Turkish Journal of Mathematics

http://journals.tubitak.gov.tr/math/

Research Article

Turk J Math
(2013) 37: 719 – 731
© TÜBİTAK
doi:10.3906/mat-1011-14

Second order approximations in sequential point estimation of the probability of zero in Poisson distribution

Eisa MAHMOUDI,* **Mohammad HATAMI KAMIN**
Department of Statistics, Yazd University, P.O. Box 89195-741, Yazd, Iran

Received: 04.11.2010 • **Accepted:** 12.05.2012 • **Published Online:** 12.06.2013 • **Printed:** 08.07.2013

Abstract: In the analysis of the count data, the Poisson model becomes overtly restrictive in the case of over-dispersed or under-dispersed data. When count data are under-dispersed, specific models such as generalized linear models (GLM) are proposed. Other examples are the zero-inflated Poisson model (ZIP) and zero-truncated Poisson model (ZTP), which have been used in literature to deal with an excess or absence of zeros in count data. Thus having a knowledge of the probability of zeros and its estimation in Poisson distribution can be significant and useful.

Some estimation problems with unknown parameter cannot attain minimum risk where the sample size is fixed. To resolve this captivity, working with a sequential sampling procedure can be useful. In this paper, we consider sequential point estimation of the probability of zero in Poisson distribution. Second order approximations to the expected sample size and the risk of the sequential procedure are derived as the cost per observations tends to zero. Finally, a simulation study is given.

Key words: Poisson distribution, regret, second-order approximations, sequential estimation

1. Introduction

The Poisson distribution with the probability density function

$$f(x;\theta) = \frac{e^{-\theta}\theta^x}{x!}, \qquad x = 0, 1, 2, \ldots, \quad \theta > 0, \tag{1}$$

was first studied by Poisson [13] as a limiting case of the binomial distribution. This distribution is very important among the discrete distributions. Johnson et al. [10] have discussed the genesis of Poisson distribution in detail. The Poisson distribution has wide applications in many fields. It is used as an approximation to the binomial and other distributions; used to describe when events occur randomly in time or space; used in certain models for the analysis of contingency tables; for the empirical treatment of count data; in quality control to characterize the number of defective items per batch (see, e.g., Walsh [22]; van der Waerden [20]; and Chatfield [4]); as a limiting form for the hypergeometric distribution and hence an approximation for sampling without replacement; in quantum statistics; and in the theory of photographic plates (see Feller [6], p. 59); and on the analysis of quadrant data which have been collected extensively in ecology, geology, geography, and urban studies (e.g., Greig-Smith [8]).

*Correspondence: emahmoudi@yazduni.ac.ir

2010 *AMS Mathematics Subject Classification:* 62L12; 62L15.

In spite of the wide applications of Poisson distribution, it becomes restrictive in analysis of some count data. Count data distributions are characterized by exhibiting a concentration of values on a few small discrete values, skewness to the right, and intrinsic heteroscedasticity with variance increasing or decreasing with the mean [3]. However, count data are often over-dispersed or under-dispersed relative to the Poisson distribution. That is, the variance is larger or smaller than mean for the data, so the Poisson model becomes overtly restrictive.

Under-dispersions have qualitatively similar consequences to heteroscedasticity in the linear regression models. When count data are under-dispersed specific models such as generalized linear models (GLM)are proposed. For example, the zero-inflated Poisson model (ZIP) and the zero-truncated Poisson model (ZTP) have been used in literature to deal with an excess or absence of zeros in count data [11, 24].

Poisson regression models provide a standard framework for the analysis of count data. In practice, however, count data are often over-dispersed relative to the Poisson distribution. One frequent manifestation of over-dispersion is that the incidence of zero counts is greater than expected for the Poisson distribution; this is of interest because zero counts frequently have special status. For example, in counting disease lesions on plants, a plant may have no lesions either because it is resistant to the disease, or simply because no disease spores have landed on it. This is the distinction between structural zeros, which are inevitable, and sampling zeros, which occur by chance. In recent years there has been considerable interest in models for count data that allow for excess zeros, particularly in the econometric literature. These models complement more conventional models for over-dispersion that concentrate on modeling the variance-mean relationship correctly.

According to what was mentioned above, having knowledge about the probability of zero i.e. $\gamma(\theta) = P(X=0) = e^{-\theta}$ and its estimation in Poisson distribution can be significant and useful.

The problem of sequential estimation refers to any estimation technique for which the total number of observations used is not a degenerate random variable. In some estimation problems with unknown parameter the sequential estimation must be used because no procedure using a preassigned non-random sample size can achieve the desired objective.

Wald [21] was the first person to introduce the concept of sequential probability ratio test (SPRT). Other authors such as Handle [9] and Stein [15] used the sequential methods to tackle some unsolved problems in point and interval estimation. Sequential estimation of the scale parameter of exponential and gamma distributions has been considered by Starr and Woodroofe [14], Woodroofe [23], Ghosh and Mukhopadhyay [7], Mukhopadhyay et al. [12], Takada [16, 17] and Uno and Isogai [18]. Ali and Isogai [1] derived the sequential point estimation of the powers of an exponential scale parameter. Also, Uno et al. [19] obtained the sequential point estimation of a function of the scale parameter of an exponential distribution subject to the loss function given as a sum of the squared error and a linear cost. In estimating the mean, their results coincided with that of Woodroofe [23].

Consider the weighted squared error loss function of the form

$$L_n = L(\gamma(\hat{\theta}), \gamma(\theta)) = Aw(\theta)(e^{-\bar{X}_n} - e^{-\theta})^2 + cn, \tag{2}$$

where $A(>0)$ and $w(\theta)(>0)$ are known to the experimenter and $c(>0)$ is the cost per unit sample for measured the accuracy of the estimate. This loss function is suitable in estimation problems for which the overestimation is more serious than the underestimation orvice versa. In this paper we consider sequential point estimation of the probability of zero in Poisson distribution under the above loss function.

