret $f(x;\mu,\tau)$ as $f(x;\mu) \equiv P_{\mu}(X=x)$. In what follows, we use the notation I_A or $I(A)$ interchangeably to denote the indicator function of set A.

The mean parameter μ is unknown, but τ may or may not be known. We have included both situations where τ is assumed known (Sect. 2) or τ is assumed unknown (Sect. 3).

The parameter μ used in (1) is interpreted as the average insect count or the average number of weed per sampling unit, whereas τ indicates the degree of clumping or thatching of infestation per sampling unit. The mean and variance for the distribution (1) are given by:

$$E_{\mu,\tau}[X] \equiv \mu \quad \text{and} \quad V_{\mu,\tau}[X] \equiv \sigma^2 = \mu + \tau^{-1}\mu^2. \tag{2}$$

Again, this notation pretends that μ, τ are both unknown. If τ is known, then we will interpret $E_{\mu,\tau}[.]$ and $V_{\mu,\tau}[.]$ as $E_{\mu}[.]$ and $V_{\mu}[.]$, respectively.

Some selected references interfacing a NB model and sequential and/or multistage sampling strategies in agriculture and biology include: Bliss and Owen (1958), Kuno (1972), Mukhopadhyay (2002), Mukhopadhyay and de Silva (2005), Mukhopadhyay and Banerjee (2014, 2015), and Banerjee and Mukhopadhyay (2016). For a general overview in the broad area of sequential and multistage inference methodologies, one may refer to Mukhopadhyay and Solanky (1994), Ghosh et al. (1997), Mukhopadhyay et al. (2004), Mukhopadhyay and de Silva (2009), Zacks (2009), and other sources.

Willson and Folks (1983) and Willson et al. (1984) contained a wide variety of sequential problems arising from estimation of a NB mean. One may refer to Mukhopadhyay and Diaz (1985) for an overview of a two-stage point estimation problem. Mukhopadhyay and Banerjee (2014, 2015) included an extensive set of literature review. A majority of sources had dealt with problems of sequential fixed-width or fixed-accuracy confidence intervals, point estimation under a *squared error loss* (SEL), or tests for μ.

A customary Linex loss, introduced originally by Varian (1975), combined a linear component with an exponential component. Such an asymmetric loss has been widely used in situations where one assigns unequal penalties for over-estimation and under-estimation. Varian's (1975) Linex loss in estimating a generic parameter θ with $\widehat{\theta}_n$ is defined as follows:

$$L_n \equiv L_n\left(\widehat{\theta}_n, \theta\right) = \exp\left\{a(\widehat{\theta}_n - \theta)\right\} - a(\widehat{\theta}_n - \theta) - 1, a \in R. \tag{3}$$

One may refer to Zellner (1986) and Chattopadhyay (1998, 2000) to gain a broader perspective. One may additionally refer to Mukhopadhyay and Bapat (2016a, b) for further practical applications under appropriate modifications of (3) under negative exponential models. In this paper, we develop only purely sequential methodologies for bounded-risk point estimation. Next, we describe the layout of this paper.

1.1 Layout of this paper

Let us provide a precise outline of this paper for a clear and crisp road-map. In all sections, our primary goal is estimation of unknown μ with a preassigned risk-bound $\omega(> 0)$.

(a) Section 2 assumes τ known:
- Section 2.1 develops an estimation strategy under modified Linex loss defined in (5).
- Section 2.2 develops an estimation strategy under squared error loss (SEL) defined in (32).
- Both subsections show first-order asymptotic efficiency and risk efficiency properties (Theorems 1, 2).

(b) Section 3 assumes τ unknown:
- Section 3.1 develops an estimation strategy under squared error loss (SEL) defined in (51).
- Section 3.2 develops an estimation strategy under customary Linex loss defined in (56).
- Both subsections derive first-order asymptotic efficiency and risk efficiency properties (Theorems 3, 4).

(c) Section 4 summarizes performances obtained from simulations. All requisite computer programs were prepared using our own R (R Core Team 2014) codes:
- Sections 4.1, 4.2 address the two problems from Sects. 2.1, 2.2 (τ known).
- Sections 4.3, 4.4 address the two problems from Sects. 3.1, 3.2 (τ unknown).

(d) Section 5 summarizes performances obtained from real data applications from ecology:
- Sections 5.1, 5.2 address the problems from Sects. 2.1, 2.2 (τ known).
- Sections 5.3, 5.4 address the problems from Sects. 3.1, 3.2 (τ unknown).

(e) Section 6 gives some brief concluding thoughts.

2 The thatch parameter τ is known

This section develops purely sequential bounded risk estimation methodologies for μ when the thatch parameter τ is assumed known under two different formulations. Specifically, Sects. 2.1 and 2.2 formulate and investigate appropriated approaches associated with (i) modified Linex loss under CV and (ii) squared error loss, respectively.

2.1 Modified Linex loss under CV approach

In this section, we introduce an appropriate modification to the conventional form of Linex loss shown in (3). We assume that we have available a sequence $\{X_1, \ldots, X_n, \ldots\}$ of *independent and identically distributed* (i.i.d.) random variables from a NB(μ, τ) population. Then, we develop a purely sequential bounded-risk methodology to estimate the NB mean μ under modified Linex loss (5) when τ is assumed known.

2.1.1 A modified Linex loss

We revisit the *coefficient of variation* (CV) approach originated by Willson and Folks (1983) which was further developed by Mukhopadhyay and Diaz (1985) and Mukhopadhyay and de Silva (2005). The customary Linex loss from (3) may be mildly modified by taking the CV approach into account. We may let:

$$W_n^* \equiv W_n^*\left(\widehat{\theta}_n, \theta\right) = \exp\left\{\frac{a(\widehat{\theta}_n - \theta)}{\theta}\right\} - \frac{a(\widehat{\theta}_n - \theta)}{\theta} - 1, a \in R, \tag{4}$$

pretending that the unknown θ parameter is non-zero and it is estimated by a *generic* estimator $\widehat{\theta}_n$.

Having recorded $X_1, \ldots, X_n$, we denote the sample mean $\overline{X}_n = n^{-1}\Sigma_{i=1}^n X_i$ which estimates μ and in the light of (4), we propose a modified Linex loss function for the point estimation problem on hand as follows. We define:

$$L_n \equiv L_n\left(\overline{X}_n, \mu\right) = \exp\left\{\frac{a(\overline{X}_n - \mu)}{\mu}\right\} - \frac{a(\overline{X}_n - \mu)}{\mu} - 1, \; a \in R. \tag{5}$$

Next, we express the risk function as follows:

$$\begin{aligned} E_\mu[L_n] &= E_\mu\left[\exp\left\{\frac{a(\overline{X}_n - \mu)}{\mu}\right\}\right] - E_\mu\left[a\left(\frac{\overline{X}_n - \mu}{\mu}\right)\right] - 1 \\ &= e^{-a}\left[1 + \frac{\mu}{\tau}(1 - e^t)\right]^{-n\tau} - 1, \text{using m.g.f. of } \overline{X}_n \\ &= \exp\left(\frac{a^2}{2n\mu}\right) + \frac{a^2}{2n\tau} + o\left(\frac{1}{n}\right) - 1. \end{aligned} \tag{6}$$

Thus, from (6), the risk associated with the modified Linex loss function (5) reduces to:

$$R_n \equiv E_\mu[L_n] = \frac{a^2}{2n}\left(\tau^{-1} + \mu^{-1}\right) + o(n^{-1}). \tag{7}$$

2.1.2 Sequential bounded risk estimation

The idea here is to bound the risk R_n given in (7) from above by a suitable constant, namely the risk-bound, $\omega(> 0)$ for all μ. This leads us to obtain the optimal fixed sample size n^* approximately as follows:

$$n \geq \frac{a^2}{2\omega}\left\{\tau^{-1} + \mu^{-1}\right\} = \frac{a^2}{2\omega}\frac{\sigma^2}{\mu^2} = n^*, \text{ say.} \tag{8}$$

The magnitude of n^* remains unknown even though its expression is given by (8). Hence, we resort to developing a purely sequential bounded risk estimation strategy next.

Now, Looking back at the expression of n^* from (8), we see clearly that

$$n^* > \frac{a^2}{2\omega\tau}, \tag{9}$$

and hence, we determine the pilot size m as follows: Let

$$m \equiv m(\omega) = \left\lfloor \frac{a^2}{2\omega\tau} \right\rfloor + 1, \tag{10}$$

where $\lfloor u \rfloor$ denotes the largest integer $< u(> 0)$ and we first gather pilot data $X_i, i = 1, \ldots, m$.

After pilot data, we gather one additional observation at-a-time, as needed, according to the stopping rule that is defined next. Since $\overline{X}_n$ may be zero with a positive probability, whatever be n, we fix a number $\gamma(> \frac{1}{2})$ and define:

$$N = \inf\left\{n \geq m : n \geq \frac{a^2}{2\omega}\left[\tau^{-1} + \left(\overline{X}_n + n^{-\gamma}\right)^{-1}\right]\right\}. \tag{11}$$

Here, $\overline{X}_n + n^{-\gamma}$ is an estimator of μ that is positive with (P_μ) probability one and that the classical *central limit theorem* (CLT) will remain in effect for $n^{1/2}\left(\overline{X}_n + n^{-\gamma} - \mu\right)$ since $\gamma > \frac{1}{2}$. At termination, based on the fully gathered data $\{N, X_1, \ldots, X_m, \ldots, X_N\}$, we propose to estimate μ by the sample mean $\overline{X}_N$.

