Journal of Statistical Planning and Inference 17 (1987) 51-65
North-Holland

ASYMPTOTIC EXPANSIONS FOR FIXED WIDTH CONFIDENCE INTERVAL

Abd-El-Hakim MESLEM*

Department of Statistics, University of Michigan, Ann Arbor, MI 48109, USA

Received 30 June 1986; revised manuscript received 20 October 1986
Recommended by M.L. Puri

Abstract: A sequential procedure for setting a fixed width confidence interval for the natural parameter of an exponential family is proposed and studied. Average confidence levels are considered instead of the classical confidence curves. The procedure is shown to perform very well in a simulation study.

AMS Subject Classification: 62L12.

Key words and phrases: Stopping time; Average coverage probability; Baysian model; Excess over the boundary; Asymptotic expansions.

1. The problem

Let Ω be an open interval of the real line. A family F_ω, $\omega \in \Omega$, of probability distributions is said to form an exponential family if

$$dF_\omega(x) = \exp\{\omega x - \psi(\omega)\}\, d\Lambda(x),$$

where the dominating measure Λ is a non degenerate sigma finite measure on the Borel sets of the real line. Let Ω be the natural parameter space; that is, the set of all ω for which $\int e^{\omega z}\, d\Lambda(x)$ is finite. It is known that the natural parameter space is an interval. Ω° and $\bar{\Omega}$ will respectively denote the interior and the closure of Ω. The function ψ is convex and real analytic on Ω° and its derivative ψ' is increasing. The mean and variance of F_ω are

$$\theta = \psi'(\omega) \quad \text{and} \quad \sigma^2 = \psi''(\omega).$$

Let $\Theta = \psi'(\Omega) = (\underline{\theta}, \bar{\theta})$. Θ is sometimes called the expectation space and the family of distributions above can be reparametrized in terms of θ as well.

Let $X_1, X_2, \dots$ be independent and identically distributed with common distribution F_ω where $\omega \in \Omega$ is unknown. Let

$$\hat{\theta}_n = \frac{S_n}{n} \quad \text{where} \quad S_n = X_1 + X_2 + \cdots + X_n, \quad n \geq 1.$$

* Research supported by the U.S. Army under DAAG 29-85-K-0008.

Thus, $\hat{\theta}_n$ is the maximum likelihood estimate of θ. Define $\hat{\omega}_n$, $n \geq 1$, as follows: if $\theta \leq \underline{\theta}$, then $\hat{\omega}_n = \underline{\omega}$; and if $\underline{\theta} < \hat{\theta}_n < \bar{\theta}$, then $\psi'(\hat{\omega}_n) = \hat{\theta}_n$; and if $\hat{\theta}_n \geq \bar{\theta}$, then $\hat{\omega}_n = \bar{\omega}$. When $\hat{\theta}_n \in \Theta$, $\hat{\omega}_n$ is the maximum likelihood estimate of ω. Finally, let

$$U_n = \max\{1/n, \min\{n, \psi^n\}\};$$

extend U_n continuously to all of $\bar{\Omega}$ for each $n \geq 1$; and let

$$\hat{\sigma}_n^2 = U_n(\hat{\omega}_n), \quad n \geq 1.$$

Then $\hat{\sigma}_n^2$, $n \geq 1$, is an asymptotically efficient sequence of positive estimates of σ^2.

Given $h > 0$ and a fixed γ, $0 < \gamma < 1$, we are interested in finding a stopping time $t = t(X_1, X_2, \ldots)$ such that:

$$\gamma_h(t, \omega) := P_\omega\{|\omega - \hat{\omega}_t| \leq h\} \simeq \gamma, \quad \forall h \simeq 0, \ \forall \omega \in \Omega.$$

Apparently, Dantzig (1944) was the first to point out that many such problems cannot be solved by a single stage; that is, it may be impossible to prescribe an integer n and give a confidence interval of preassigned length and confidence based on a sample of size n. The earliest result of a positive nature is the paper by Stein (1945) who devised a two stage sampling procedure for estimating the mean of a normal distribution with unknown variance. Anscombe (1953) and Chow and Robbins (1965) developed a more efficient sequential procedure which uses a smaller sample size to achieve a confidence interval with very nearly the same coverage probability. An asymptotic analysis of the latter procedure is given by the work of Stein (1949) and Anscombe (1953). Starr (1966) made an extensive study of the normal case with some numerical computations. Simmons (1968) proved that the expected sample size $E_\sigma(t)$ is equal to $N + O(1)$ where N is the hypothetical sample size needed had σ^2 been known. Woodroofe (1977) extended the latter result to a second order approximation using non-linear renewal theory. Hall (1981) appended a third stage to Stein's sampling method and described the asymptotic theory of triple sampling as it pertains to the estimation of a normal mean.

The second order approximations described above exploit the independence of the sample mean and variance in the normal case in a fundamental way. Here we obtain second order approximations for the natural parameter of an exponential family, where there is no analogue of this independence. The papers by Woodroofe (1985a,b) are related.