The remainder of this paper is organized as follows. The stopping rule N is introduced and some properties are given in Section 2. In Section 3, we give second-order approximations of the $E(N)$ and asymptotic expansions of the risk R_N of the purely sequential procedure as $c \to 0$. Section 4 is devoted to sequential point estimation of the probability of zero in Poisson distribution, where $w(\theta) = 1/\theta$ in (2). A simulation study is given in Section 5.

2. The stopping rule and its properties

Suppose that $X_1, X_2, \ldots$ is a sequence of independent and identically distributed random variables having Poisson distribution (1). It is interesting to estimate the probability of $x = 0$: $\gamma(\theta) = P(X = 0) = e^{-\theta}$. Prevalently, we will estimate the function $\gamma(\theta)$ by $\gamma(\hat{\theta}) = e^{-\bar{X}_n}$, where $\bar{X}_n$ is the sample mean of a random sample of size n from the Poisson distribution with mean θ and variance θ. The risk function of the estimator $\gamma(\hat{\theta}) = e^{-\bar{X}_n}$ in estimating $\gamma(\theta) = e^{-\theta}$, under the loss function (2) is given by the relation

$$R_n = E[L(\gamma(\hat{\theta}), \gamma(\theta))] = Aw(\theta)E(e^{-\bar{X}_n} - e^{-\theta})^2 + cn. \tag{3}$$

Taylor expansion of $e^{-\bar{X}_n}$ around $e^{-\theta}$ gives

$$R_n \simeq \frac{A}{n} w(\theta)\theta e^{-2\theta} + cn, \tag{4}$$

for sufficiently large n. The risk function in (4) is approximately minimized at

$$n_0 \simeq \left(\frac{A}{c}\right)^{1/2} \sqrt{\theta w(\theta)} e^{-\theta} = n^* \quad \text{(say)}. \tag{5}$$

If θ is known, then n^* can be found to minimize the risk function R_n. But since θ is unknown, we cannot use the best fixed sample size procedure n_0 to achieve the minimum risk of R_{n_0}. Thus it is necessary to find a sequential sampling procedure. First, take a sample of size $m = 1$ from the Poisson distribution. If $m < \sqrt{\frac{A\bar{X}_m}{c} w(\bar{X}_m)} e^{-\bar{X}_m}$, then we take one more observation, X_{m+1}; otherwise sampling is terminated. Therefore we introduce the stopping rule

$$N = N_c = \inf\left\{ n \geq m : n \geq \sqrt{\frac{A\bar{X}_n}{c} w(\bar{X}_n)} e^{-\bar{X}_n} \right\}, \tag{6}$$

where $m \geq 1$ is the pilot sample size. In estimating the parameter $\gamma(\theta) = e^{-\theta}$ by $\gamma(\hat{\theta}) = e^{-\bar{X}_N}$, the risk is given by the relation

$$R_N = Aw(\theta)E(e^{-\bar{X}_N} - e^{-\theta})^2 + cE(N). \tag{7}$$

The performance of the procedure is measured by the regret $R_N - 2cn^*$. In the succeeding section, we will give second-order approximations of the $E(N)$ and asymptotic expansions of the risk R_N of the purely sequential procedure as $c \to 0$.

Proposition 1 *The stopping rule N in (6) has the following properties:*

(i) $P(N < \infty) = 1$,

(ii) $\lim_{c\to 0} N/n^* = 1$ *a.s.*,

(iii) $\lim_{c\to 0} E[(N/n^*)^2] = 1$,

Proof See Propositions 1 and 2 of Aras and Woodroofe [2]. □

3. Second-order approximations of $E(N)$

According to the stopping rule (6) we define the function $h(x)$ by

$$h(x) = \frac{e^x}{\sqrt{xw(x)}}, \qquad x > 0.$$

The stopping rule N in (6) can be written as

$$N = \inf\{n \geq m : Z_n \geq n^*\},$$

where $Z_n = n\frac{h(\bar{X}_n)}{h(\theta)}$.

Let $Y_i = (X_i - \theta)$, for $i = 1, 2, \ldots$, $S_n = \sum_{i=1}^{n} Y_i$ and $\bar{Y}_n = n^{-1}S_n$, respectively. Taylor's theorem gives

$$h(\bar{X}_n) = h(\theta) + h'(\theta)(\bar{X}_n - \theta) + \frac{h''(\eta_n)}{2}(\bar{X}_n - \theta)^2,$$

where η_n is a random variable lying between θ and $\bar{X}_n$. Using the above expression for $h(\bar{X}_n)$ we have

$$Z_n = n + \alpha S_n + \xi_n,$$

where

$$\alpha = 1 - \frac{1}{2\theta} - \frac{w'(\theta)}{2w(\theta)}, \quad \text{and} \quad \xi_n = n(\bar{X}_n - \theta)^2\frac{h''(\eta_n)}{2h(\theta)}, \tag{8}$$

with

$$h''(x) = \frac{e^x}{\sqrt{xw(x)}}\left\{1 - \frac{1}{x} + \frac{3}{4x^2} + \frac{w'(x)}{2xw(x)} - \frac{w'(x)}{w(x)} + \frac{3}{4}[\frac{w'(x)}{w(x)}]^2 - \frac{w''(x)}{2w(x)}\right\}.$$

Define

$$t = \inf\{n \geq 1 : n + \alpha S_n > 0\} \quad \text{and} \quad \rho = \frac{E(t + \alpha S_t)^2}{2E(t + \alpha S_t)}. \tag{9}$$

Consider the following assumptions:

($\mathbf{A}_1$) $\left\{\left[\left(Z_n - \frac{n}{\varepsilon_0}\right)^+\right]^3, n \geq m\right\}$ is uniformly integrable for some $0 < \varepsilon_0 < 1$.