Now, we prove a lemma which will be useful in the sequel in proving asymptotic risk efficiency property for the estimation strategy $(N, \overline{X}_N)$.

Lemma 1 *For the estimation strategy $(N, \overline{X}_N)$ defined via (11), for each fixed $\mu \in R^+, \tau \in R^+, \gamma > \frac{1}{2}$, and $s \in R^+$, we have as $\omega \to 0$:*

$$E_\mu\left[\left|\overline{X}_N - \mu\right|^s\right] \to 0.$$

Proof Consider s fixed. We begin with the following inequality:

$$0 \le \overline{X}_N \le \sup_{n\ge1}\overline{X}_n = W, \text{ say,} \tag{12}$$

where certainly all positive powers of W are integrable in view of Wiener's (1939) ergodic theorem because all positive moments of the X's are finite. Obviously, $\overline{X}_N \to \mu$ in probability (P_μ) as $\omega\to0$. Hence, we can claim:

$$E_\mu\left[\overline{X}_N^k\right] \to \mu^k, \tag{13}$$

since $\overline{X}_N^k$ is uniformly integrable in view of (12), for all $k > 0$.

Now, for a fixed positive integer $r(> s)$, we apply Jensen's inequality to write:

$$E_\mu\left[\left|\overline{X}_N - \mu\right|^r\right] = \mu^r E_\mu\left[\left|\frac{\overline{X}_N}{\mu} - 1\right|^r\right] \le \mu^r E_\mu^{1/2}\left[\left|\frac{\overline{X}_N}{\mu} - 1\right|^{2r}\right]. \tag{14}$$

But, since $2r$ is an even positive integer, we can express:

$$E_\mu\left[\left|\frac{\overline{X}_N}{\mu} - 1\right|^{2r}\right] = \Sigma_{i=0}^{2r}(-1)^i\binom{2r}{i}E_\mu\left[\left(\frac{\overline{X}_N}{\mu}\right)^{2r-i}\right] < \infty, \tag{15}$$

in view of binomial theorem and (13). Combining (14), (15), we can immediately claim uniform integrability of $\left|\overline{X}_N - \mu\right|^s$ which completes the proof. □

We now establish a number of attractive first-order asymptotic properties in Theorem 1 for the proposed purely sequential estimation strategy $(N, \overline{X}_N)$. Their proofs are outlined in Sect. 2.1.3.

Theorem 1 *With loss function L_N, pilot size m, and terminal sample size N defined in* (5), (10) *and* (11)*, respectively, under the purely sequential estimation rule $(N, \overline{X}_N)$ from* (11)*, for each fixed $\mu \in R^+$ and $\tau \in R^+$ we have as $\omega\to0$:*

(i) $N/n^* \overset{P_\mu}{\to} 1$ *if* $\gamma > \frac{1}{2}$*;*
(ii) $E_\mu\left[(N/n^*)^s\right]\to1$ *for all* s*, if* $\gamma > \frac{1}{2}$ *(or* > 1*) when* $s <$ *(or* $>$*)*0 *[asymptotic first-order efficiency];*
(iii) $E_\mu\left[L_N\right]/\omega \to 1$ *if* $\gamma > 1$ *[asymptotic risk efficiency];*

where n^* comes from (8).

Part (ii) shows that the sequential methodology (11) is asymptotically efficient in the sense of Chow and Robbins (1965) and asymptotically first-order efficient in the sense of Ghosh and Mukhopadhyay (1981). That is, we may expect the terminal sample size N to hover around the optimal fixed sample size n^* when n^* is large. Part (iii) shows that the achieved risk $E_\mu\left[L_N\right]$ may be expected to hover around the preassigned risk-bound ω when n^* is large.

2.1.3 Proof of Theorem 1

Part (*i*):

From (11), we get the following inequality:

$$\begin{aligned} &\frac{a^2}{2\omega}\left\{\left(\overline{X}_N + N^{-\gamma}\right)^{-1} + \tau^{-1}\right\} \\ &\quad \le N \le \frac{a^2}{2\omega}\left\{\left(\overline{X}_{N-1} + (N-1)^{-\gamma}\right)^{-1} + \tau^{-1}\right\} I(N > m) \\ &\quad\quad + (m-1)I(N=m) + 1 \quad \text{w.p.1}(P_\mu). \end{aligned} \tag{16}$$

Dividing (16) throughout by n^* we get:

$$\begin{aligned} &\left\{\left(\overline{X}_N + N^{-\gamma}\right)^{-1} + \tau^{-1}\right\}(\mu^{-1} + \tau^{-1})^{-1} \le N/n^* \\ &\quad \le \left\{\left(\overline{X}_{N-1} + (N-1)^{-\gamma}\right)^{-1} + \tau^{-1}\right\} \\ &\quad\quad \times (\mu^{-1} + \tau^{-1})^{-1} + (m-1)n^{*-1} I(N=m) \\ &\quad\quad + \frac{2\omega}{a^2}(\mu^{-1} + \tau^{-1})^{-1} \quad \text{w.p.1}(P_\mu). \end{aligned} \tag{17}$$

Next, taking limits on all sides of (17) and noting the following facts: $N \to \infty$ w.p.1(P_μ), $\overline{X}_N \to \mu$ w.p.1(P_μ), $m/n^* = O(1)$, and $P_\mu(N = m) \to 0$ as $\omega \to 0$, completes the proof of Part (i).

Part (*ii*):

Case 1: $s < 0$.
From (11) we get the following inequality w.p.1(P_μ):

$$\begin{aligned} &\frac{N}{n^*} \ge \frac{1}{n^*}\frac{a^2}{2\omega}\left(\frac{1}{\overline{X}_N + N^{-\gamma}}\right) = \frac{\mu^2}{\sigma^2}\left(\frac{1}{\overline{X}_N + N^{-\gamma}}\right) \\ &\quad \Rightarrow \frac{n^*}{N} \le \frac{\sigma^2}{\mu^2}\left(\overline{X}_N + N^{-\gamma}\right) \le \frac{\sigma^2}{\mu^2}\sup_{n\ge 1}\left(\overline{X}_n + 1\right) = W, \text{ say}, \Rightarrow \left(\frac{N}{n^*}\right)^s \le W^s. \end{aligned}$$

But, again W^s is clearly integrable in view of Wiener's (1939) ergodic theorem. Thus, using part (i), we conclude:

$$E_\mu\left[\left(\frac{N}{n^*}\right)^s\right] \to 1 \text{ as } \omega \to 0 \text{ when } s < 0. \tag{18}$$

Note that $\gamma > \frac{1}{2}$ suffices when $s < 0$.

Case 2: $s > 0$.
This part follows along the lines of Willson and Folks (1983) who improvised upon some of the original techniques from Mukhopadhyay (1974). Accordingly, we first fix

some arbitrarily small $\epsilon > 0$ and define $\beta = (1+\epsilon)^{1/s} n^*$. We tacitly disregard that β may not be an integer. Then, we may write:

$$E_\mu[N^s] \le (\beta+1)^s n^{*s} P_\mu(N \le \beta+1) + T(\beta), \text{ say},$$

where $T(\beta) = \Sigma_{n>\beta+1}^{\infty} n^s P_\mu(N=n)$.

Next, using the moment generating function of $\Sigma_{i=1}^{n} X_i$, which is distributed as $NB(n\mu, n\tau)$, we show that $T(\beta) \le \Sigma_{n=1}^{\infty} d_n$ where $d_n > 0$ and $d_n^{1/n} \to d, 0 < d < 1$, as $n \to \infty$. Hence, we can claim:

$$\limsup_{\omega \to 0} E_\mu\left[\left(N/n^*\right)^s\right] \le 1+\epsilon, \tag{19}$$

but $\epsilon(>0)$ is arbitrary. Also, from part (i) and Fatou's lemma, we have:

$$\liminf_{\omega\to 0} E_{\mu,\tau}\left[\left(N/n^*\right)^s\right] \ge E_{\mu,\tau}\left[\liminf_{\omega\to 0}\left(N/n^*\right)^s\right] = 1. \tag{20}$$

Essential details are included in Bapat (2017). Now, combining (18)–(20) completes the proof of Part (ii).

Part (iii):

We will improvise upon some of the techniques developed recently by Mukhopadhyay and Zacks (2017). For a clear presentation, we split the proof into a number of (main) steps as follows:

Step 1:

From (5), we express:

$$\begin{aligned}
&\omega^{-1} E_\mu[L_N] \\
&\quad = \frac{n^*}{\sigma^2} E_\mu\left[(\overline{X}_N - \mu)^2\right] + \frac{a}{3\mu} E_\mu\left[\frac{n^*}{\sigma^2}(\overline{X}_N-\mu)^3 e^{\frac{a}{\mu}\xi_N}\right] \\
&\quad = \frac{n^*}{\sigma^2} E_\mu[I_1] + \frac{a}{3\mu\sigma^2} n^* E_\mu[I_2], \text{ say},
\end{aligned} \tag{21}$$

where ξ_N is a random variable between 0 and $(\overline{X}_N - \mu)$.