The arguments used in our asymptotic analysis differ from those of the classical one. Instead of considering the coverage probability at fixed but arbitrary ω, a collection of average coverage probabilities over ω in a neighborhood of a fixed but arbitrary ω_0 are considered; that is, if Π is a density on Ω°, the average coverage probability under Π is defined as follows: for $h > 0$,

$$\bar{\gamma}_h(\Pi) = \int_\Omega \gamma_h(\omega)\Pi(\omega)\,d\omega.$$

In the statement of the problem, we interpret $\simeq$ to mean that

$$\bar{\gamma}_h(\Pi) = \gamma + o(h^2) \quad \text{as } h \to 0$$

for a large collection of Π's.

The following are some rationals for considering such averages:

(1) When our problem is related to Bayesian models, the expansions for $\bar{\gamma}_h(\omega)$ are easier to obtain and simpler to work with than expansions for fixed ω. The latter may be difficult to derive as noted by Sigmund (1985) and often very complicated to work with in the case where obtainable.

(2) Average confidence levels may provide a better measure of frequentist properties. To see how, suppose that an experiment produce an outcome Y from which a confidence set $C(Y)$ for an unknown parameter ω is to be constructed. Let $\gamma(\omega) = P_\omega\{\omega \in C(Y)\}$ be the coverage probability when ω is the state of nature. Suppose now that the experiment is repeated N times with parameters ω_i and outcomes Y_i, $i = 1, 2, \ldots, N$. Suppose also that the ω_i's are generated by a random process with distribution G, then the expected relative frequency of coverage is

$$\int_\Omega \gamma(\omega) G(\mathrm{d}\omega) = \bar{\gamma}(G)$$

which is a first approximation to the actual frequency of coverage by the law of large numbers. Having good frequentist properties would require that $\bar{\gamma}(G)$ is large. Requiring $\bar{\gamma}(G) \geq \gamma$ for all G is equivalent to requiring $\gamma(\omega) \geq \gamma$ for all ω. However, if the inequality is replaced by an approximation, then the two conditions ($\bar{\gamma}(G) \approx \gamma$ for all G and $\gamma(\omega) \approx \gamma$ for all ω) may be quite different as noted by Woodroofe (1985a, Section 6).

This paper proceeds as follows. A sequential procedure for estimating the unknown parameter ω of an exponential family by the maximum likelihood estimate to accuracy $[-h, h]$ for given $h > 0$ is defined in Section 2 and studied in Sections 3 and 4. A numberical example will be worked in Section 5.

2. The procedure

Let Φ and ϕ denotes the standard normal distribution and density functions, and let c solve the equation

$$2\Phi(c) - 1 = \gamma.$$

Let also $\hat{\sigma}_n$, $n \geq 1$, be the asymptotically efficient sequence of positive estimators of σ defined above. We seek a stopping time t for which

$$\gamma_h(t, \omega) = P_\omega\{|\omega - \hat{\omega}_t| \leq h\} \geq \gamma, \tag{2.1}$$

i.e.

$$P_\omega\{\sqrt{t}\hat{\sigma}_t |\omega - \hat{\omega}_t| \leq \sqrt{t}\hat{\sigma}_t h\} \geq \gamma.$$

Had a sample of fixed (non-random) size n been taken, we would achieve approximately the precision needed if

$$\sqrt{n}\sigma h \geq c, \quad \text{i.e.} \quad n \geq \frac{c^2}{\sigma^2 h^2}.$$

For unknown σ, our procedure consists of estimating σ at each stage n and stopping as soon as $n \geq c^2/(\hat{\sigma}_n^2 h^2)$. To avoid underestimation at the termination, we shall consider a simple modification which continue sampling until $n \geq c_n^2/(\hat{\sigma}_n^2 h^2)$, where

$$c_n = c\sqrt{1-(1/n)\hat{g}_n}, \tag{2.2}$$

and $\hat{g}_n = g_n(\hat{\omega}_n)$ for all $n = 1, 2, \ldots$, where g_n are real valued functions on $\bar{\Omega}$ for which $g_n > -n$. Hence our stopping time will be of the form

$$t_h = \inf\left\{n \geq m : n \geq \frac{c_n^2}{\hat{\sigma}_n^2 h^2}\right\}, \tag{2.3}$$

where m is a pilote sample size. If we let $a = c^2/h^2$, then the stopping time may be written as

$$t_a = \inf\{n \geq m : Z_n \geq a\}, \tag{2.4}$$

where

$$Z_n = nc^2\hat{\sigma}_n^2/c_n^2. \tag{2.5}$$

This form will allow us to deduce properties of the stopping time t from the non-linear renewal theorem of Lai and Sigmund (1977, 1979). To see how, note that Z_n, $n \geq 1$, may be written in the form

$$Z_n = T_n + \xi_n,$$

where

$$T_n = n\psi''(\omega) + \frac{\psi'''(\omega)}{\omega''(\omega)}[S_n - n\psi'(\omega)], \qquad \xi_n = Z_n - T_n. \tag{2.6}$$

for $n = 1, 2, \ldots$. Since each T_n, $n \geq 1$, is a random walk, this is of the form considered in the non-linear renewal theorem.