($\mathbf{A}_2$) $\sum_{n=m}^{\infty} nP\{\xi_n < -n\varepsilon_1\} < \infty$ for some $0 < \varepsilon_1 < 1$.

The following theorem gives the expression of $E(N)$ as the cost per observation tends to zero.

Theorem 1 *If (A_1) and (A_2) hold, then*

$$E(N) = n^* + \rho - \theta\ell + o(1), \qquad as \quad c \to 0,$$

where

$$\ell = \frac{1}{2}\left\{1 - \frac{1}{\theta} + \frac{3}{4\theta^2} + \frac{w'(\theta)}{2\theta w(\theta)} - \frac{w'(\theta)}{w(\theta)} + \frac{3}{4}[\frac{w'(\theta)}{w(\theta)}]^2 - \frac{w''(\theta)}{2w(\theta)}\right\}. \tag{10}$$

Proof Let $W_n = \sqrt{n}(\bar{X}_n - \theta)$. Then the random variable W_n tends to the random variable W having normal distribution with mean 0 and variance θ as $n \to \infty$. Note that $h''(\eta_n) \to h''(\theta)$, since η_n is a random variable lying between θ and $\bar{X}_n$. Combining these two facts gives

$$\xi_n = \frac{h''(\eta_n)}{2h(\theta)} n(\bar{X}_n - \theta)^2 \xrightarrow{d} \frac{h''(\theta)}{2h(\theta)} W^2 = \xi \ \text{ as } n \to \infty,$$

where $\xi = \frac{h''(\theta)}{2h(\theta)} W^2 = \ell W^2$ and ℓ is given in (10). We shall check conditions (C1) to (C6) of Aras and Woodroofe [2]. Clearly, (C1) holds since $E(Y_i) = 0$ and $E(Y_i^2) < \infty$ for $i = 1, 2, \cdots$. (C2) with $p = 3$ is identical with (A_1). (C3) coincides with (A_2). Taking $g(x) = \frac{h(x+\theta)}{h(\theta)}$, (C4), (C5) and (C6) follow from Proposition 4 of Aras and Woodroofe [2] since $Z_n = ng(\bar{Y}_n) = n\frac{h(\bar{X}_n)}{h(\theta)}$, $g(x)$ is twice continuously differentiable on some neighborhood of 0, $\alpha = g'(0) = \frac{h'(\theta)}{h(\theta)}$ and $E|Y_1|^3 < \infty$. Hence, from Theorem 1 of Aras and Woodroofe [2],

$$E(N) = n^* + \rho - E(\xi) + o(1) = n^* + \rho - \theta\ell + o(1), \qquad as \quad c \to 0,$$

which concludes the theorem. □

The following proposition gives sufficient conditions for (A_2) which are useful in actual estimation problems.

Proposition 2

(i) *If* $h''(\eta_n) \geq 0$ *for all* $n \geq m$, *then* (A_2) *holds.*

(ii) *If* $\sup_{n \geq m} E|h''(\eta_n)|^s < \infty$ *for some* $s > 2$, *then* (A_2) *holds.*

Proof See Proposition 1 of Uno et al. [19] for more details. □

We shall now assess the regret $R_N - 2cn^*$. Using Taylor's theorem we obtain,

$$e^{-\bar{X}_N} = e^{-\theta} - e^{-\theta}(\bar{X}_N - \theta) + \frac{1}{2}e^{-\theta}(\bar{X}_N - \theta)^2 - \frac{1}{6}e^{-\varphi_c}(\bar{X}_N - \theta)^3, \tag{11}$$

where φ_c is a random variable lying between $\bar{X}_N$ and θ. Consider the following condition:

(A_3) For some $a > 1$, $u > 1$ and $c_0 > 0$,

$$\sup_{0<c\leq c_0} \left\{c^{-au} E\left|\bar{X}_N - \theta\right|^{4au}\right\} < \infty, \quad \sup_{0<c\leq c_0} E\left|e^{-\varphi_c}\right|^{\frac{2au}{u-1}} < \infty,$$

where φ_c is the random variable lying between $\bar{X}_N$ and θ.

Remark 1 *The second part of (A_3) is satisfied here, since $|e^{-x}| \leq 1$ for each $x > 0$.*

Theorem 2 *If (A_1), (A_2) and (A_3) hold, then*

$$R_N - 2cn^* = \left\{ \tfrac{3}{4}\theta + \tfrac{5}{4\theta} - 1 + \tfrac{9}{4}\theta[\tfrac{w'(\theta)}{w(\theta)}]^2 - 2\theta\tfrac{w'(\theta)}{w(\theta)} + \tfrac{3}{2}\tfrac{w'(\theta)}{w(\theta)} - \theta\tfrac{w''(\theta)}{w(\theta)} \right\} c + o(c). \tag{12}$$

Proof Using the expressions (7) and (11), we have

$$\begin{aligned} &R_N - 2cn^* = Aw(\theta)E(e^{-\bar{X}_N} - e^{-\theta})^2 + cE(N) - 2cn^* \\ &= Aw(\theta)e^{-2\theta}E(\bar{X}_N - \theta)^2 + cE(N) - 2cn^* + Aw(\theta)\tfrac{e^{-2\theta}}{4}E(\bar{X}_N - \theta)^4 \\ &+ \tfrac{A}{36}w(\theta)E[e^{-2\varphi_c}(\bar{X}_N - \theta)^6] - Aw(\theta)e^{-2\theta}E(\bar{X}_N - \theta)^3 \\ &+ \tfrac{A}{3}w(\theta)e^{-\theta}E[e^{-\varphi_c}(\bar{X}_N - \theta)^4] - \tfrac{A}{6}w(\theta)e^{-\theta}E[e^{-\varphi_c}(\bar{X}_N - \theta)^5]. \end{aligned} \tag{13}$$