Step 2:

We first address the term $n^* E_\mu[I_2]$ from (21) and show that it is $o(1)$. From Anscombe's (1952) *random CLT*, with $W_N = \Sigma_{i=1}^{N} X_i$, we can claim:

$$U \equiv U_N = \frac{W_N - N\mu}{\sigma\sqrt{n^*}} \xrightarrow{£} N(0,1), \text{ as } \omega \to 0. \tag{22}$$

With U from (22), we may express $n^* I_2$ as follows w.p.1(P_μ):

$$\frac{n^{*5/2}}{N^3}\sigma^3 U^3 e^{\frac{a}{\mu}\xi_N} = \sigma^3 e^{\frac{a}{\mu}\xi_N}\left(\frac{n^*}{N}\right)^3 n^{*-1/2} U^3. \tag{23}$$

Case 1: On the set where $\frac{a}{\mu}\xi_N < 0$, (23) gives w.p.1(P_μ):

$$n^* |I_2| I\left(\frac{a}{\mu}\xi_N < 0\right) \le \sigma^3 \left(\frac{n^*}{N}\right)^3 n^{*-1/2} |U|^3. \tag{24}$$

From Theorem 2 of Chow et al. (1979), we can claim: $|U|^s$ is uniformly integrable for all $s > 0$ when $\gamma > 1$. Also, from Theorem 1, part (ii), it follows that $(n^*/N)^s$ is uniformly integrable for fixed $s > 0$ when $\gamma > \frac{1}{2}$. Using Cauchy-Schwartz inequality, we claim:

$$E_\mu\left[\left|\left(\frac{n^*}{N}\right)^3 U^3\right|\right] \le E_\mu^{1/2}\left[\left(\frac{n^*}{N}\right)^6\right] E_\mu^{1/2}\left[\left|U^6\right|\right] = O(1),$$

which combined with (24) shows (when $\gamma > 1$):

$$E_\mu\left[n^* |I_2| I\left(\frac{a}{\mu}\xi_N < 0\right)\right] = o(1). \tag{25}$$

Case 2: On the set where $\frac{a}{\mu}\xi_N \ge 0$, (23) gives w.p.1(P_μ):

$$n^* |I_2| I\left(\frac{a}{\mu}\xi_N \ge 0\right) = \sigma^3 e^{\frac{a}{\mu}\xi_N} \left(\frac{n^*}{N}\right)^3 n^{*^{-1/2}} |U|^3 I\left(\frac{a}{\mu}\xi_N \ge 0\right), \tag{26}$$

and we will now show that

$$E_\mu\left[\left(\frac{n^*}{N}\right)^3 e^{\frac{a}{\mu}\xi_N} |U|^3 I\left(\frac{a}{\mu}\xi_N \ge 0\right)\right] = O(1).$$

Upon repeated uses of Holder's inequality to split the expectations of various terms within (26), with appropriate choices of $\alpha > 1, \beta > 1, \alpha^{-1} + \beta^{-1} = 1$ and $\alpha' > 1, \beta' > 1,\ \alpha'^{-1} + \beta'^{-1} = 1$, we claim (when $\gamma > 1$):

$$E_\mu\left[\left(\frac{n^*}{N}\right)^3 e^{\frac{a}{\mu}\xi_N} |U|^3 I\left(\frac{a}{\mu}\xi_N \ge 0\right)\right] \le O(1) E_\mu^{1/\beta\alpha'}\left[e^{\frac{a}{\mu}\beta\alpha'\xi_N} I\left(\frac{a}{\mu}\xi_N \ge 0\right)\right]. \tag{27}$$

Next, we verify:

$$E_\mu\left[e^{\lambda\xi_N} I\left(\frac{a}{\mu}\xi_N \ge 0\right)\right] = O(1),$$

where we see $\lambda = \frac{a}{\mu}\beta\alpha'$ from (27) and thus write w.p.1(P_μ):

$$e^{\lambda\xi_N} I\left(\frac{a}{\mu}\xi_N \ge 0\right) \le e^{\left|\lambda\left(\overline{X}_N - \mu\right)\right|} = \Sigma_{s=0}^{\infty} \frac{1}{s!}\left|\lambda\left(\overline{X}_N - \mu\right)\right|^s, \tag{28}$$

with each term positive and integrable. So, by applying the monotone convergence theorem, we have from (28):

$$E_\mu\left[e^{\lambda\xi_N} I\left(\frac{a}{\mu}\xi_N \geq 0\right)\right] \leq \Sigma_{s=0}^{\infty}\frac{1}{s!}E_\mu\left[\left|\lambda\left(\overline{X}_N-\mu\right)\right|^s\right]. \tag{29}$$

Then, after using Lemma 1, from (29), we can claim (when $\gamma > 1$):

$$\lim_{\omega\to 0} E_\mu\left[e^{\lambda\xi_N} I\left(\frac{a}{\mu}\xi_N \geq 0\right)\right] = 0. \tag{30}$$

A combination of (23), (25), and (30) leads us to conclude (when $\gamma > 1$):

$$n^* E_\mu[I_2] = o(1). \tag{31}$$

Step 3:
Next, we go back to the term $\frac{n^*}{\sigma^2}E_\mu[I_1]$ from (21), and show that it converges to 1 as $\omega \to 0$. Upon improvising the proof found in Mukhopadhyay (1978), a precursor of Ghosh and Mukhopadhyay (1979), we can conclude that $\lim_{\omega\to 0}\frac{n^*}{\sigma^2}E_\mu[I_1] = 1$. This completes the proof of Part (iii). □

Remark 1 It will be a fair question to ask whether one could develop a purely sequential bounded risk approach under customary Linex loss in the spirit of (3). Yes, that is certainly possible. In Sect. 4 of the original version of this paper, we had included such details. In this present version, however, we have omitted that for brevity. Much related details are still available in Bapat (2017).

2.2 Squared error loss approach

In this section, we develop a purely sequential estimation strategy for estimating the mean of a NB(μ, τ) population under *squared error loss* (SEL). Having recorded $X_1, \ldots, X_n$, recall that the sample mean $\overline{X}_n = n^{-1}\Sigma_{i=1}^n X_i$ estimates μ and we propose SEL:

$$L_n \equiv L_n(\overline{X}_n, \mu) = b(\overline{X}_n - \mu)^2,\ b > 0. \tag{32}$$

We express the risk function associated with (32) as:

$$R_n \equiv E_\mu[L_n] = \frac{b}{n}\left(\mu + \tau^{-1}\mu^2\right). \tag{33}$$

2.2.1 Sequential bounded risk estimation

We will again bound the risk R_n given in (33) from above by $\omega(> 0)$. This leads us to the optimal fixed sample size n^* approximately as follows:

$$n \geq \frac{b}{\omega}\left(\mu + \tau^{-1}\mu^2\right) = n^*,\ \text{say}. \tag{34}$$

The magnitude of n^* remains unknown even though its expression is in (34). Again, we resort to developing a purely sequential estimation strategy next.

A major difference between (8) and (34) is the fact that we do not have a known and positive lower bound along the line of (9) for n^* defined in (34). Thus, we first fix $m(\geq 1)$ and gather pilot data $X_i, i = 1, \ldots, m$ of size m from the NB population and define our stopping variable as:

$$N = \inf\left\{n \geq m : n \geq \frac{b}{\omega}\left[\overline{X}_n + \tau^{-1}\overline{X}_n^2 + n^{-\gamma}\right]\right\}, \tag{35}$$

where $\gamma(> \frac{1}{2})$ is fixed. Based on full data $\{N, X_1, \ldots, X_m, \ldots, X_N\}$, we propose to estimate μ by the sample mean $\overline{X}_N$.

Theorem 2 shows a set of attractive first-order asymptotic properties for the proposed purely sequential estimation methodology $(N, \overline{X}_N)$ obtained from (35). Interpretations of these results stay similar to those explained under Theorem 1.

Theorem 2 *With loss function L_N and terminal sample size N defined in* (32) *and* (35)*, respectively, under the purely sequential estimation rule $(N, \overline{X}_N)$ from* (35)*, for each fixed $\mu \in R^+$ and $\tau \in R^+$ we have as $\omega \to 0$:*

(i) $N/n^* \xrightarrow{P_\mu} 1$ *if* $\gamma > \frac{1}{2}$;
(ii) $E_\mu\left[(N/n^*)^s\right] \to 1$ *for all* $s > 0$, *if* $\gamma > \frac{1}{2}$ *[asymptotic first-order efficiency];*
(iii) $E_\mu[L_N]/\omega \to 1$ *if* $\gamma > 1$ *[asymptotic risk efficiency]*;

where n^* comes from (34).

2.2.2 Proof of Theorem 2

We may construct a proof of Part (i) along the line of our proof of Part (i) in Theorem 1. Hence, we omit it.

Part (*ii*):

From (35), we express the following inequality (for sufficiently large n^*) w.p.1(P_μ):

$$N/n^* \leq \left(\overline{X}_{N-1} + \tau^{-1}\overline{X}_{N-1}^2 + (m-1)^{-\gamma}\right)(\mu + \tau^{-1}\mu^2)^{-1} + m. \tag{36}$$

Now, denoting $\sup_{n\geq 2}\left(\overline{X}_n + \tau^{-1}\overline{X}_n^2\right)$ as W we can claim w.p.1(P_μ):

$$N/n^* \leq (\mu + \tau^{-1}\mu^2)^{-1}\{W + 1\} + m. \tag{37}$$

The right-hand side of (37) is free from ω and using Wiener's (1939) ergodic theorem we can claim the uniform integrability of all positive powers of N/n^*. Next, appealing to Part (i), we complete the proof. Here, $\gamma > \frac{1}{2}$ suffices.

Part (*iii*): This proof is split into a number of steps for clarity by improvising on the techniques that were originally developed by Ghosh and Mukhopadhyay (1979) and then moved further along by Sen and Ghosh (1981). Throughout, we fix an arbitrary ϵ in (0, 1).