The following condition will be needed: let Ω_0 be a compact subset of the natural parameter space,

$$E_{\Omega_0}(f) = \sup_{\omega \in \Omega_0} E_\omega(f),$$

for non-negative random variables f, possibly depending on ω.

Condition A. The functions $\hat{g}_n$ in (2.2) are continuous functions on $(-\infty, \infty)$ for which $\hat{g}_n \geq -\frac{1}{2}n$ on $(-\infty, \infty)$, $\hat{g}_n \to g$ uniformly on compacts, and

$$E_{\Omega_0}\left[\sup_n \hat{g}_n^4\right] < \infty, \tag{2.7}$$

on all compact subset of Ω_0 of $\Omega°$.

Lemma 2.1. *If Ω_0 is a compact subset of $\Omega°$, then*

$$E_{\Omega_0}\left[\sup_{k>n}|U_n(\hat{\omega}_n)-\psi''(\omega)|^p\right]\to 0 \quad \text{as } n\to\infty$$

for all $p>0$.

Proof. See Woodroofe (1985b).

The key idea in what follows is to relate average coverage probabilities to Bayesian models. To see how, let Π be a density and t a stopping time; then

$$\bar{\gamma}(t,\Pi)=\int_\Omega \gamma(t,\omega)\Pi(\omega)\,\mathrm{d}\omega=\int_\Omega P_\omega\{|\omega-\hat{\omega}_t|\le h\}\Pi(\omega)\,\mathrm{d}\omega.$$

If we consider the Bayesian model in which ω has prior density Π and $X_1, X_2, \ldots$ are conditionally independent and identically distributed F_ω, then

$$\bar{\gamma}(t,\Pi)=P^\Pi\{|\omega-\hat{\omega}_t|\le h\}$$

where P^Π denotes the probability in the Bayesian model above.

3. Expansions for posterior distributions

Throughout this section, Π denotes a density on $\Omega°$ which satisfies the following conditions: for some $q\ge 3$,

$$\Pi(\omega)=(\omega-\omega_0)_+^q(\omega_1-\omega)_+^q\Pi_0(\omega), \quad \omega\in\Omega°, \tag{3.1}$$

where $[\omega_0,\omega_1]=\Omega_0\subset\Omega°$ and Π_0 is a positive three times continuously differentiable function in a neighborhood of Ω_0. Let

$$\omega_k=\psi^{(k)}/\sigma^k \quad \text{and} \quad \Pi_k=\Pi^{(k)}/\Pi\sigma^k$$

where (k) denotes the k-th derivative; that is, $\psi^{(k)}(\omega)=(\mathrm{d}^k\psi/\mathrm{d}\omega^k)(\omega)$. We let also

$$w_n^*=\sqrt{n}\hat{\sigma}_n(\omega-\hat{\omega}_n).$$

Johnson (1967, 1970) derived the following expansion for the expectation of $f(\omega_n^*)$ where f is a bounded measurable function on $(-\infty,\infty)$:

$$E(f(\omega_n^*)\mid\mathscr{D}_n)=\int_{-\infty}^{\infty}f\,\mathrm{d}\Phi+\frac{1}{\sqrt{n}}\varrho_1(\hat{\omega}_n,x)+\frac{1}{n}\varrho_2(\hat{\omega}_n,x)+\frac{1}{n^{3/2}}R_n^1 \tag{3.2}$$

where R_n^1 is a remainder term, which is discussed below.

For our problem, the function f of interest is the indicator of an interval, namely $f(\cdot)=I_{(-c_n,c_n)}(\cdot)$. Then, in this case the function ϱ_1 vanishes by symmetry and the function ϱ_2 may be written as follows:

$$\varrho_2(\omega, y) = \tfrac{1}{2}Q_2(y)\Pi_2(\omega) - \tfrac{1}{6}Q_4(y)\psi_3(\omega)\Pi_1(x) - \tfrac{1}{24}Q_4(y)\psi_4(\omega) + \tfrac{1}{72}Q_6(y)\psi_3^2(\omega) \tag{3.3}$$

where

$$Q_2(y) = -2y\phi(y), \qquad Q_4(y) = -2(3y + y^3)\phi(y),$$
$$Q_6(y) = -2(15y + 5y^3 + y^5)\phi(y)$$

for $-\infty < y < \infty$.

