

Metrika (2014) 77:617–634  
DOI 10.1007/s00184-013-0456-6

---

# Distributions of stopping times in some sequential estimation procedures

Alicja Jokiel-Rokita · Ryszard Magiera

Received: 30 December 2012 / Published online: 17 August 2013  
© The Author(s) 2013. This article is published with open access at Springerlink.com

**Abstract** A class of sequential estimation procedures is considered in the case when relevant data may become available only at random times. The exact distributions of the optimal stopping time and the number of observations at the moment of stopping are derived in some sequential procedures. The results obtained in an explicit form are applied to derive the expected time of observing the process, the average number of observations and the expected loss of sequential estimation procedures based on delayed observations. The use of the results is illustrated in a special model of normally distributed observations and the Weibull distributed lifetimes. The probabilistic characteristics are also derived for an adaptive sequential procedures and the behavior of the adaptive procedure is compared with the corresponding optimal sequential procedure.

**Keywords** Stopping time · Optimal stopping · Bayes sequential estimation · Distribution of a stopping time · Boundary crossing probability

## 1 Introduction

In many practical problems, for example in reliability or in clinical research, data are available only at random times. Namely, it is assumed that we observe simultaneously  $n$  objects which, for example, undergo failure at random times. In these random times we observe the realizations of random variables from a distribution  $P_{\vartheta} \in \mathcal{P}$ , where  $\mathcal{P}$  is a given parametric family of distributions. It is also assumed that the distribution  $P_{\vartheta}$  is independent of a distribution of failures. For example in papers of [Starr et al.](#)

---

A. Jokiel-Rokita (✉) · R. Magiera  
Institute of Mathematics and Computer Science, Wrocław University of Technology,  
Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland  
e-mail: [alicja.jokiel-rokita@pwr.wroc.pl](mailto:alicja.jokiel-rokita@pwr.wroc.pl)

(1976), Magiera (1982), Jokiel-Rokita and Magiera (1999, 2010) and Jokiel-Rokita (2006) a class of problems of estimating the unknown parameter  $\vartheta$  in some statistical models is considered on the basis of such data which are called delayed observations. In the papers mentioned above it is assumed that if observation is stopped at time  $t$ , the loss incurred is given by

$$\mathcal{L}_t(\vartheta, d) = \mathcal{L}(\vartheta, d) + c_A K(t) + c(t), \quad (1)$$

where the function  $\mathcal{L}(\vartheta, d)$  denotes the loss associated with the estimation error, when  $\vartheta$  is the true value of the parameter and  $d$  is the chosen estimate. The function  $c(t)$  represents the cost of observing the process up to time  $t$ ,  $c_A \geq 0$  is the cost of taking one observations and  $K(t)$  is the number of observations which have been made by time  $t$ . In the papers mentioned above the sequential procedures  $(\tau, d(\tau))$ , where  $\tau$  is the optimal stopping time and  $d(\tau)$  is the corresponding terminal decision (estimator) are derived in two cases:

1. when the distribution of the random variables which determine the moments of recording the observations is exactly known (papers of Starr et al. 1976; Magiera 1982; Jokiel-Rokita and Magiera 1999, 2010; Jokiel-Rokita 2006);
2. when the distribution of the random variables which determine the moments of obtaining the observations is exponential with an unknown parameter (papers of Starr et al. 1976; Jokiel-Rokita and Magiera 1999; Jokiel-Rokita 2006) or it belongs to a time transformed exponential family with an unknown parameter (paper Jokiel-Rokita and Magiera 2010).

In the papers of Starr et al. (1976) and Jokiel-Rokita and Magiera (1999) adaptive strategies are also proposed which require knowledge of neither  $n$  nor  $F$ . These strategies perform nearly as well as it is possible when  $n$  is large for large class of  $F$ .