Step 1: We note:

$$E_\mu[L_N]/\omega = \frac{b}{\omega}E_\mu\left[(\overline{X}_N-\mu)^2\right] = \frac{n^*}{\sigma^2}E_\mu\left[(\overline{X}_N-\mu)^2\right].$$

Now, we need to verify that $\frac{n^*}{\sigma^2}E_\mu\left[\left(\overline{X}_N-\mu\right)^2\right] \to 1$ as $\omega \to 0$ when $\gamma > 1$. This can be shown as follows:

In the spirit of (21), we express:

$$\frac{n^*}{\sigma^2}E_\mu\left[\left(\overline{X}_N-\mu\right)^2\right] = E_\mu\left[U_N^2\right]+E_\mu\left[U_N^2\left\{\frac{{n^*}^2}{N^2}-1\right\}\right] = E_\mu\left[I_1\right]+E_\mu\left[I_2\right], \text{ say,} \tag{38}$$

where

$$U_N = \frac{W_N - N\mu}{\sigma\sqrt{n^*}} \text{ with } W_N = \Sigma_{i=1}^N X_i \text{ and } N \text{ comes from (35).}$$

In a straightforward fashion, we can claim that $E_\mu\left[I_1\right] \to 1$ as $\omega \to 0$ when $\gamma > \frac{1}{2}$. Next, we have to verify:

$$E_\mu\left[I_{12}\right] \to 0 \quad \text{as} \quad \omega \to 0. \tag{39}$$

Step 2:
On the set $|N-n^*| \le \epsilon n^*$, we can express:

$$\frac{1}{1+\epsilon} \le \frac{n^*}{N} \le \frac{1}{1-\epsilon} \Rightarrow \frac{-\epsilon(2+\epsilon)}{(1+\epsilon)^2} \le \left(\frac{{n^*}^2}{N^2}-1\right) \le \frac{\epsilon(2+\epsilon)}{(1-\epsilon)^2},$$

which shows:

$$\left|\left(\frac{n^*}{N}\right)^2-1\right| I_{[|N-n^*|\le\epsilon n^*]} \le \epsilon(2+\epsilon)(1-\epsilon)^{-2}. \tag{40}$$

Step 3:
We recall that $U_N \xrightarrow{\pounds} N(0,1)$ by Anscombe's (1952) random CLT so that $U_N^2 \xrightarrow{\pounds} \chi_1^2$ as $\omega \to 0$, and since $E_\mu\left[U_N^2\right] \to 1$ as $\omega \to 0$, we claim that U_N^2 is uniformly integrable. Then, in view of (38), we have w.p.1(P_μ):

$$U_N^2\left|\left(\frac{n^*}{N}\right)^2-1\right| I_{[|N-n^*|\le\epsilon n^*]} \le \epsilon(2+\epsilon)(1-\epsilon)^{-2}U_N^2.$$

Thus, $U_N^2 \left| \left(\frac{n^*}{N}\right)^2 - 1 \right| I_{[|N-n^*| \le \epsilon n^*]}$ is also uniformly integrable when $\gamma > \frac{1}{2}$. In view of Part (i), we have:

$$U_N^2 \left| \left(\frac{n^*}{N}\right)^2 - 1 \right| I_{[|N-n^*| \le \epsilon n^*]} \to 0 \text{ in probability } (P_\mu) \text{ as } \omega \to 0,$$

so that we have:

$$E_\mu \left[U_N^2 \left| \left(\frac{n^*}{N}\right)^2 - 1 \right| I_{[|N-n^*| \le \epsilon n^*]} \right] \to 0 \text{ as } \omega \to 0. \tag{41}$$

Step 4:
Next, on the set $N > n^*(1+\epsilon)$, we observe $\left| \left(\frac{n^*}{N}\right)^2 - 1 \right| < 1$ so that we can express

$$U_N^2 \left| \frac{{n^*}^2}{N^2} - 1 \right| I_{[N-n^* > \epsilon n^*]} \le U_N^2.$$

Thus, $U_N^2 \left| \left(\frac{n^*}{N}\right)^2 - 1 \right| I_{[N-n^* > \epsilon n^*]}$ is uniformly integrable when $\gamma > \frac{1}{2}$. In view of Part (i), we have:

$$U_N^2 \left| \left(\frac{n^*}{N}\right)^2 - 1 \right| I_{[N-n^* > \epsilon n^*]} \to 0 \text{ in probability } (P_\mu) \text{ as } \omega \to 0,$$

so that we conclude:

$$E_\mu \left[U_N^2 \left| \left(\frac{n^*}{N}\right)^2 - 1 \right| I_{[N-n^* > \epsilon n^*]} \right] \to 0 \text{ as } \omega \to 0. \tag{42}$$

Step 5:
Next, we outline a proof of the following claim:

$$E_\mu \left[U_N^2 \left| \frac{{n^*}^2}{N^2} - 1 \right| I_{[N-n^* < -\epsilon n^*]} \right] \to 0 \text{ as } \omega \to 0. \tag{43}$$

Let $q(> 0)$ be a generic constant that does not involve ω. From (35), we can claim that $N \ge \frac{b}{\omega} N^{-\gamma}$ w.p.1(P_μ) so that $N \ge \left(\frac{b}{\omega}\right)^{1/(1+\gamma)} = O(n^{*1/(1+\gamma)})$ w.p.1(P_μ). This implies that $\left(\frac{n^*}{N}\right)^2 \le qn^{*2\gamma/(1+\gamma)}$ w.p.1(P_μ) and then we can write:

$$\left|E_\mu\left[U_N^2\left|\frac{n^{*2}}{N^2}-1\right|I_{[N-n^*<-\epsilon n^*]}\right]\right| \le qn^{*(2\gamma-1)/(1+\gamma)}E_\mu\left[\left(W_{N_k}-N_k\mu\right)^2 I_{[N\le k]}\right], \tag{44}$$

where we denote: $k=\lfloor n^*(1-\epsilon)\rfloor = O(n^*)$, $N_k=\min(N,k)$. In the sequel, we also use $\gamma'=E_\mu[(X_1-\mu)^3]$ and $\gamma''=E_\mu[(X_1-\mu)^4]$.

Now, using Cauchy-Schwartz inequality and Wald's fourth lemma (Theorem 7, Chow et al. (1965); Theorem 2.4.7, Ghosh et al. (1997)), we have:

$$\begin{aligned} &E_\mu\left[\left(W_{N_k}-N_k\mu\right)^2 I_{[N\le k]}\right] \\ &\quad\le \left\{6\sigma^2 E_\mu\left[N_k\left(W_{N_k}-N_k\mu\right)^2\right]+4\gamma' E_\mu\left[N_k\left(W_{N_k}-N_k\mu\right)\right]\right. \\ &\qquad\left.+\gamma'' E_\mu[N_k]\right\}^{1/2} P_\mu^{1/2}(N\le k). \end{aligned} \tag{45}$$

But, we have $N_k \le k$ w.p.1(P_μ). Next, we further use Wald's second equation (Theorem 8, Chow et al. (1965); Theorem 2.4.5, Ghosh et al. (1997)) and Cauchy-Schwartz inequality in (45) to obtain:

$$E_\mu\left[\left(W_{N_k}-N_k\mu\right)^2 I_{[N\le k]}\right] \le \left\{6\sigma^4k^2+4\gamma' k^{3/2}\sigma+\gamma'' k\right\}^{1/2} P_\mu^{1/2}(N\le k). \tag{46}$$

Step 6:
Let $h=\left\lfloor\left(\frac{b}{\omega}\right)^{1/(1+\gamma)}\right\rfloor+1$. We may pick ω small enough so that $h\le k$ where k was defined underneath (44) and write:

$$P_\mu\left\{N\le n^*(1-\epsilon)\right\} = \Sigma_{n=h}^k P_\mu(N=n) \le \Sigma_{n=h}^k P_\mu\left\{n\ge \frac{n^*}{\sigma^2}\overline{X}_n\right\}. \tag{47}$$

Now, with fixed but otherwise arbitrary $\nu(\ge 1)$, we may rewrite (47):

$$P_\mu\left\{N\le n^*(1-\epsilon)\right\} \le (\epsilon\mu)^{-\nu}\Sigma_{n=h}^k E_\mu\left[\left|\overline{X}_n-\mu\right|^{2\nu}\right] \le q\,\Sigma_{n=h}^k n^{-\nu}. \tag{48}$$

The last step in (48) follows from the lemma of Sen and Ghosh (1981). One may also refer to Lemma 9.2.3 in (Ghosh et al. (1997), pp. 275–276).

Next, whatever be $\nu(\ge 1)$, (48) leads to:

$$P_\mu\left\{N\le n^*(1-\epsilon)\right\} \le qkh^{-\nu} \le O(n^*)O(n^{*-\nu/(1+\gamma)}) = O\left(n^{*(1+\gamma-\nu)/(1+\gamma)}\right). \tag{49}$$

Now, by combining (44)–(46) with (48), (49), we obtain:

$$E_\mu\left[U_N^2\left|\frac{n^{*2}}{N^2}-1\right|I_{[N-n^*<-\epsilon n^*]}\right] = O(n^{*(3\gamma-\nu)/(1+\gamma)}), \tag{50}$$

which validates (43) for any $\gamma > 1/2$, provided that we pick $\nu > \max(2, 3\gamma)$ which is certainly possible to do. Proof of Theorem 2 is thus complete. □

Remark 2 (Negative moments of N/n^* for estimation strategy (35)). Now, we may summarize asymptotic behavior of $E_\mu\left[(N/n^*)^s\right]$ when $s < 0$. Clearly, $E_\mu\left[(N/n^*)^s I\left(N > \frac{1}{2}n^*\right)\right] \to 1$. But, then, in view of (49), we also have:

$$E_\mu\left[(N/n^*)^s I\left(N \leq \frac{1}{2}n^*\right)\right] = O\left(n^{*-s+((1+\gamma-\nu)/(1+\gamma))}\right) = o(1),$$

if we pick $\nu > \max(2, (1+\gamma)(1-s))$ which is certainly possible to do. Thus, it follows that $E_\mu\left[(N/n^*)^s\right] \to 1$ when $\gamma > \frac{1}{2}, s < 0$.