Using Theorems 2 and 3 of Aras and Woodroofe [2], we obtain (14)–(19) below, as $c \to 0$:

$$\begin{aligned} &Aw(\theta)e^{-2\theta}E(\bar{X}_N - \theta)^2 + cE(N) - 2cn^* \\ &= c\left\{ {n^*}^2 E\left[\sqrt{\tfrac{1}{\theta}}\bar{Y}_N\right]^2 + E(N) - 2n^* \right\} \\ &= c\left\{ 2E\left(\tfrac{1}{\theta}\xi W^2 - \xi\right) + 3\theta\alpha^2 + \tfrac{2}{\theta}\alpha E(Y_1)^3 \right\} \\ &= \left\{ \tfrac{5}{4\theta} + 5\theta - 3 + \tfrac{9}{4}\theta[\tfrac{w'(\theta)}{w(\theta)}]^2 - 5\theta\tfrac{w'(\theta)}{w(\theta)} + \tfrac{3}{2}\tfrac{w'(\theta)}{w(\theta)} - \theta\tfrac{w''(\theta)}{w(\theta)} \right\} c + o(c). \end{aligned} \tag{14}$$

Here, $\xi = \ell W^2$, $W \sim N(0, \theta)$ and α and ℓ are given in (8) and (10). Also,

$$\begin{aligned} Aw(\theta)e^{-2\theta}E(\bar{X}_N - \theta)^3 &= \sqrt{\theta}c\left\{ {n^*}^2 E(\sqrt{\tfrac{1}{\theta}}\bar{Y}_N)^3 \right\} \\ &= \sqrt{\theta}c\left\{ \sqrt{\tfrac{1}{\theta}}6\theta(1 - \tfrac{1}{2\theta}) + \sqrt{\tfrac{1}{\theta}}\tfrac{1}{\theta}\theta \right\} \\ &= \left\{ 6\theta - 2 - 3\theta\tfrac{w'(\theta)}{w(\theta)} \right\} c + o(c). \end{aligned} \tag{15}$$

$$\begin{aligned} Aw(\theta)\tfrac{e^{-2\theta}}{4}E(\bar{X}_N - \theta)^4 &= \tfrac{1}{4}\theta c\left\{ {n^*}^2 E(\sqrt{\tfrac{1}{\theta}}\bar{Y}_N)^4 \right\} \\ &= \left\{ \tfrac{3}{4}\theta \right\} c + o(c). \end{aligned} \tag{16}$$

For $r > 1$, $s > 1$, $k = \frac{s}{s-1}$ and $v = \frac{u}{u-1}$, we have

$$\begin{aligned} &E\left| \left\{ (n^*)^{\frac{1}{2}}\bar{Y}_N \right\}^4 e^{-2\varphi_c}\left(\bar{Y}_N\right)^2 \right|^r \\ &\leq \left\{ E\left|(n^*)^{\frac{1}{2}}\bar{Y}_N\right|^{4rus} \right\}^{\frac{1}{us}} \left\{ E\left|e^{-\varphi_c}\right|^{2rvs} \right\}^{\frac{1}{vs}} \left\{ E\left|\bar{Y}_N\right|^{2rk} \right\}^{\frac{1}{k}}. \end{aligned}$$

Doob's maximal inequality gives

$$E(\bar{Y}_N)^{2rk} \leq E\left\{ \sup_{n\geq 1}(\bar{Y}_n)^{2rk} \right\} \leq (\frac{2rk}{2rk-1})^{2rk}E(Y_1)^{2rk} < \infty.$$

Choosing r and s such that $a = rs$, then (A_3) yields the uniform integrability of

$$\left\{\left\{(n^*)^{\frac{1}{2}}\bar{Y}_N\right\}^4 e^{-2\varphi_c}\left(\bar{Y}_N\right)^2, 0 < c \le c_0\right\}.$$

Since

$$\left\{(n^*)^{\frac{1}{2}}\bar{Y}_N\right\}^4 e^{-2\varphi_c} \xrightarrow{d} e^{-2\theta}W^4,$$

and

$$\left(\bar{Y}_N\right)^2 \to 0, \quad as \quad c \to 0,$$

we get

$$\begin{aligned} &\tfrac{A}{36}w(\theta)E[e^{-2\varphi_c}(\bar{X}_N-\theta)^6] \\ &= \tfrac{c}{36\theta}e^{2\theta}E\left\{\left\{(n^*)^{\frac{1}{2}}\bar{Y}_N\right\}^4 e^{-2\varphi_c}\left(\bar{Y}_N\right)^2\right\} = o(c). \end{aligned} \tag{17}$$

By arguments similar to (17), we obtain

$$\begin{aligned} &\tfrac{A}{3}w(\theta)e^{-\theta}E[e^{-\varphi_c}(\bar{X}_N-\theta)^4] \\ &= \tfrac{c}{3\theta}e^{\theta}E\left\{\left\{(n^*)^{\frac{1}{2}}\bar{Y}_N\right\}^4 e^{-\varphi_c}\right\} = \{\theta\}c + o(c), \end{aligned} \tag{18}$$

and

$$\begin{aligned} &\tfrac{A}{6}w(\theta)e^{-\theta}E[e^{-\varphi_c}(\bar{X}_N-\theta)^5] \\ &= \tfrac{c}{6}e^{\theta}E\left\{\left\{(n^*)^{\frac{1}{2}}\bar{Y}_N\right\}^4 e^{-\varphi_c}\left\{\bar{Y}_N\right\}\right\} = o(c). \end{aligned} \tag{19}$$