For the first time the problem of estimation from delayed observations was considered by Starr et al. (1976). They derived, among others, the sequential procedures (Bayes and minimax) in the problem of estimating a mean of a normal distribution with known variance under quadratic loss and linear cost for observation time. However in their paper, as in the other papers concerning the problem of estimation from delayed observations, the distributions of the optimal stopping times in the obtained procedures were not derived. Although in the papers mentioned above the problems of estimation of an unknown parameter were considered in different statistical models, the optimal sequential procedures obtained are determined by the stopping times which can be given in the following form

$$\tau_0 = \inf \left\{ t \geq 0 : [n - K(t)][h(K(t)) - h(K(t) + 1)] \leq \frac{c'(t)}{\rho(t)} \right\}, \quad (2)$$

where the function  $h(k)$  depends on the model considered,  $\rho(t)$  is a failure rate of the random variables which determine the moment of recording the observations.

In the statistical literature the problems of deriving distributions of stopping times were considered mainly in the context of estimation of the unknown parameter under so called purely sequential sampling scheme. The determination of the stopping time for

sequential estimation in a frequentist context was first considered by [Robbins \(1959\)](#). He obtained the exact distribution of the stopping time in a problem of estimating the mean of a normal population with an unknown variance, but for a “two-item at a time” sampling plan with a recurrence relation for its probabilities. In the purely sequential sampling scheme the optimal stopping time of a procedure frequently turns out to be the smallest positive integer  $k$  for which the sum  $S_k = X_1 + \dots + X_k$  crosses a boundary  $f(k)$ , where  $X_1, X_2, \dots$  is a sequence of independent and identically distributed random variables with finite variance. In the asymptotic theory of sequential analysis instead of a fixed boundary  $f(k)$ , we have a set of boundaries  $f_c(k)$  indexed by a parameter  $c$  tending to 0. [Siegmund \(1968\)](#) developed asymptotic distributions of one-sided stopping times. In sequential analysis the famous Anscombe’s random central limit theorem ([Anscombe 1952](#)) also gives a tool for obtaining asymptotic distributions for appropriately standardized stopping times (see, [Mukhopadhyay and Chattopadhyay 2012](#); [Gut 2012](#)). In the paper of [Bhattacharya and Mallik \(1973\)](#) the asymptotic distribution of the stopping time of the procedure due to [Robbins \(1959\)](#) is shown to be normal. Their result was generalized very substantially in a paper of [Ghosh and Mukhopadhyay \(1975\)](#). In the paper of [Woodrooffe \(1977\)](#) the second order approximations are obtained for the expected sample size and the risk of some sequential estimation procedures, for example, the sequential procedures proposed by [Robbins \(1959\)](#), [Chow and Robbins \(1965\)](#), [Starr \(1966a,b\)](#) and [Starr and Woodrooffe \(1972\)](#). [Pham-Gia \(1998\)](#) presented a method for computing the distribution of the stopping time for the Bayesian sequential one step look-ahead procedure in the problem of estimating the parameter of a Bernoulli experiment under a squared error loss. Also under purely sequential sampling schemes, the problems of exact determination of the distributions of stopping rules and risks of some sequential procedures were considered in papers of [Zacks and Mukhopadhyay \(2006, 2007\)](#), [Mukhopadhyay and Zacks \(2007\)](#) and [Zacks \(2005, 2009\)](#).

In this paper, we derive the exact distribution of the optimal stopping time  $\tau_0$ , given by (2), and of the number of observations  $K(\tau_0)$  at the moment of stopping. We then apply our results to obtain exact formulas for the expected time of observing the process, for the average number of observations at the moment of stopping and for the expected loss (risk) of some sequential procedures. We also compare the risks of the optimal strategies with the risks of an adaptive strategy in estimation problems in a special model.