We will next derive a bound for $\varrho_2(\hat{\omega}_n, x)$ and the remainder term R_n^1. Let A_n, $n \geq 1$, denote the events

$$A_n = \left\{\omega_0 + \frac{\operatorname{Log} n}{\sqrt{n}} \leq \hat{\omega}_n \leq \omega_1 - \frac{\operatorname{Log} n}{\sqrt{n}}\right\}. \tag{3.4}$$

Then there exist constants $K_i = K(\Pi)$, $i = 1, 2$, depending on Π but not on n, g, or $\mathscr{D}_n$ (if $|f| \leq 1$), for which

$$|\varrho_2(\hat{\omega}_n, x) I_{A_n}| \leq K_1[(\hat{\omega}_n - \omega_0)_+^2 + (\omega_1 - \hat{\omega}_n)_+^{-2}] \tag{3.5}$$

and

$$|R_n^1 I_{A_n}| \leq K_2[(\hat{\omega}_n - \omega_0)_+^{-3} + (\omega_1 - \hat{\omega}_n)_+^{-3}] \tag{3.6}$$

for all $n \geq 2$. These inequalities are similar to that derived by Ghosh, Sinha and Joshi (1982).

Lemma 3.1. *Let Π and A_n be as defined above. Then*

$$P^{\Pi}\left\{\bigcup_{k=n}^{\infty} A_k'\right\} \leq K\left(\frac{\operatorname{Log} n}{\sqrt{n}}\right)^{q+1} \tag{3.7}$$

for all $n \geq 2$, where K is some constant which depends only on Π and is independent of n; and

$$E^{\Pi}\left\{\sup_{n \geq 1} [(\hat{\omega}_n - \omega_0)_+^{-q} + (\omega_1 - \hat{\omega}_n)_+^{-q}] I_{A_n}\right\} < \infty. \tag{3.8}$$

Proof. See Woodroofe (1985a, pp. 8, 9).

The following lemma shows some properties of our stopping time which will be needed when integrating the above expansions.

Lemma 3.2. *For all $\Omega_0 = (\omega_0, \omega_1) \subset \Omega^\circ$, there exists $\eta, \eta = \eta(\Omega^\circ) > 0$, such that*

$$\lim_{a \to \infty} a \int_{\Omega_0} P_\omega\{t_a \leq a\eta\}\, \mathrm{d}\omega = 0 \tag{3.9}$$

and

$$E_{\Omega_0}\{t_a I_{[t_a>a/\eta]}\} = \mathrm{o}(1) \quad \text{as } a\to\infty. \tag{3.10}$$

Proof. Let K be a compact neighborhood of Ω_0, let $\eta = 1/4M$ where $M = \max_{\omega\in\Omega_0}\psi''(\omega)$, and let δ be the distance from Ω_0 to K'. Then

$$P_\omega\{t_a \le C \operatorname{Log} a\} \le \sum_{k\le C\operatorname{Log} a} P_\omega\{t=k\} \le \sum_{k\le C\operatorname{Log} a} P_\omega\left\{\hat{\sigma}_k^2 > \frac{a}{2k}\right\}$$

$$\le \sum_{k\le C\operatorname{Log} a}\left(\frac{2k}{a}\right)^2 E_\omega(\hat{\sigma}_k^4) = \mathrm{O}(a^{-2}\operatorname{Log}^3 a) = \mathrm{o}(1/a)$$

by Lemma 2.1; and, if $\bar{X}_k \in K$ and $k\le \eta a$, then $Z_k = kc^2\hat{\sigma}_k^2/c_k^2 \le 2kM \le \frac{1}{2}a$, so that $t>k$. So, by Bernstein's inequality,

$$P_\omega\{C\operatorname{Log} a < t_a < \eta a\} \le P_\omega\{|\bar{X}_k - \theta| \ge \delta,\ \exists k > C\operatorname{Log} a\}$$

$$\le K_0 \exp\{-\varepsilon_0 C \operatorname{Log} a\},$$

for some $K_0 = K_0(\Omega_0)$ and $\varepsilon_0 = \varepsilon_0(\Omega_0)$. By appropriately choosing C, we may make the right hand side of the above to be $\mathrm{o}(1/a)$, as $a\to\infty$. Thus, the first assertion of the lemma is established.

For the second assertion, see Woodroofe (1985b).

Lemma 3.3.

$$\frac{t_a}{a} \to \sigma^2$$

with probability one as $a\to\infty$.

Proof. Follows easily from the fact that $\xi_n/n \to 0$ w.p. 1.

Let

$$B_a = \{t_a \ge \eta a,\ \omega_0 + (\operatorname{Log} t/\sqrt{t}) \le \hat{\omega}_t \le \omega_1 - (\operatorname{Log} t/\sqrt{t})\}.$$

Then from Lemmas 3.1 and 3.2, we have

$$P^\Pi(B_a') = \mathrm{o}(1/a) \quad \text{as } a\to\infty.$$

Therefore

$$\bar{\gamma}_a(\Pi) = P^\Pi\{\sqrt{t}|\hat{\omega}_t - \omega|\sigma_t \le c_t;\ B_a\} + \mathrm{o}(1/a) = \Delta_a + \mathrm{o}(1/a), \quad \text{say.}$$