Substituting (14)–(19) into (13), we get

$$\begin{aligned} R_N - 2cn^* &= \left\{\tfrac{3}{4}\theta + \tfrac{5}{4\theta} - 1 + \tfrac{9}{4}\theta[\tfrac{w'(\theta)}{w(\theta)}]^2 - 2\theta\tfrac{w'(\theta)}{w(\theta)} + \tfrac{3}{2}\tfrac{w'(\theta)}{w(\theta)} - \theta\tfrac{w''(\theta)}{w(\theta)}\right\}c \\ &+ o(c). \end{aligned}$$

The proof is completed. □

4. Sequential point estimation of the probability of zero where $w(\theta) = 1/\theta$

In this section we consider sequential point estimation of the probability of zero in Poisson distribution, i.e. $\gamma(\theta) = e^{-\theta}$ under the weighted squared error loss function (2) with $w(\theta) = 1/\theta$. The main idea of choosing this weight is that the sample mean $\bar{X}_n$ is admissible and a minimax estimator for θ under the weighted squared error loss function $L(\hat{\theta}, \theta) = \frac{1}{\theta}(\bar{X}_n - \theta)^2$.

In estimating $\gamma(\theta) = e^{-\theta}$ by $\gamma(\hat{\theta}) = e^{-\bar{X}_n}$ under the loss function

$$L(\gamma(\hat{\theta}), \gamma(\theta)) = \frac{A}{\theta}(e^{-\bar{X}_n} - e^{-\theta})^2 + cn, \tag{20}$$

the risk is approximately given by

$$R_n \simeq \frac{A}{n}e^{-2\theta} + cn, \tag{21}$$

which is finite for $n \geq 1$. Then the optimal fixed-sample size is

$$n^* = (\frac{A}{c})^{1/2}e^{-\theta}. \tag{22}$$

The stopping rule N in (6) is given by

$$N = \inf\left\{n \geq m; \quad n \geq (\frac{A}{c})^{1/2}e^{-\bar{X}_n}\right\}. \tag{23}$$

Since $w'(\theta) = -1/\theta^2$ and $w''(\theta) = 2/\theta^3$, we have

$$\alpha = 1 \qquad \text{and} \qquad \ell = 1/2. \tag{24}$$

The stopping rule t in (9) becomes $t = \inf\{n \geq 1; \quad n + \alpha S_n > 0\} = 1$ and

$$\rho = \frac{E(1+Y_1)^2}{2E(1+Y_1)} = \frac{\theta+1}{2}. \tag{25}$$

Second-order approximations of the expected sample size are given in the following theorem.

Theorem 3 *Suppose that $m \geq 1$. Consider the sequential point estimation of a function $\gamma(\theta) = e^{-\theta}$ under the weighted squared error loss function (20). Then as $c \to 0$*

$$E(N) = n^* + \frac{1}{2} + o(1). \tag{26}$$

Proof The stopping rule N in (23) can be written as

$$N = \inf\{n \geq m; \quad Z_n \geq n^*\},$$

where $Z_n = n\frac{h(\bar{X}_n)}{h(\theta)} = ng(\bar{Y}_n)$, in which $g(x) = e^x$. Note that the function $g(x)$ is convex and

$$E[\{g(Y_1)\}^+]^3 = E\{e^{Y_1}\}^3 = e^{-3\theta}M_X(3) = e^{16.08\theta} < \infty,$$

where $M_X(t)$ denotes the moment generating function of the Poisson distribution. Thus from Proposition 5 of Aras and Woodroofe [2] (A1) and (A2) hold. Substituting α and ℓ from (24) and ρ from (25) in Theorem 1 gives

$$E(N) = n^* + \frac{1}{2} + o(1),$$

as $c \to 0$. □

We need the following two lemmas to obtain the regret $R_N - 2cn^*$ of the sequential point estimation of the probability of zero in Poisson distribution i.e., $\gamma(\theta) = e^{-\theta}$. Let M stand for a generic positive constant not depending on c and let $c_0 > 0$ be a constant such that $n^* \geq 1$.

Lemma 1 *Let $q \geq 1$. Consider sequential point estimation of the function $\gamma(\theta) = e^{-\theta}$ with corresponding stopping rule N in (23). The following results are satisfied:*

(i) $\left\{(\frac{N}{n^*})^{-q}, c > 0\right\}$ *is uniformly integrable.*

(ii) *If* $m > q$, *then* $\left\{(\frac{N}{n^*})^{q}, 0 < c \leq c_0\right\}$ *is uniformly integrable.*

Proof Part(i): It's enough to show $E[\sup_{c>0}(\frac{N}{n^*})^{-q}] < \infty$. In sequential point estimation of the function $\gamma(\theta) = e^{-\theta}$, the optimal fixed-sample size is given by $n^* = (A/c)^{1/2}e^{-\theta}$. According to (23) the stopping rule N is given by

$$N = \inf\left\{n \geq m; \quad n \geq (A/c)^{1/2}e^{-\bar{X}_n}\right\}.$$

For $q \geq 1$ we have $(\frac{n^*}{N})^q \leq (\frac{e^{\bar{X}_N}}{e^{\theta}})^q$. On the other hand, $\sup_{c>0} e^{\bar{X}_N} \leq \sup_{n\geq 1} e^{\bar{X}_n}$, since $q \geq 1$ and $N \leq n$ on the set $\{n \geq 1\}$.
Finally, we have

$$E[\sup_{c>0}(\frac{N}{n^*})^{-q}] \leq e^{-q\theta}E[\sup_{c>0} e^{q\bar{X}_N}] \leq e^{-q\theta}E[\sup_{n\geq 1} e^{q\bar{X}_n}] < \infty,$$

and the proof of Part (i) is completed.