3 The thatch parameter τ is unknown

This section develops purely sequential bounded risk estimation methodologies for μ when the thatch parameter τ is assumed unknown under two different formulations. Specifically, Sects. 3.1 and 3.2 formulate and investigate appropriated approaches associated with (i) squared error loss and (ii) customary Linex loss.

3.1 Squared error loss approach

We develop a purely sequential estimation strategy for μ in a NB(μ, τ) population under SEL assuming that **both** parameters are unknown. Having recorded $X_1, \ldots, X_n$, we propose SEL as follows:

$$L_n \equiv L_n(\overline{X}_n, \mu) = b(\overline{X}_n - \mu)^2, \ b > 0. \tag{51}$$

for estimating μ with $\overline{X}_n$. The loss function (51) looks similar to (32), but τ remains unknown too.

We express the risk function as follows:

$$R_n \equiv E_{\mu,\tau}[L_n] = \frac{b}{n}\sigma^2, \tag{52}$$

parallel to (33) with σ^2 from (2).

3.1.1 Sequential bounded risk estimation

The goal is to bound the risk R_n given in (52) from above by $\omega(> 0)$. This leads to the optimal fixed sample size n^* as follows:

$$n \geq \frac{b}{\omega}\sigma^2 = n^*, \ \text{say}. \tag{53}$$

The magnitude of n^* remains unknown and thus we develop a purely sequential estimation strategy in the spirit of Chow and Robbins (1965).

We first fix $m(\geq 2)$ and begin with pilot data $X_i, i = 1, \ldots, m$ of size m from the NB population. Since σ^2 is unknown, we estimate it using the customary sample variance, $S_n^2 \equiv (n-1)^{-1}\Sigma_{i=1}^{n}(X_i - \overline{X}_n)^2, n \geq m$. However, since S_n^2 may be zero with a positive probability, whatever be n, we also fix a number $\gamma(> \frac{1}{2})$ and define a final sample size:

$$N = \inf\left\{n \geq m : n \geq \frac{b}{\omega}\left[S_n^2 + n^{-\gamma}\right]\right\}. \tag{54}$$

We ensure that our estimated variance $S_n^2 + n^{-\gamma}$ remains positive with probability 1 $(P_{\mu,\tau})$ for all $n \geq m$.

Next, based on the fully gathered data $\{N, X_1, \ldots, X_m, \ldots, X_N\}$, we propose to estimate μ by $\overline{X}_N$. The following Theorem gives a set of attractive first-order asymptotic properties for the purely sequential methodology (54). Interpretations of these results stay similar to those explained under Theorem 1.

Theorem 3 *With loss function L_N and terminal sample size N defined in (51) and (54), respectively, under the purely sequential estimation rule $(N, \overline{X}_N)$ from (54), for each fixed $\mu \in R^+$ and $\tau \in R^+$ we have as $\omega \to 0$:*

(i) $N/n^* \overset{P_{\mu,\tau}}{\to} 1$;
(ii) $E_{\mu,\tau}\left[(N/n^*)^s\right] \to 1$ *for alls[asymptotic first-order efficiency]*;
(iii) $E_{\mu,\tau}\left[L_N\right]/\omega \to 1$ *if* $\gamma > 1$*[asymptotic risk efficiency]*;

where n^* comes from (53).

3.1.2 An outline of a Proof of Theorem 3

Part (i) follows from the basic inequality:

$$\left(S_N^2 + N^{-\gamma}\right)\sigma^{-2} \leq N/n^* \leq \left(S_{N-1}^2 + (N-1)^{-\gamma}\right)\sigma^{-2} + n^{*-1} \text{ w.p.1}(P_{\mu,\tau}), \tag{55}$$

and the facts that $N \to \infty$, $S_N^2 \to \sigma^{-2}$, $S_{N-1}^2 \to \sigma^{-2}$ w.p.1$(P_{\mu,\tau})$ as $\omega \to 0$.

Next, for small enough ω, observe that the right-hand side of (55) implies w.p.1$(P_{\mu,\tau})$:

$$\begin{aligned} N/n^* &\leq \sigma^{-2}\left\{2N^{-1}\Sigma_{i=1}^{N}(X_i - \mu)^2 + 1\right\} + 1 \\ &\leq \sigma^{-2}\sup_{n\geq 2}\left\{2n^{-1}\Sigma_{i=1}^{n}(X_i - \mu)^2 + 1\right\} + 1. \end{aligned}$$

Now, when $s > 0$, Part (ii) follows in the spirit of our proofs of Theorems 1, 2. When $s < 0$, a proof in the spirit of Remark 2 can be easily put together since S_n^2 is a U-statistic. Part (iii) can be proved along the line of the proof of Theorem 2, part (iii). Details are left out for brevity. □

3.2 Customary Linex loss approach

We develop a purely sequential estimation strategy for μ in a NB(μ, τ) population under customary Linex loss in the spirit of (3). Having recorded $X_1, \ldots, X_n$, we propose the following loss:

$$L_n \equiv L_n(\overline{X}_n, \mu) = \exp\left\{a(\overline{X}_n - \mu)\right\} - a(\overline{X}_n - \mu) - 1,\ a \in R. \tag{56}$$

Next, we express the associated risk function as follows:

$$R_n \equiv E_{\mu,\tau}[L_n] = E_{\mu,\tau}\left[\exp\left\{a(\overline{X}_n - \mu)\right\} - a(\overline{X}_n - \mu) - 1\right]. \tag{57}$$

Upon simplifying the risk from (57) along the lines of (6), it reduces to:

$$R_n = \frac{a^2}{2n}\sigma^2 + o(n^{-1}). \tag{58}$$

3.2.1 Sequential bounded risk estimation

The goal is to bound the risk R_n given in (58) from above by $\omega(> 0)$. This leads to the optimal fixed sample size n^* as follows:

$$n \geq \frac{a^2}{2\omega}\sigma^2 = n^*,\ \text{say}. \tag{59}$$

Again, the magnitude of n^* remains unknown even though its expression is given by (59). Hence, we resort to developing a purely sequential bounded risk estimation strategy.

We fix $m(\geq 2)$ and gather pilot data $X_i, i = 1, \ldots, m$ of size m from the NB population. We also fix a number $\gamma(> \frac{1}{2})$ and define:

$$N = \inf\left\{n \geq m : n \geq \frac{a^2}{2\omega}\left[S_n^2 + n^{-\gamma}\right]\right\}. \tag{60}$$

Next, based on the fully gathered data $\{N, X_1, \ldots, X_m, \ldots, X_N\}$, we propose to estimate μ by the sample mean $\overline{X}_N$. Theorem 4 gives a set of attractive first-order asymptotic properties for the proposed purely sequential methodology (60). For brevity, it is stated without explicitly providing a proof. Interpretations of these results stay similar to those explained under Theorem 1.

Theorem 4 *With loss function L_N and terminal sample size N defined in* (56) *and* (60)*, respectively, under the purely sequential estimation rule*$(N, \overline{X}_N)$ *from* (60)*, for each fixed $\mu \in R^+$ and $\tau \in R^+$ we have as $\omega \to 0$:*

(i) $N/n^* \overset{P_{\mu,\tau}}{\to} 1$;

(ii) $E_{\mu,\tau}\left[(N/n^*)^s\right] \to 1$ *for all s [asymptotic first-order efficiency]*;
(iii) $E_{\mu,\tau}\left[L_N\right]/\omega \to 1$ *if* $\gamma > 1$ *[asymptotic risk efficiency]*;

where n^* *comes from* (59).

4 Summaries from simulations

In this section, we provide summaries derived from sets of simulation studies. These were run to examine performances of our proposed purely sequential estimation strategies from Sects. 2 and 3 for both small and moderate values of n^*.

4.1 Performances of estimation strategy (11) from Sect. 2.1.2: τ known

We first generated pseudorandom observations from the distribution (1) with combinations of choices for μ and τ. We fixed values $a = 1, \gamma = 1.5$, and determined m from (10). Each row in Table 1 corresponds to averages from 10,000 replications which were run under a given configuration. In order to represent varying sample sizes, we show results for fixed values of $n^* = 50$ (small), 200 (medium).

Table 1 shows n^* (column 3), ω (column 4), the estimated values $\overline{x}$ and $s_{\overline{x}}$ (column 5), estimated values $\overline{n}$, $s_{\overline{n}}$ (column 6), the ratio $\overline{n}/n^*$ (column 7), the values of $\overline{z}$ and $s_{\overline{z}}$ (column 8) and the ratio $\overline{r}/\overline{n}$ (column 9) where we denote:

$$\begin{aligned}
&N = n_i, r_i = \frac{a^2}{2n_i}\left(\frac{1}{\mu} + \frac{1}{\tau}\right) \text{as in (7) under the } i\text{th replication},\\
&\text{and } \overline{r} = H^{-1}\Sigma_{i=1}^{H} r_i, s_{\overline{r}} = \sqrt{(H^2 - H)^{-1}\Sigma_{i=1}^{H}(r_i - \overline{r})^2},\\
&\text{so that } \overline{z} = \overline{r}/\omega, s_{\overline{z}} = s_{\overline{r}}/\omega, H = 10000. \qquad (61)
\end{aligned}$$

The column 9 quantifies estimated average risk per unit average number of samples. We say a bit more about this at the end of Sect. 4.