We now use the fact that posterior distributions are unaffected by optional stopping to get

$$\Delta_a = \int_{B_a} P^\Pi\{-c_t \le \hat{\omega}_t \le c_t \mid \mathscr{D}_t\}\,\mathrm{d}P^\Pi$$

$$= \int_{B_a} (2\Phi(c_t)-1)\,\mathrm{d}P^{\Pi} + \frac{1}{a}\int_{B_a}\left(\frac{a}{t}\right)\varrho_2(\hat{\omega}_t, c_t)\,\mathrm{d}P^{\Pi}$$
$$+ \frac{1}{a^{3/2}}\int_{B_a}\left(\frac{a}{t}\right)^{3/2} R_t^1\,\mathrm{d}P^{\Pi}$$
$$= \beta_0(a) + \frac{1}{a}\beta_1(a) + \frac{1}{a^{3/2}}\beta_2(a), \quad \text{say.} \tag{3.11}$$

By (3.6) and Lemma 3.1, $\beta_2(a) = \mathrm{O}(1)$ as $a \to \infty$, and hence,

$$(1/a^{3/2})\beta_2(a) = \mathrm{o}(1/a) \quad \text{as } a \to \infty.$$

Next, by the Law of large numbers and the fact that Π is twice continuously differentiable and ψ is real analytic on $\Omega°$, we get that

$$\varrho_2(\hat{\omega}_t, c_t) \to \varrho_2(\omega, c) \quad \text{as } a \to \infty,$$

since $c_t \to c$. Also by (3.5), the integrand in $\beta_1(a)$ is bounded by

$$K_2[(\hat{\omega}_t - \omega_0)_+^{-2} + (\omega_1 + \hat{\omega}_t)_+^{-2}]$$

on B_a, for all $a \geq 1$. Thus,

$$\int_{B_a} \frac{a}{t}\varrho_2(\hat{\omega}_t, c_t) \to \int_{\Omega} \sigma^2 \varrho_2(\omega, c)\Pi(\omega)\,\mathrm{d}\omega$$

by the dominated convergence theorem. Finally, to estimate $\beta_0 = \int_{B_a}(2\Phi(c_t)-1)\,\mathrm{d}P^{\Pi}$, observe that $c_t^2 \geq c^2 + \hat{g}_t/t$ and expand $\Phi(\sqrt{x})$ in a Taylor series about c^2 to get $\Phi(c_t) = \Phi(c) + c\hat{g}_t\phi(c_t'')/2t$ for some intermediate point c_t'' between c^2 and c_t^2. Thus

$$\beta_0(a) \geq \gamma + \frac{c}{a}\int_{B_a}\left(\frac{a}{t}\right)\hat{g}_t\phi(c_t'')\,\mathrm{d}P^{\Pi} + \mathrm{o}\left(\frac{1}{a}\right).$$

It is easily seen that $(a/t)\hat{g}_t(\hat{\omega}_t)\phi(c_t'') \to \sigma^2 g(\omega)\phi(c)$ in P^{Π} probability as $a \to \infty$. Finally, since the g_n's are dominated, we get

$$\int_{B_a}\left(\frac{a}{t}\right)\hat{g}_t\phi(c_t'')\,\mathrm{d}P^{\Pi} \to \phi(c)\int_{\Omega}\sigma^2 g(\omega)\Pi(\omega)\,\mathrm{d}\omega$$

as $a \to \infty$, and

$$\beta_0(a) \geq \gamma + \frac{c\phi(c)}{a}\int_{\Omega}\sigma^2 g(\omega)\Pi(\omega)\,\mathrm{d}\omega + \mathrm{o}\left(\frac{1}{a}\right) \quad \text{as } a \to \infty,$$

Substituting these two limits into (3.11) yields to the following asymptotic lower bound for Δ_a,

$$\Delta_a \geq \gamma + \frac{\phi(c)}{a}\int_{\Omega}\sigma^2[cg(\omega) + \varrho_2(\omega, c)/\phi(c)]\Pi(\omega)\,\mathrm{d}\omega + \mathrm{o}\left(\frac{1}{a}\right).$$

Thus, we have proved the following theorem:

Theorem 3.1. *If condition* A *is satisfied, then*

$$\bar{\gamma}(\Pi) \geq \gamma + \frac{\phi(c)}{a} \int_{\Omega} \sigma^2 [cg(\omega) + \varrho_2(\omega, c)/\phi(c)] \Pi(\omega)\, \mathrm{d}\omega + \mathrm{o}\left(\frac{1}{a}\right)$$

for all densities Π of the form (3.1).

A more attractive way of writing $\bar{\gamma}(\Pi)$ follows in the next corollary. This is obtained by simplifying the expression of ϱ_2 using integration by parts of its terms to get finally:

Corollary 1.

$$\bar{\gamma}(\Pi) \geq \gamma + \frac{\phi(c)}{a} \int_{\Omega} \left[cg(\omega) - \frac{1}{4}(3c + c^3)\psi_4(\omega) + \frac{1}{36}(21c + 7c^3 - c^5)\psi_3^2 \right] \sigma^2 \Pi(\omega)\, \mathrm{d}\omega + \mathrm{o}\left(\frac{1}{a}\right)$$

The functions g_n, $n \geq 1$, which appear in the definition of t_a, are design parameters. Corollary 1 indicates how they may be chosen to make the average confidence levels be at least $\gamma + \mathrm{o}(h^2)$.