Part(ii): To show (ii), observe that $\frac{n^*}{N-1} > \frac{e^{\bar{X}_{N-1}}}{e^{\theta}}$ which gives $n^* > (N-1)\frac{e^{\bar{X}_{N-1}}}{e^{\theta}}$, on the set $\{N \geq m\}$. Now we can write

$$\frac{N}{n^*} = \frac{N}{n^*}I_{\{N>m\}} + \frac{N}{n^*}I_{\{N=m\}} \leq \frac{N}{n^*}I_{\{N>m\}} + \frac{m}{n^*}.$$

On the other hand,

$$\begin{aligned} \tfrac{N}{n^*}I_{\{N>m\}} = \{\tfrac{N-1}{n^*} + \tfrac{1}{n^*}\}I_{\{N>m\}} \quad &\leq e^{\theta}e^{-\bar{X}_{N-1}} + \tfrac{1}{n^*}\}I_{\{N>m\}} \\ &\leq e^{\theta}e^{-\bar{X}_{N-1}}I_{\{N>m\}} + \tfrac{1}{n^*}. \end{aligned} \tag{27}$$

Thus we have

$$\frac{N}{n^*} \leq e^{\theta}e^{-\bar{X}_{N-1}}I_{\{N>m\}} + \frac{m+1}{n^*} \leq e^{\theta}e^{-\bar{X}_{N-1}}I_{\{N>m\}} + m + 1. \tag{28}$$

According to expression (28) for $0 < c \leq c_0$ we have

$$(\frac{N}{n^*})^q \leq M\{(e^{\theta}e^{-\bar{X}_{N-1}})^q I_{\{N>m\}} + (m+1)^q\}.$$

Hence

$$\begin{aligned} \sup_{0<c\leq c_0}\{(\tfrac{N}{n^*})^q\} \quad &\leq M\left\{e^{q\theta}\sup_{0<c\leq c_0}\{e^{-q\bar{X}_{N-1}}\}I_{\{N>m\}} + (m+1)^q\right\} \\ &\leq M\left\{e^{q\theta} + (m+1)^q\right\} < \infty, \end{aligned}$$

since $|e^{-x}| \leq 1,\ \forall\, x > 0$. Thus the second assertion holds. □

The next lemma follows from Theorem 2 of Chow et al. [5].

Lemma 2 *Let* $q \geq 1$. *If* $\left\{ \left(\frac{N}{n^*}\right)^q, 0 < c \leq c_0 \right\}$ *is uniformly integrable, then*

$$\left\{ \left| (n^*)^{-\frac{1}{2}} \sum_{i=1}^{N} (X_i - \theta) \right|^q, \ 0 < c \leq c_0 \right\}$$

is uniformly integrable.

The following theorem gives the regret $R_N - 2cn^*$ of sequential point estimation of the probability of zero in a Poisson distribution.

Theorem 4 *Consider sequential point estimation of the function* $\gamma(\theta) = e^{-\theta}$ *under weighted squared error loss function (20). If* $m > 12$, *then we have*

$$R_N - 2cn^* = \{\frac{3}{4}\theta + 1\}c + o(c),$$

as $c \to 0$.

Proof According to Theorem 2 we must show that conditions (A1), (A2) and (A3) are satisfied. Obviously, conditions (A1) and (A2) are satisfied using Theorem 3. Thus we must show condition (A3). The second part of condition (A3) is also satisfied using Remark 1. We shall show the first part of (A3) with $u = 3$. Choose $a > 1$, $p > 1$ and $q = \frac{p}{p-1}$, such that $m > 12a$. Hölder inequality gives

$$\begin{aligned} c^{-3a} \left\{ E \left| \bar{X}_N - \theta \right|^{12a} \right\} &= E \left\{ \left| c^{-1/4} (\bar{X}_N - \theta) \right|^{12a} \right\} \\ &= E \left\{ \left| \tfrac{e^{\theta/2}}{A^{1/4}} {n^*}^{1/2} (\bar{X}_N - \theta) \right|^{12a} \right\} \\ &\leq ME \left\{ \left| \tfrac{n^*}{N} {n^*}^{-1/2} \sum_{i=1}^{N} (X_i - \theta) \right|^{12a} \right\} \\ &\leq ME \left\{ (\tfrac{n^*}{N})^{12aq} \right\}^{1/q} E \left\{ \left| {n^*}^{-1/2} \sum_{i=1}^{N} (X_i - \theta) \right|^{12ap} \right\}^{1/p}, \end{aligned}$$

which together with Lemmas 1 and 2 implies

$$\sup_{0 < c \leq c_0} \left\{ c^{-3a} E \left| \bar{X}_N - \theta \right|^{12a} \right\} < \infty.$$

Thus the first part of condition (A3) is satisfied. Therefore, Theorem 2 with

$$R_N - 2cn^* = \{\frac{3}{4}\theta + 1\}c + o(c)$$

proves Theorem 4. □

5. Simulation study

In order to justify the results of Theorems 3 and 4 in the previous section we shall give brief simulation results. Monte Carlo simulation is performed to illustrate the behavior and performance of the stopping rule in the proposed sequential procedure as $c \to 0$. The results of the Monte Carlo simulation, based on the sequential rule are summarized in Tables 1, 2 and 3 which show several choices of the parameter θ, namely $\theta = 0.5$, 2 and 5 with corresponding values 0.6065, 0.1353 and 0.0067 for the probability of zero, i.e., $\gamma(\theta) = e^{-\theta}$. Since the cost c is sufficiently small in our theorems, the values of c are chosen such that $n^* = 50$, 75, 100 and 200. The pilot sample size is set at $m = 13$ and the values of A are 2, 5 and 10.