Table 1 Simulation results from 10,000 replications for the purely sequential procedure (11) with m from (10): $a = 1, \gamma = 1.5$

μ	τ	n^*	ω	$\overline{x}$ $s_{\overline{x}}$	$\overline{n}$ $s_{\overline{n}}$	$\overline{n}/n^*$	$\overline{z}$ $s_{\overline{z}}$	$\overline{r}/\overline{n}$
2	3	50	0.0083	2.0178 0.0026	51.65 0.0683	1.0330	1.0155 9.72×10^{-6}	1.63×10^{-4}
		200	0.0020	2.0032 0.0013	201.70 0.1275	1.0085	1.0084 1.24×10^{-6}	9.99×10^{-6}
3	4	50	0.0058	3.0178 0.0032	51.21 0.0504	1.0242	1.0072 5.32×10^{-6}	1.14×10^{-4}
		200	0.0014	3.0080 0.0016	201.33 0.0973	1.0066	1.0040 6.79×10^{-7}	6.98×10^{-6}

Table 2 Simulation results from 10,000 replications for the purely sequential procedure (35): $b = 1$, $\gamma = 1.5$, $m = 5$

μ	τ	n^*	ω	$\overline{x}$ $s_{\overline{x}}$	$\overline{n}$ $s_{\overline{n}}$	$\overline{n}/n^*$	$\overline{z}$ $s_{\overline{z}}$	$\overline{r}/\overline{n}$
2	3	50	0.0667	2.0035	49.53	0.9907	1.0626	1.43×10^{-3}
				0.0026	0.0943		2.58×10^{-4}	
		200	0.0167	2.0001	199.37	0.9968	1.0148	8.50×10^{-5}
				0.0013	0.1817		5.61×10^{-5}	
3	4	50	0.1050	3.0056	49.90	0.9981	1.0307	2.10×10^{-3}
				0.0032	0.0783		2.11×10^{-4}	
		200	0.0262	3.0006	200.14	1.0007	1.0055	1.31×10^{-4}
				0.0016	0.1568		2.11×10^{-5}	

The $\overline{x}$ values are very close to the corresponding means in each case with very small standard error values $s_{\overline{x}}$. These become even closer for $n^* = 200$. The values of $\overline{n}$ seem to estimate n^* very accurately across the rows. We see that the $\overline{n}$ values overestimate n^* by a small margin. The last column shows that both sequential methodologies tend to provide a risk-bound close (or smaller) to preset goal ω for small (moderate) n^*. Clearly, the proposed strategy (11) performs very well.

4.2 Performances of estimation strategy (35) from Sect. 2.2.1: τ known

A summary is provided from simulations that were carried out rather analogously as we had explained in Sect. 4.1. One difference however is that we had to fix m arbitrarily in the present scenario. We fixed the values $b = 1$, $m = 5$, and $\gamma = 1.5$. The sets of notation used in Tables 1 and 2 are very similar with one exception: $r_i = \frac{b}{n_i}\left(\mu + \tau^{-1}\mu^2\right)$.

Table 2 shows performances similar to those summarized in Table 1. Clearly, the proposed strategy (35) performs very well.

4.3 Performances of estimation strategy (54) from Sect. 3.1.1: τ unknown

A summary is provided from simulations that were carried out rather analogously as we had explained in Sect. 4.1. One difference however is that we had to fix m arbitrarily in the present scenario. We fixed the values $a = 1$, $m = 5$, and $\gamma = 1.5$. The sets of notation used in Tables 1, 2 and 3 are also very similar with one exception: $r_i = \frac{b}{n_i}\sigma^2$.

Table 3 shows performances similar to those summarized in Tables 1 and 2. Clearly, the proposed strategy (54) performs very well.

Table 3 Simulation results from 10,000 replications for the purely sequential procedure (54): $b = 1$, $\gamma = 1.5$, $m = 5$

μ	τ	n^*	ω	$\overline{x}$ $s_{\overline{x}}$	$\overline{n}$ $s_{\overline{n}}$	$\overline{n}/n^*$	$\overline{z}$ $s_{\overline{z}}$	$\overline{r}/\overline{n}$
2	3	50	0.0667	1.9989	46.15	0.9231	1.2634	1.82×10^{-3}
				0.0029	0.1491		4.72×10^{-4}	
		200	0.0167	1.9972	197.69	0.9884	1.0358	8.74×10^{-5}
				0.0013	0.2939		2.79×10^{-5}	
3	4	50	0.1050	3.0036	46.10	0.9220	1.2731	2.89×10^{-3}
				0.0036	0.1438		8.54×10^{-4}	
		200	0.0262	2.9995	196.92	0.9846	1.0369	1.37×10^{-4}
				0.0016	0.2758		4.11×10^{-5}	

Table 4 Simulation results from 10,000 replications for the purely sequential procedure (60) with $a = 1$, $\gamma = 1.5$, $m = 5$

μ	τ	n^*	ω	$\overline{x}$ $s_{\overline{x}}$	$\overline{n}$ $s_{\overline{n}}$	$\overline{n}/n^*$	$\overline{z}$ $s_{\overline{z}}$	$\overline{r}/\overline{n}$
2	3	50	0.0333	2.0006	43.67	0.8735	1.5633	1.19×10^{-3}
				0.0032	0.1660		5.15×10^{-4}	
		200	0.0083	1.9999	195.74	0.9787	1.0545	4.47×10^{-5}
				0.0013	0.3035		4.08×10^{-5}	
3	4	50	0.0525	2.9960	43.91	0.8782	1.5894	1.90×10^{-3}
				0.0040	0.1590		9.66×10^{-4}	
		200	0.0131	3.0007	195.50	0.9775	1.0562	7.07×10^{-5}
				0.0016	0.2785		9.09×10^{-5}	

4.4 Performances of estimation strategy (60) from Sect. 3.2.1: τ unknown

A summary is provided from simulations that were carried out rather analogously as we had explained in Sect. 4.1. One difference however is that we had to fix m arbitrarily in the present scenario. We fixed the values $a = 1$, $m = 5$, and $\gamma = 1.5$. The sets of notation used in Tables 3 and 4 are similar.

Table 4 shows performances similar to those summarized in Tables 1, 2 and 3. Clearly, the proposed strategy (60) performs very well.

4.5 Which loss function to use in practice?

This is a difficult question to answer precisely and mathematically. In fact, there may not be such a resolution. Under the umbrella of statistical decision making, a loss

Table 5 Ad hoc comparisons of loss functions

μ	τ	n^*	Average risk per average sample size, $\overline{r}/\overline{n}$, values from column 9 in Tables 1, 2, 3 and 4			
			τ known		τ unknown	
			Table 1	Table 2	Table 3	Table 4
2	3	50	0.00016300	0.001430	0.0018200	0.0011900
3	4	50	0.00011400	0.002100	0.0028900	0.0019000
2	3	200	0.00000999	0.000085	0.0000874	0.0000447
3	4	200	0.00000698	0.000131	0.0001370	0.0000707

function is an essential input that must be arrived by the practitioner in the substantive field.

In the case of the four bounded-risk problems which have been under our considerations in Sects. 2 and 3, we may however attempt to suggest some ad hoc guidelines supported by the presented data analyses from Tables 1, 2, 3 and 4. Since $\overline{n}$ values are very close to the n^* values all across Tables 1, 2, 3 and 4, we suggest utilizing the information obtained from column 9 from each table regarding the achieved values $\overline{r}/\overline{n}$ which indeed estimate the "risk per unit sample size." In Table 5, we present them all in one place to facilitate comparisons in our earnest hope that a kind of data-based theme may emerge to help a practitioner.

For the range of values of μ, τ under consideration, Table 5 shows that the estimated values of "risk per unit sample size" associated with the modified or customary Linex loss come out sizably smaller than those associated with SEL. This feature holds whether the thatch parameter τ is known or unknown. Such limited empirical evidence may suggest that one could treat the modified or customary Linex loss (Sects. 2.1 and 3.2) more seriously than SEL (Sects. 2.2 and 3.1) while handling the class of problems under discussion.

4.6 How to pick the risk-bound ω?

Again, there may not be a precise mathematical approach to quantify ω, the risk-bound. In statistical decision making, a practitioner or his/her team in a substantive applied field must come up with a risk-bound that may be acceptable and relevant to them in the context of an estimation problem on hand.

The papers of Willson and Folks (1983) and Willson et al. (1984) were derived from their practical field-work on entomological studies with the Agricultural Experiment Station, Oklahoma State University. The values of μ, τ, ω used in Tables 1, 2, 3 and 4 are reasonably consistent under a number of practical scenarios summarized by them. In our view, we should not attempt to prescribe a choice of ω quantifying an acceptable error bound that is supposed to work across all problems.