## 2 The model

Let us assume that the observations  $Y_1, \dots, Y_n$ , where  $n$  is fixed and known, are obtained at random times  $t_1, \dots, t_n$ , which are the order statistics of positive, independent and identically distributed random variables  $T_1, \dots, T_n$  which are independent of the observations  $Y_1, \dots, Y_n$ . The random variables  $T_1, \dots, T_n$  can be interpreted as the lifetimes of  $n$  objects. We assume that the random variables  $T_1, \dots, T_n$  have a known distribution function  $F$  and the random variables  $Y_1, \dots, Y_n$  are independent and have a common distribution  $P_\vartheta$  with unknown parameter  $\vartheta$ . The aim is to find a sequential procedure  $(\tau, d(\tau))$ , where  $\tau$  is a stopping time and  $d(\tau)$  is a terminal

decision function, in the problem of estimating the unknown parameter  $\vartheta$  under a loss function given by (1), where

$$K(t) = \sum_{i=1}^n 1_{[0,t]}(T_i), \quad t \geq 0,$$

is the number of observations which have been made by time  $t \geq 0$ . Usually the form of the terminal decision function, for example the Bayes decision function, can be easily given. Then the problem reduces to determining the optimal stopping time. Let

$$\mathcal{F}_t = \sigma\{K(s), s \leq t, Y(1), \dots, Y(K(t))\},$$

be the information which is available at time  $t$ . We assume that if we observe the process for  $t$  units of time, then the loss incurred by time  $t$  is determine by the conditional expected loss, given  $\mathcal{F}_s, s \leq t$ , having the following form

$$\mathcal{L}_h(t) = h(K(t)) + c(t), \tag{3}$$

where  $h$  be a given real valued function on  $E_N = \{0, 1, \dots, N\}$ , and such that  $0 \leq h(k) < \infty$  for each  $k \in E_N$ . The function  $c(t)$  is assumed to be a differentiable and non-decreasing function with  $c(0) = 0$ . The function  $\mathcal{L}_h(t)$  can be interpreted as the total loss incurred if the process is stopped at time  $t$ . For example, the conditional expected losses of Bayes or minimax estimators in the statistical models considered in the papers of Starr et al. (1976), Magiera (1982), Jokiel-Rokita and Magiera (1999, 2010) and Jokiel-Rokita (2006) are of the form given by (3). By the optimal stopping time  $\tau_0$  with respect to  $\mathcal{F}_t, t \geq 0$ , we mean the random variable which minimizes

$$\mathcal{V}_h(\tau) = E[\mathcal{L}_h(\tau)] = E[h(K(\tau)) + c(\tau)],$$

over all stopping times  $\tau$  with respect to  $\mathcal{F}_s, s \leq t$ .

Suppose that the distribution function  $F$  satisfies the following conditions:  $F(0) = 0$ ;  $F(t) > 0$  for  $t > 0$ ;  $F$  is absolutely continuous with density  $f$ ; and  $f$  is the right hand derivative of  $F$  on  $(0, \infty)$ . The class of such  $F$  will be denoted by  $\mathcal{G}$ . Let  $\zeta = \sup\{t : F(t) < 1\}$ , and  $\rho(t) = f(t)[1 - F(t)]^{-1}, 0 \leq t < \zeta$ , denote the failure rate. The process  $K(t), 0 \leq t \leq \zeta$ , is a non-stationary Markov chain with respect to  $\mathcal{F}_t, 0 \leq t \leq \zeta$ , and its infinitesimal operator is

$$\mathcal{A}_t h(k) = (n - k)[h(k + 1) - h(k)]\rho(t),$$

for  $k \in E_n = \{0, 1, \dots, n\}$  and all real valued functions  $h$  on  $E_n$  (see, for instance Starr et al. 1976, p. 104).

**Lemma 1** (Jokiel-Rokita and Magiera 2010, Lemma 1) *Suppose that  $h(k) - h(k + 1)$  is non-increasing for  $k \leq n - 1$  and that  $F \in \mathcal{G}$ . Moreover, assume that the function  $c'(t)/\rho(t)$  is non-decreasing. Then the stopping time*

$$\begin{aligned} \tau_0 &= \inf\{t \geq 0 : \mathcal{A}_t h(K(t)) + c'(t) \geq 0\} \\ &= \inf\{t \geq 0 : [n - K(t)][h(K(t)) - h(K(t) + 1)] \leq c'(t)/\rho(t)\}, \end{aligned}$$

minimizes  $\mathcal{V}_h(\tau)$ .