The simulation results in Tables 1, 2 and 3 are based on 1,000,000 repetitions by means of the stopping rule N defined by (23). Each table contains the selected value of c, the optimal fixed sample size n^*, the estimate of $\hat{\theta} = \bar{X}_N$ and $\gamma(\hat{\theta}) = e^{-\bar{X}_N}$, the average of the stopping time, $E(N)$, the average risk associated with the stopping time N, R_N, the regret $R_N - 2cn^*$ and the values $(R_N - 2cn^*)/c$.

Tables 1, 2 and 3 show that the results of Theorems 3 and 4 are justified. Further, it appears that the estimates $E(\hat{\theta}) = E(\bar{X}_N)$ for θ and $E(\gamma(\hat{\theta})) = E(e^{-\bar{X}_N})$ for $\gamma(\theta) = e^{-\theta}$ in the three tables are very close to the true values. Therefore, our sequential procedure seems to be effective and useful. From Tables 1, 2 and 3, we see that the expected stopping time $E(N)$ is very close to the optimal fixed sample size n^* and is uniformly larger than the optimal stopping time n^*. That is, the suggested procedure requires larger sample sizes than the fixed-sample procedure. We also observe that as the optimal fixed sample size n^* becomes larger, the average risk R_N and the regret $R_N - 2cn^*$ tend to zero. Also, the average risk under stopping time N decrease as $c \to 0$. As $c \to 0$, the quantity $\frac{R_N - 2cn^*}{c}$ become close to $\frac{3}{4}\theta + 1$, according to Theorem 4.

Table 1. Simulation results for sequential point estimation of the probability of zero in a Poisson distribution, with $A = 2$ and $m = 13$.

				$\theta = 0.5$			
n^*	c	$\hat{\theta}$	$\exp(\hat{\theta})$	$E[N]$	R_N	$R_N - 2cn^*$	$\frac{R_N - 2cn^*}{c}$
50	2.94×10^{-4}	0.509946	0.603712	50.398	0.029796	3.66×10^{-4}	1.24349
75	1.31×10^{-4}	0.506830	0.604499	75.392	0.019784	1.65×10^{-4}	1.25795
100	7.35×10^{-5}	0.504701	0.605238	100.419	0.014809	9.41×10^{-5}	1.27829
200	1.84×10^{-5}	0.502743	0.605636	200.337	0.007382	2.43×10^{-5}	1.30479
				$\theta = 2$			
50	1.46×10^{-5}	2.64650	0.132386	49.976	1.51×10^{-3}	4.38×10^{-5}	2.9938
75	6.51×10^{-6}	2.02870	0.133510	75.044	9.96×10^{-4}	1.89×10^{-5}	2.9083
100	3.66×10^{-6}	2.02160	0.133904	100.048	7.42×10^{-4}	9.34×10^{-6}	2.5522
200	9.16×10^{-7}	2.01080	0.134581	199.949	3.68×10^{-4}	2.27×10^{-6}	2.4868
				$\theta = 5$			
50	3.63×10^{-8}	5.13570	0.006354	48.702	3.81×10^{-6}	1.84×10^{-7}	5.0726
75	1.61×10^{-8}	5.09410	0.006454	73.385	2.51×10^{-6}	7.98×10^{-8}	4.9421
100	9.08×10^{-9}	5.06230	0.006550	98.775	1.85×10^{-6}	4.35×10^{-8}	4.7912
200	2.27×10^{-9}	5.02890	0.006639	198.651	9.18×10^{-7}	1.08×10^{-8}	4.7423

Table 2. Simulation results for sequential point estimation of the probability of zero in a Poisson distribution, with $A = 5$ and $m = 13$.

				$\theta = 0.5$			
n^*	c	$\hat{\theta}$	$\exp(\hat{\theta})$	$E[N]$	R_N	$R_N - 2cn^*$	$\frac{R_N-2cn^*}{c}$
50	7.36×10^{-4}	0.510370	0.603455	50.3738	0.074440	8.66×10^{-4}	1.17775
75	3.27×10^{-4}	0.506368	0.604753	75.424	0.049454	4.04×10^{-4}	1.23420
100	1.84×10^{-4}	0.505274	0.604887	100.367	0.037026	2.39×10^{-4}	1.29821
200	4.59×10^{-5}	0.502572	0.605741	200.369	0.018454	6.18×10^{-5}	1.34410
				$\theta = 2$			
50	3.66×10^{-5}	2.04610	0.132440	49.972	3.77×10^{-3}	1.11×10^{-4}	3.0396
75	1.63×10^{-5}	2.02930	0.133418	74.9944	2.49×10^{-3}	4.46×10^{-5}	2.7401
100	9.16×10^{-6}	2.02211	0.133820	99.937	1.85×10^{-3}	2.40×10^{-5}	2.6212
200	2.29×10^{-6}	2.01047	0.134619	199.990	9.22×10^{-4}	5.79×10^{-6}	2.5301
				$\theta = 5$			
50	9.08×10^{-8}	5.13750	0.006342	48.617	9.49×10^{-6}	4.11×10^{-7}	4.5269
75	4.04×10^{-8}	5.09150	0.006467	73.528	6.24×10^{-6}	1.85×10^{-7}	4.5921
100	2.27×10^{-8}	5.06475	0.006538	98.581	4.65×10^{-6}	1.06×10^{-7}	4.6513
200	5.68×10^{-9}	5.02780	0.006647	198.866	2.30×10^{-6}	2.68×10^{-8}	4.7212

Table 3. Simulation results for sequential point estimation of the probability of zero in a Poisson distribution, with $A = 10$ and $m = 13$.