Corollary 2. *Let*

$$g(\omega) = \tfrac{1}{4}(3 + c^2)\psi_4(\omega) - \tfrac{1}{36}(21 + 7c^2 - c^4)\psi_3^2.$$

Then

$$\bar{\gamma}(\Pi) \geq \gamma + \mathrm{o}\left(\frac{1}{a}\right) \quad \textit{as } a \to \infty$$

for all Π of the form (3.1).

4. Expected stopping time

In this section, an approximation to the mean of the stopping time t_a is derived. The notation and conditions of Section 2 and 3 are used throughout. Let $N_a = a\sigma^2$, $a > 0$ and let Ω_0 as defined before.

Lemma 4.1. *The process ξ_n, $n \geq 1$, is slowly changing and satisfies*

$$\sup_n E_{\Omega_0} \left\{ \max_{k<n} |\xi_{n+k}|^2 \right\} < \infty$$

and

$$\sum_{n=1}^{\infty} P\{\xi_n \leq -n\varepsilon\} < \infty \quad \forall \varepsilon > 0.$$

Proof. The slow change of ξ_n, $n \geq 1$, follows directly from Example 4.1 of Woodroofe (1982, pp. 41, 42), the second and third assertion follows from Lemma 1 of Woodroofe (1985b) with $K = \max_J |\psi_4 - \psi_3^2|$ and J a compact neighborhood of Ω_0.

Theorem 4.1. *If condition* A *is satisfied on* Ω_0, *then*

$$E_\omega(t_a) - N_a = \frac{1}{\psi''(\omega)} (\psi''(\omega) - \tfrac{1}{2}\psi^{(4)}(\omega) + \varrho_4(\omega))$$

where ϱ_a *denotes the mean of the asymptotic distribution of the excess over the boundary* $R_a = Z_{t_a} - a$.

Proof. Follows from Theorem 4.5 of Woodroofe (1982).

5. Numerical illustrations

In this section, we specialize the results of Section 3 and 4 to the Poisson distribution and present the results of numerical studies implemented sing Pascal programs in the Amdahl computer at the University of Michigan. We present the results of simulation studies for the sequential procedure presented in Section 2 when the observations come from a Poisson distribution with mean θ. In the case where the observations come from a Poisson process, the exact distribution of the stopping time may be found and exact computations are presented. The form of the distribution is given below. Observe that in the Poisson case, $\omega = \operatorname{Log} \theta$; so fixed width accuracy for ω is fixed proportional accuracy for the mean θ.

5.1. Simulations

Throughout our simulations we implemented the procedure of Section 2 for the family of Poisson densities

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

with a grid of values of θ: $0.97, 0.98, 0.99, 1.0, 1.01, 1.02, 1.03$. For each value of θ, we obtain estimates of the coverage probability of various choices of h, where $2h$ is the length of the confidence interval.

Throughout the simulation we start with an initial sample size $m = 1$ and proceed sequentially until the rule given in (2.4) indicates that sampling terminates. For each pair (θ, h) we repeated this procedure a total of 10 000 times. Each time we ceased sampling, we determined the observed value of t, say $t(i)$, the estimate of $\theta, \theta(i)$, then counted the number S of these estimates that fall inside the confidence interval.

Throughout the tables,

$$\bar{t}=\frac{1}{N}\sum_{i=1}^{N} t(i), \qquad SD_t=\sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(t(i)-E(t))^2},$$

$$\hat{p}=\frac{S}{N} \quad \text{and} \quad \hat{\sigma}(p)=\sqrt{\frac{\hat{p}(1-\hat{p})}{N}},$$

with $N=10\,000$. We quite naturally take $\hat{p}$ as an estimate of our coverage probability and estimate its standard deviation by $\hat{\sigma}(p)$. We also take $\bar{t}$ as an estimate of the expected value of our stopping time, and estimate its variance by $SD_t/\sqrt{N}$. The desired coverage probability γ has been set at 0.95. If the procedure performs well, we would expect that $\hat{p}\approx\gamma$ and the variations would be small.

The results indicate good agreement with the asymptotic formulas of Section 3. Thus, the procedure accomplishes its objectives: the values of $\hat{p}$ are close to 0.95. They exhibit small variations but are within two standard deviation for the most part. It's evident that within a row (i.e. for fixed values of h) the values of $\hat{p}$ show some oscillations as θ varies. These seem to get smaller as h decreases. Woodroofe and Keener (1985) show for a related problem that such oscillations are of order h. This fact provide an important rational for considering average coverage probabilities instead of regular coverage probabilities.