For example, we may revisit SEL from (32) with $b = 1$. Now, if one knows a'priori that μ is rather small, then $\left|\overline{X}_n - \mu\right|^2$ may be expected to be much smaller in view of consistency, perhaps leading one to pick $\omega = (0.01)^2$ or even smaller than that. However, if one knows a'priori that μ is rather medium or large, then $\left|\overline{X}_n - \mu\right|^2$ may not be expected to be all that small for comparable sample sizes. In this case ω could be chosen bit larger.

The point is this: Estimating average weekly salary as opposed to annual salary in a population must be treated differently. It may suffice to estimate average weekly salary within ± 5 dollars or ± 10 dollars based on sample size n. With a comparable sample size, it may be impossible to estimate the average annual salary within ± 5 dollars. On the other hand, it may be more reasonable to estimate the average annual salary with ± 500 dollars or ± 1000 dollars bringing down the required sample size substantially.

The bottom line: A practically realistic-appropriate-useful choice of ω must be problem-dependent. Subject matter practitioners should not shy away from specifying his/her practically realistic-appropriate-useful choice of ω while handling a particular problem or scenario.

5 Real data illustrations

Sections 5.1–5.4 highlight performances of our estimation strategies using real data from statistical ecology. We have used emphasized (i) weed count data (Sects. 5.1 and 5.3) of different species from a field in Netherlands and (ii) count data of migrating woodlarks (Sects. 5.2 and 5.4) from the Hanko bird sanctuary in Finland.

5.1 Estimation strategy (11) with τ known: weed count data

We resort to dataset from ecology and make use of the weed count data presented by Heijting (2013). Heijting et al. (2007) recorded data from quadrats on part of an arable maize field in Wageningen, Netherlands, prior to herbicide application. We looked at data from the year 2001 between 18 and 21 June, and applied our methodology (11) on *Capsella bursa-pastoris L.* (Shephard's purse). The field was cultivated and sown in north-south direction and the observation area was divided into $16 \times 67 = 1072$ quadrats of 0.75×0.75 meters. Further information are available from:

http://dx.doi.org/10.17026/dans-zu9-r7y8 and Heijting et al. (2007).

The dataset consists of 1072 rows and a NB fit was seen as appropriate with a p-value of 0.84. We found $\widehat{\mu} = 0.3$ and $\widehat{\tau} = 3.98$ from full data. This dataset was treated as our population with unknown mean μ with a known value 3.98 for τ. We fixed $a = 1$ and $\gamma = 1.5$, implemented the methodology (11) by drawing observations from the full set of data without replacement. Sampling with or without replacement made practically no difference.

Table 6 shows results from implementing the estimation strategy (11) in single runs with 3 different preset values of the risk-bound ω. The terminal estimated value of μ appear close to the value of $\widehat{\mu} = 0.3$ obtained from full data. The n^* values (obtained by pretending that $\mu = 0.3$ and $\tau = 3.98$) are provided as a vehicle for ad hoc comparison with the observed n values. We have not used these n^* values in our

Table 6 Analysis of weed count data with sequential strategy (11) assuming known $\tau = 3.98$ and $a = 1, \gamma = 1.5$

n^*	ω	$\hat{\mu}: \overline{x}_n$	n	n/n^*	$\widetilde{z}$
75	0.0182	0.91	82	1.09	1.2069
200	0.0068	0.44	151	0.75	1.1339
500	0.0027	0.41	607	1.21	0.9052

Table 7 Analysis of woodlarks data with sequential strategy (35) assuming known $\tau = 0.23$ and $b = 1, \gamma = 1.5,\ m = 7$

n^*	ω	$\hat{\mu}: \overline{x}_n$	n	n/n^*	$\widetilde{z}$
30	1.4498	2.61	31	1.03	0.7490
40	1.0873	2.84	39	0.97	0.8939
50	0.8699	3.47	46	0.92	1.3950

implementation. We note that the ratio n/n^* is reasonably close to 1 which is desired. In the last column of Table 6, we show a value $\widetilde{z}$ obtained from a single run:

$$N = n, \widetilde{r} = \frac{a^2}{2n}\left(\frac{1}{\overline{x}_n} + \frac{1}{\tau}\right) \text{ in the spirit of (61) under one replication, so that } \widetilde{z} = \widetilde{r}/\omega. \tag{62}$$

Column 6 shows $\widetilde{z}$ values mainly to grasp a sense of how close the estimated risk may or may not be when compared with the preset goal ω. The erratic behavior is due to the fact that $\widetilde{z}$ values were obtained from single runs.

5.2 Estimation strategy (35) with τ known: woodlarks data

In this real data illustration, we resort to an ecological count dataset on migrating woodlarks at Hanko bird sanctuary situated in southwestern Finland. This data were used by Linden and Mantyniemi (2011) and are available in the *Ecological Archives* E092-120-S2. One could also refer to Supplement 2 in Linden and Mantyniemi (2011) for more details.

We used the migration data for Autumn season, 1 September-10 November, 2009 and worked with the daily counts of migrating birds during first four hours of daylight after sunrise. The dataset included 71 rows and a NB fit was seen as appropriate with a p-value of 0.35. We found $\widehat{\mu} = 3.05$ and $\widehat{\tau} = 0.23$ from full data which was treated as our population with unknown μ but with known value 0.23 for τ.

We picked $a = 1$ and $\gamma = 1.5$, implemented the methodology (35) by drawing observations from the full set of data without replacement. Sampling with or without replacement made practically no difference.

Table 7 shows results from implementing the estimation strategy (35).in single runs with 3 different preset values of the risk-bound ω. The terminal estimated values of μ look bit erratic and these are not too close to the value $\widehat{\mu} = 3.05$, obtained from full

Table 8 Analysis of weed count data with sequential strategy (54) assuming unknown τ and $b = 1, \gamma = 1.5, m = 10$

n^*	ω	$\hat{\mu}$: $\overline{x}_n$	n	n/n^*	$\widetilde{z}$
75	0.0148	1.06	65	0.87	1.3860
200	0.0055	0.87	224	1.12	0.8406
500	0.0022	0.92	538	1.07	0.9584

data. The n^* values (obtained by pretending that $\mu = 3.05$ and $\tau = 0.23$) are merely provided as a vehicle for ad hoc comparison with the observed n values. We have not used these n^* values in our implementation. We note that the sample sizes (Table 7) are small in the range of 30–50, however the ratio n/n^* remains close to 1 which is desired. In the last column of Table 7, we show associated value $\widetilde{z}$ obtained from a single run:

$$N = n, \widetilde{r} = \frac{b}{n}\left(\overline{x}_n + \tau^{-1}\overline{x}_n^2\right) \text{ in the spirit of (61) under}$$
$$\text{one replication, so that } \widetilde{z} = \widetilde{r}/\omega. \tag{63}$$

We observe that $\widetilde{z}$ values are smaller than one which indicates that the estimated risk is smaller than the preset goal ω.

5.3 Estimation strategy (54) with τ unknown: weed count data

We again return to use weed count data in this illustration. We applied our methodology (54) on *Polygonum aviculare L.* (knotweed) from the full set of data without replacement. A NB fit was seen as appropriate with a p-value of 0.94. We found $\widehat{\mu} = 0.91$ and $\widehat{\tau} = 4.11$ from full data.

Table 8 shows the real data illustration. The choices of γ and n^* are consistent with those in Sect. 5.1 along with $b = 1$ and $m = 10$. In column 6 of Table 8, we show a value $\widetilde{z}$ obtained from a single run:

$$N = n, \widetilde{r} = \frac{b}{n}s_n^2 \text{ in the spirit of (61)under}$$
$$\text{one replication, so that } \widetilde{z} = \widetilde{r}/\omega. \tag{64}$$

Again, column 6 shows $\widetilde{z}$ values mainly to grasp a sense of how close the estimated risk may or may not be when compared with the preset goal ω. A slight erratic behavior is due to the fact that $\widetilde{z}$ values were obtained from single runs.

5.4 Estimation strategy (60) with τ unknown: woodlarks data

We applied the methodology (60) on the daily counts of migrating birds during first four hours of daylight after sunrise by drawing observations from the full set of data without replacement. Table 9 shows this real data illustration. Column 6 of Table 9

Table 9 Analysis of woodlarks data with sequential strategy (60) assuming unknown τ and $a = 1, \gamma = 1.5,\ m = 7$

n^*	ω	$\hat{\mu}: \overline{x}_n$	n	n/n^*	$\widetilde{z}$
30	0.7249	2.30	42	1.40	0.7900
40	0.5436	3.36	52	1.30	0.9276
50	0.4349	2.63	39	0.78	0.9640

shows a value $\widetilde{z}$ obtained from a single run in the spirits of (63). The $\widetilde{z}$ values help in grasping a sense of how close the estimated risk may or may not be when compared with the preset goal ω. A slight erratic behavior is due to the fact that $\widetilde{z}$ values were obtained from single runs.

6 Concluding thoughts

In Sect. 2.1, we noted a known lower bound for n^* in (9) which led to a specific choice of m in (10). Such a lower bound for n^* was first noted by Willson and Folks (1983) in developing their purely sequential methodology. The same lower bound for n^* led Mukhopadhyay and Diaz (1985) to obtain an associated two-stage estimation methodology.

In a different but closely related route, Mukhopadhyay and Duggan (1997) developed a two-stage fixed-width confidence interval methodology for the normal mean when the unknown population variance had a known positive lower bound. They developed asymptotic second-order properties for an appropriately modified two-stage estimation methodology. Proliferation of such core ideas from Mukhopadhyay and Duggan (1997) in many directions has been widespread and continues to spread in areas including big data as well as small n large p problems. For brevity, we only mention Aoshima and Takada (2000), Mukhopadhyay and Duggan (2000, 2001), and Aoshima and Yata (2010).