				$\theta = 0.5$			
n^*	c	$\hat{\theta}$	$\exp(\hat{\theta})$	$E[N]$	R_N	$R_N - 2cn^*$	$\frac{R_N-2cn^*}{c}$
50	1.47×10^{-3}	0.510120	0.603611	50.391	0.149218	0.062067	1.40510
75	6.54×10^{-4}	0.507307	0.604204	75.353	0.098921	8.20×10^{-4}	1.25352
100	3.68×10^{-4}	0.504720	0.605229	100.419	0.074109	5.09×10^{-4}	1.38190
200	9.19×10^{-5}	0.502538	0.605758	200.377	0.036885	1.25×10^{-4}	1.36218
				$\theta = 2$			
50	7.33×10^{-5}	2.04660	0.132394	49.957	0.007553	2.27×10^{-4}	3.1005
75	3.26×10^{-5}	2.02919	0.133445	75.011	0.004977	9.31×10^{-5}	2.8606
100	1.83×10^{-5}	2.02108	0.133952	100.020	0.003711	5.09×10^{-5}	2.7832
200	4.58×10^{-6}	2.01062	0.134603	199.972	0.001844	1.18×10^{-6}	2.5802
				$\theta = 5$			
50	1.82×10^{-7}	5.13620	0.006349	48.671	1.91×10^{-5}	9.26×10^{-7}	5.1012
75	8.08×10^{-8}	5.09389	0.006450	73.335	1.25×10^{-5}	4.03×10^{-7}	4.9912
100	4.54×10^{-8}	5.06430	0.006540	98.633	9.30×10^{-6}	2.20×10^{-7}	4.8502
200	1.14×10^{-8}	5.02597	0.006658	199.197	4.59×10^{-6}	5.35×10^{-8}	4.7112

References

[1] Ali, M., Isogai, E.: Sequential point estimation of the powers of an exponential scale parameter. Scientiae Mathematicae Japonicae 58, 39–53 (2003).

[2] Aras, G., Woodroofe, M.: Asymptotic expansions for the moments of a randomly stopped average. The Annals of Statistics 21, 503–519 (1993).

[3] Cameron, A.C., Trivedi, P. K.: Essentials of count data regression, A Companion to Theoretical Econometrics. Oxford. Blackwell 2001.

[4] Chatfield, C.: Statistics for Technology, 3nd ed. London. Chapman Hall 1983.

[5] Chow, Y.S., Hsiung, C.A., Lai, T.L.: Extended renewal theory and moment convergence in Anscombes theorem. The Annals of Probability 7, 304–318 (1979).

[6] Feller, W.: An Introduction to Probability Theory and Its Applications, 3nd ed., vol. 1. New York. Wiley 1968.

[7] Ghosh M., Mukhopadhyay, N.: Sequential estimation of the percentiles of exponential and normal distributions. The South Africa Statistics Journal 23, 117–136 (1989).

[8] Greig-Smith, P.: Quantitative Plant Ecology, 2nd ed. London. Butterworth 1964.

[9] Handle, J.B.S.: On a method of estimating frequencies. Biometrika 33, 222–225 (1945).

[10] Johnson, N.L., Kotz, S., Kemp, A.W.: Univariate discrete distributions, 2nd ed. New York. John Wiley 1992.

[11] Lee, A.H., Wang, K., Yau, K.K.W.: Analysis of zero inflated Poisson data incorporating extent of exposure. Biometrical Journal 43, 963–975 (2001).

[12] Mukhopadhyay, N., Padmanabhan, A.R., Solanky, T.K.S.: On estimating the reliability after sequentially estimating the mean: the exponential case. Metrika 45, 235–252 (1997).

[13] Poisson, S.D.: Recherches sur la probabilite des Jugements en Matiere Criminelle et en Matiere Civile, Precedees des Regles. Generales du calcul des Probabilities, Paris: Bachelier,Imprimeur Libraire pour les mathematiques, la physique, etc., 1837.

[14] Starr, N., Woodroofe, M.: Further remarks on sequential estimation. The Annals of Mathematics Statistics 43, 1147–1154 (1972).

[15] Stein, C.: A two sample test for a linear hypothesis whose power is independent of the variance. Ann. Math. Statist. 16, 243–258 (1945).

[16] Takada, Y.: Non-existence of fixed sample size procedures for scale families. Sequential Analysis 5, 93–100 (1986).

[17] Takada, Y.: Fixed-width confidence intervals for a function of normal parameters. Sequential Analysis 16, 107–117 (1997).

[18] Uno C., Isogai, E.: Sequential point estimation of the powers of a normal scale parameter. Metrika 55, 215–232 (2002).

[19] Uno, C., Isogai, E., Lou Lim, D.: Sequential point estimation of a function of the exponential scale parameter. Austrian Journal of Statistics 33(3), 281–291 (2004).

[20] van der Waerden, B.L.: Sampling inspection as a minimum loss problem. Annals of Mathematical Statistics 31, 369–384 (1960).

[21] Wald, A.: Sequential tests of statistical hypothesis. Annals of Mathematical Statistics 16, 117–186 (1945).

[22] Walsh, J.E.: The Poisson distribution as a limit for dependent binomial events with unequal probabilities. Operations Research 3, 198–209 (1955).

[23] Woodroofe, M.: Second order approximations for sequential point and interval estimation. The Annals of Statistics 5, 984–995 (1977).

[24] Yau, K.K.W., Wang, K., Lee, A.H.: Zero-inflated negative binomial mixed regression modeling of over-dispersed count data with extra zeros. Biometrical Journal 45, 437–452 (2003).