Table 1
Simulated coverage probabilities

	λ						
h	0.97	0.98	0.99	1.00	1.01	1.02	1.03
0.25	0.9513	0.9519	0.9518	0.9483	0.9505	0.9510	0.9469
	0.0022	0.0022	0.0021	0.0022	0.0021	0.0022	0.0022
0.26	0.9510	0.9522	0.9491	0.9502	0.9468	0.9499	0.9474
	0.0022	0.0021	0.0022	0.0022	0.0021	0.0022	0.0022
0.27	0.9505	0.9508	0.9545	0.9518	0.9525	0.9489	0.9473
	0.0022	0.0022	0.0021	0.0021	0.0022	0.0022	0.0022
0.28	0.9456	0.9507	0.9522	0.9545	0.9485	0.9446	0.9525
	0.0021	0.0022	0.0021	0.0021	0.0022	0.0022	0.0021
0.29	0.9467	0.9559	0.9542	0.9482	0.9553	0.9481	0.9470
	0.0022	0.0022	0.0021	0.0022	0.0022	0.0021	0.0021
0.30	0.9533	0.9480	0.9568	0.9487	0.9498	0.9555	0.9514
	0.0021	0.0022	0.0020	0.0022	0.0022	0.0022	0.0022
0.40	0.9507	0.9531	0.9590	0.9563	0.9529	0.9501	0.9513
	0.0022	0.0022	0.0020	0.0020	0.0021	0.0021	0.0022
0.50	0.9582	0.9588	0.9558	0.9523	0.9571	0.9555	0.9516
	0.0022	0.0022	0.0021	0.0020	0.0022	0.0021	0.0022

The upper entry is the simulated value of the average coverage probability, based on 10 000 replications; the lower entry is its standard deviation; $2h$ is the length of the confidence interval.

Next, recall from (2.4) that

$$t_a = \inf\{n \geq m: Z_n \geq a\},$$

where $Z_n = nc^2\hat{\sigma}_n^2/c_n^2$. In our case $\hat{\sigma}_n^2 = \bar{X}_n$ and $c_n^2 = c^2(1+g/S_n)$, where g is a constant. Therefore

$$Z_n = n\bar{X}_n/(1+g/S_n) = S_n^2/(g+S_n)$$
$$= S_n - gS_n/(g+S_n) = S_n + \xi_n.$$

Theorem 4.1 provides us with an approximation of the asymptotical expected sample size to be used to compare with the results of our Monte-Carlo study in Table 2. It shows very good agreement between the average simulated stopping time

Table 2
Expected stopping times

h	λ 0.97	0.98	0.99	1.00	1.01	1.02	1.03
0.25	65.54	64.81	64.15	63.49	62.93	62.29	61.72
	0.082	0.081	0.079	0.079	0.079	0.078	0.078
	65.16	64.61	63.87	63.85	62.64	62.04	61.45
0.26	60.35	59.72	59.08	58.53	57.98	57.37	56.87
	0.079	0.078	0.077	0.076	0.076	0.075	0.074
	60.38	59.78	59.19	58.61	57.27	57.49	56.18
0.27	56.22	55.64	55.04	54.48	54.02	53.46	52.48
	0.076	0.075	0.073	0.074	0.070	0.073	0.072
	56.12	55.56	55.02	54.48	53.96	53.44	52.94
0.28	52.08	51.58	51.01	50.50	50.05	49.52	49.09
	0.073	0.072	0.071	0.070	0.068	0.070	0.069
	52.31	51.79	51.28	50.79	50.30	49.82	49.35
0.29	48.99	48.51	47.98	47.50	47.09	46.58	46.17
	0.071	0.070	0.069	0.068	0.068	0.068	0.067
	48.89	48.40	47.93	47.46	47.01	46.56	46.13
0.30	45.88	45.43	44.96	44.48	44.12	43.65	43.24
	0.069	0.068	0.067	0.067	0.066	0.066	0.065
	45.80	45.35	44.90	44.47	44.04	43.63	43.22
0.40	26.29	26.09	25.77	25.51	25.38	25.07	24.85
	0.052	0.051	0.050	0.050	0.050	0.050	0.049
	26.55	26.29	26.04	25.80	25.55	25.32	25.09
0.50	18.04	17.89	17.57	17.49	17.44	17.23	17.11
	0.043	0.042	0.042	0.042	0.042	0.040	0.040
	17.64	17.47	17.31	17.15	17.00	16.84	16.70

In each entry, the upper figure is the average simulated stopping time; the middle figure is its standard deviation and the lower figure is the asymptotic value of the stopping time obtained using Theorem 4.1.

and the approximate value of the stopping time obtained using the asymptotic theory.