We should emphasize that when μ and τ remain unknown, the literature on sequential estimation has been rather scarce. One may look back at Mukhopadhyay and de Silva (2005) for a lone treatment in this case that we are aware of. In this light, Sect. 3 fills a part of this void with interesting directions.

Going back to Sect. 3, one could think of stopping rules different from (54) and (60) by replacing the sample variance S_n^2 with the following estimator:

$$\widehat{\mu}_{n,\mathrm{MLE}} + \widehat{\tau}_{n,\mathrm{MLE}}^{-1}\widehat{\mu}_{n,\mathrm{MLE}}^{2},$$

where $\widehat{\mu}_{n,\mathrm{MLE}}$ and $\widehat{\tau}_{n,\mathrm{MLE}}$, respectively, denote the maximum likelihood estimators (MLE) of μ, τ obtained from $X_1, \ldots, X_n, n \geq m$. In other words, at every stage of sampling one must continue to update $\widehat{\mu}_{n,\mathrm{MLE}}$ and $\widehat{\tau}_{n,\mathrm{MLE}}$ sequentially. Then, complexities of such sequential estimation strategies will increase tremendously, but their first-order properties and numerical performances may be expected to remain comparable to those reported in Sect. 3.

Acknowledgements We are grateful to Andreas Linden for giving us the permission to use bird count data found in Linden and Mantyniemi (2011) paper. We also appreciate Kees Waterman's approval to use the dataset from Heijting et al. (2007). The Reviewer, the Associate Editor, and the Editor gave us kind encouragements to tighten the original submission by streamlining its focus. We heartily thank them all.

References

Anscombe, F. J. (1949). The statistical analysis of insect counts based on the negative binomial distribution. *Biometrics*, *5*, 165–173.

Anscombe, F. J. (1950). Sampling theory of the negative binomial and logarithmic series distributions. *Biometrika*, *37*, 358–382.

Anscombe, F. J. (1952). Large-sample theory of sequential estimation. *Proceedings of Cambridge Philosophical Society*, *48*, 600–607.

Aoshima, M., Takada, Y. (2000). Second order properties of a two stage procedure for comparing several treatments with a control. *Journal of Japan Statistical Society*, *30*, 27–41.

Aoshima, M., Yata, K. (2010). Asymptotic second order consistency for two stage estimation methodologies and its applications. *Annals of the Institute of Statistical Mathematics*, *62*, 571–600.

Banerjee, S., Mukhopadhyay, N. (2016). A general sequential fixed-accuracy confidence interval estimation methodology for a positive parameter: Illustrations using health and safety data. *Annals of the Institute of Statistical Mathematics*, *68*, 541–570. http://dx.doi.org/10.1007/s10463-015-0504-2.

Bapat, S. R. (2017). Multistage sampling strategies and inference in health studies under appropriate Linex loss functions, Ph.D. thesis, August 2017, Department of Statistics, University of Connecticut, Storrs, Connecticut, USA.

Bliss, C. I., Owen, A. R. G. (1958). Negative binomial distributions with a common κ. *Biometrika*, *45*, 37–58.

Chattopadhyay, S. (1998). Sequential estimation of normal mean under asymmetric loss function with a shrinkage stopping rule. *Metrika*, *48*, 53–59.

Chattopadhyay, S. (2000). Sequential estimation of exponential location parameter using an asymmetric loss function. *Communications in Statistics-Theory & Methods*, *29*, 783–795.

Chow, Y. S., Robbins, H. (1965). On the asymptotic theory of fixed width sequential confidence intervals for the mean. *Annals of Mathematical Statistics*, *36*, 457–462.

Chow, Y. S., Robbins, H., Teicher, H. (1965). Moments of randomly stopped sums. *Annals of Mathematical Statistics*, *36*, 789–799.

Chow, Y. S., Hsiung, C., Lai, T. L. (1979). Extended renewal theory and moment convergence in Anscombe's theorem. *Annals of Probability*, *7*, 304–318.

Ghosh, M., Mukhopadhyay, N. (1979). Sequential point estimation of the mean when the distribution is unspecified. *Communications in Statistics-Theory & Methods*, *8*, 637–652.

Ghosh, M., Mukhopadhyay, N. (1981). Consistency and asymptotic efficiency of two-stage and sequential procedures. *Sankhya, Series A*, *43*, 220–227.

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

Heijting, S. (2013). *Weed count data from Heijting et al. (2007)*. DANS. http://dx.doi.org/10.17026/dans-zu9-r7y8.

Heijting, S., Van Der Werf, W., Stein, A., Kropff, M. J. (2007). Are weed patches stable in location? Application of an Explicitly Two-Dimensional Methodology, *Weed Research*, *47*, 381–395. http://dx.doi.org/10.17026/dans-zu9-r7y8v

Kuno, E. (1972). Some notes on population estimation by sequential sampling. *Researches on Population Ecology*, *14*, 58–73.

Linden, A., Mantyniemi, S. (2011). Using negative binomial distribution to model overdispersion in ecological count data. *Ecology*, *92*, 1414–1421.

Mukhopadhyay, N. (1974). Sequential estimation of location parameter in exponential distributions. *Calcutta Statistical Association Bulletin*, *23*, 85–95.

Mukhopadhyay, N. (1978). Sequential point estimation of the mean when the distribution is unspecified, Technical Report #312. School of Statistics: Department of Theoretical Statistics, University of Minnesota, Minneapolis, Minnesota, USA.

Mukhopadhyay, N. (2002). Sequential sampling. In A. H. Sharaawi, W. W. Piegorsch (Eds.), *The encyclopedia of environmetrics* (pp. 1983–1988). Chichester: Wiley.

Mukhopadhyay, N., Banerjee, S. (2014). Purely sequential and two stage fixed-accuracy confidence interval estimation methods for count data for negative binomial distributions in statistical ecology: One-sample and two-sample problems. *Sequential Analysis*, *33*, 251–285.
Mukhopadhyay, N., Banerjee, S. (2015). Sequential negative binomial problems and statistical ecology: A selected review with new directions. *Statistical Methodology*, *26*, 34–60.
Mukhopadhyay, N., Bapat, S. R. (2016a). Multistage point estimation methodologies for a negative exponential location under a modified linex loss function: Illustrations with infant mortality and bone marrow data. *Sequential Analysis*, *35*, 175–206. http://dx.doi.org/10.1080/07474946.2016.1165532.
Mukhopadhyay, N., Bapat, S. R. (2016b). Multistage estimation of the difference of locations of two negative exponential populations under a modified Linex loss function: Real data illustrations from cancer studies and reliability analysis. *Sequential Analysis*, *35*, 387–412. http://dx.doi.org/10.1080/07474946.2016.1206386.
Mukhopadhyay, N., de Silva, B. M. (2005). Two stage estimation of the mean of a negative binomial distribution with applications to mexican bean beetle data. *Sequential Analysis*, *24*, 99–137.
Mukhopadhyay, N., de Silva, B. M. (2009). *Sequential methods and their applications*. Boca Raton: CRC Press.
Mukhopadhyay, N., Diaz, J. (1985). Two stage sampling for estimating the mean of a negative binomial distribution. *Sequential Analysis*, *4*, 1–18.
Mukhopadhyay, N., Duggan, W. T. (1997). Can a two-stage procedure enjoy second order properties? *Sankhya, Series A*, *59*, 435–448.
Mukhopadhyay, N., Duggan, W. T. (2000). On a two-stage procedure having second-order properties with applications. *Annals of the Institute of Statistical Mathematics*, *51*, 621–636.
Mukhopadhyay, N., Duggan, W. T. (2001). A two-stage point estimation procedure for the mean of an exponential distribution and second-order results. *Statistics & Decisions*, *19*, 155–171.
Mukhopadhyay, N., Solanky, T. K. S. (1994). *Multistage selection and ranking procedures*. New York: Dekker.
Mukhopadhyay, N., Zacks, S. (2017). Modified Linex two-stage and purely sequential estimation of the variance in a normal distribution with illustrations using horticultural data. *Journal of Statistical Theory and Practice***(in press)**. http://dx.doi.org/10.1080/15598608.2017.1350608.
Mukhopadhyay, N., Datta, S., Chattopadhyay, S. (2004). *Applied sequential methodologies, edited volume*. New York: Dekker.
R Core Team. (2014). *R*: *A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
Sen, P. K., Ghosh, M. (1981). Sequential point estimation of estimable parameters based on U-statistics. *Sankhya, Series A*, *43*, 331–344.
Varian, H. R. (1975). A Bayesian approach to real estate assessment. In L. J. Savage, S. E. Fienberg, A. Zellner (Eds.), *Studies in Bayesian econometrics and statistics* (pp. 195–208). North Holland: Elsevier.
Wiener, N. (1939). The ergodic theorem. *Duke Mathematical Journal*, *5*, 1–18.
Willson, L. J., Folks, J. L. (1983). Sequential estimation of the mean of the negative binomial distribution. *Communications in Statistics-Sequential Analysis*, *2*, 55–70.
Willson, L. J., Folks, J. L., Young, J. H. (1984). Multistage estimation compared with fixed-sample size estimation of the negative binomial parameter κ. *Biometrics*, *40*, 109–117.
Zacks, S. (2009). *Stage-wise adaptive designs*. New York: Wiley.
Zellner, A. (1986). Bayesian estimation and prediction using asymmetric loss functions. *Journal of American Statistical Association*, *81*, 446–451.