5.2. Exact computations

One interesting aspect of the Poisson case is that the distribution of the stopping time can be derived and used to compute exact coverage probabilities. To see how, recall that

$$t_h = \inf\{t \geq 0\colon t \geq c_t^2/(\sigma_t^2 h^2)\} = \inf\{t \geq 0\colon N_t \geq c_t^2/h^2\},$$

where $N_t = \hat{\sigma}^2 t = \hat{\theta}_t t = X_1 + X_2 + \cdots + X_t$. We may compute exact coverage probabilities for the related problem in which t is replaced by a continuous parameter and N_t by a Poisson process. These calculation are described. Recall from (2.2) that

$$c_t^2 = c^2(1 + g_t/t) = c^2(1 + g/t\hat{\theta}_t) = c^2(1 + g/N_t),$$

where g is given by Corollary 2; that is,

$$g = \frac{1}{36}(6 + 2c^2 + c^4) \quad \text{and} \quad c = \Phi^{-1}\left(\frac{1+\gamma}{2}\right).$$

Hence, the stopping time can be written as

$$t_h = \inf\left\{t \geq 0\colon N_t \geq \frac{c^2}{h^2}(1 + g/N_t)\right\} = \inf\{t \geq 0\colon N_t \geq b\},$$

where

$$b = [(a + \sqrt{a^2 + 4ag})/2] + 1 \quad \text{and} \quad a = c^2/h^2,$$

and $[\cdot]$ denotes the greatest integer function. So t_h is distributed as a Gamma with parameter b and θ; that is,

$$t_n \sim \Gamma(b, \theta) \quad \text{and} \quad \hat{\theta}_t = b/t.$$

Now

$$P_\theta\left\{e^{-h} \leq \frac{\theta}{\hat{\theta}_t} \leq e^h\right\} = P_\theta\left\{e^{-h} \leq \frac{\theta t}{b} \leq e^h\right\}$$
$$= P_1\left\{e^h \leq \frac{t}{b} \leq e^h\right\} = G(be^{-h}) - G(be^h)$$

where

$$G(u) = P_1\{t_h > u\} = \sum_{j=0}^{b-1} \frac{1}{j!} u^j e^{-u}.$$

The results of the computations follow in Table 3.

Table 3
Exact coverage probabilities

h	prob. of coverage
0.16	0.9501
0.17	0.9502
0.18	0.9506
0.19	0.9508
0.20	0.9502
0.21	0.9501
0.22	0.9511
0.23	0.9509
0.24	0.9508
0.25	0.9513
0.26	0.9507
0.27	0.9510
0.28	0.9504
0.29	0.9512
0.30	0.9513
0.40	0.9509
0.50	0.9555

As we see the results are remarkably impressive. This provides us with a good check of our simulations and give an idea of how good our procedure is in this particular case.

Acknowledgements

All praises due to Allah. I am greatly indebted to my advisor, professor Michael B. Woodroofe for introducing me to the problem as well as for his guidance, help, patience and, above all his care.

References

Anscombe, F.G. (1953). Sequential estimation. *J. Roy. Statist. Soc.* 15, 1-21.

Chow, Y.S. and H. Robbins (1965). On the asymptotic theory of fixed width confidence for the mean. *Ann. Math. Statist.* 36, 457-462.

Ghosh, J.K., B. Sinha and S. Joshi (1982). Expansions for posterior probability and integrated Bayes risk. In: S. Gupta and J. Berger, Eds., *Statistical Decision Theory and Related Topics III*, Vol. 1. Academic Press, New York, 403-456.

Hall, P. (1981). Asymptotic theory of triple sampling for sequential estimation of the mean. *Ann. Statist.* 9, 1229-1238.

Johnson, R.A. (1967). An asymptotic expansion for posterior distribution. *Ann. Math. Statist.* 38, 1899-1907.

Johnson, R.A. (1970). Asymptotic expansions associated with posterior distributions. *Ann. Math. Statist.* 41, 851-864.

Lai, T.L. and D. Siegmund (1977). A nonlinear renewal theorem with applications to sequential analysis I. *Ann. Statist.* 5, 946–954.

Lai, T.L. and D. Siegmund (1979). A nonlinear renewal theorem with applications to sequential analysis II. *Ann. Statist.* 7, 60–76, 389–402.

Siegmund, D. (1984). *Sequential Analysis.* Springer, New York.

Simons, G. (1968). On the cost of not knowing the variance when making a fixed width confidence interval for the mean. *Ann. Math. Statist.* 39, 1946–1952.

Starr, N. (1966). The performance of a sequential procedure for the fixed width interval estimation of the mean. *Ann. Math. Statist.* 37, 36–50.

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

Stein, C. (1949). Some problems in sequential estimation (abstract). *Econometrica* 17, 77–78.

Stein, C. (1981). On the coverage probability of confidence sets based on a prior distribution. Technical Report, Statistics Department, Stanford University.

Woodroofe, M. (1977). Second order approximation for sequential point and interval estimation. *Ann. Statist.* 5, 984–995.

Woodroofe, M. (1982). *Nonlinear Renewal Theory in Sequential Analysis.* Soc. Indust. Appl. Math., Philadelphia, PA.

Woodroofe, M. (1985a). Very weak expansions for sequential confidence levels. Technical Report, Statistics Department, University of Michigan.

Woodroofe, M. (1985b). Confidence interval with fixed proportional accuracy. Technical Report, Statistics Department, University of Michigan.

Woodroofe, M. (1985c). Asymptotically optimal sequential point estimation in three stages. Technical Report, Statistics Department, University of Michigan.