*Remark 1* Under the assumptions concerning the functions  $h(k)$  and  $c'(t)/\rho(t)$  the so-called monotone case holds: once the infinitesimal prospect of the loss  $\mathcal{L}_h$  for the future becomes bad (greater than zero), it remains bad. If the loss  $\mathcal{L}_h$  is in the monotone case, the optimality of  $\tau_0$  is obtained by applying Dynkin’s identity to the loss function  $\mathcal{L}_h(t)$ , and it gives the formula

$$E[\mathcal{L}_h(\tau)] - h(0) = E \left\{ \int_0^\tau [\mathcal{A}_t h(K(t)) + c'(t)] dt \right\},$$

for any stopping time  $\tau$ .

*Remark 2* If the function  $h(\cdot)$  is convex, then  $h(k) - h(k + 1)$  is non-increasing.

The main goal of this paper is to determine the exact distribution of the optimal stopping time  $\tau_0$  and of the number of observations  $K(\tau_0)$  at the moment of stopping in the sequential procedures considered.

### 3 The distribution and the expectation of the optimal stopping time

Under the assumption of Lemma 1 one can show the following lemma.

**Lemma 2** *If the functions  $h$  and  $c'(t)/\rho(t)$  satisfy the assumptions of Lemma 1, then the optimal stopping time  $\tau_0$  can be presented in the following form*

$$\tau_0 = \inf \{t \geq 0 : (K(t) \geq b(t)) \vee (K(t) = n)\}, \tag{4}$$

where the stopping boundary  $b(t)$  is a non-increasing and right continuous function. Moreover, if  $c'(t)/\rho(t) = \text{const}$  then  $b(t) = \text{const}$ .

In the sequel we will only consider the case  $b(0) > 0$ , because  $b(0) \leq 0$  implies  $P(\tau_0 = 0) = 1$ .

Let us notice that the stopping time given by (4) can be viewed as the time of the first meeting (crossing) of the sample process  $K(t), t \geq 0$ , with upper (usually non-linear) boundary  $b(t)$ .

Denote

$$b_i = \begin{cases} \inf\{t \geq 0 : b(t) \leq i\}, & \text{when } b(t_0) \leq i \text{ for some } t_0 \geq 0, \\ \infty, & \text{when } b(t) > i \text{ for } t \geq 0, \end{cases} \tag{5}$$

$i = 0, 1, \dots, n$ . The function  $b(t)$  is non-increasing and hence we have  $b_0 \geq b_1 \geq \dots \geq b_n$ . If there exists  $i_0 \in \{1, \dots, n\}$  such that  $i_0 - 1 < b(t) \leq i_0$  for each  $t \geq 0$ ,

then we put  $b_0 = \infty, b_1 = \infty, \dots, b_{i_0-1} = \infty, b_{i_0} = 0, \dots, b_n = 0$ . In the case when  $b(t) > n$  for each  $t \geq 0$ , we put  $b_0 = b_1 = \dots = b_n = \infty$ . In the case  $b(t) > n - 1$  for each  $t \geq 0$ , and  $b_n \neq 0$ , we put  $b_0 = b_1 = \dots = b_{n-1} = \infty$ .

Denote

$$\mathcal{D} = \{b_{k_0}, \dots, b_{l_0}\},$$

where

$$k_0 = \max\{0 \leq i \leq n - 1 : b_i \neq 0\},$$

and

$$l_0 = \min\{0 \leq i \leq n - 1 : b_i \neq \infty\}.$$

Let us notice that if there exists  $i_0 \in \{1, \dots, n\}$  such that  $i_0 - 1 < b(t) \leq i_0$  for each  $t \geq 0$ , or  $b(t) > n - 1$  for each  $t \geq 0$ , then  $\mathcal{D} = \emptyset$ .

**Theorem 1** *In the case when  $\mathcal{D} \neq \emptyset$ , the distribution function  $F_{\tau_0}(t)$  of the stopping time  $\tau_0$  given by (4) is of the form*

$$F_{\tau_0}(t) = \begin{cases} \sum_{j=k_0+1}^n \binom{n}{j} F^j(t)[1 - F(t)]^{n-j}, & t < b_{k_0}, \\ \sum_{j=i}^n \binom{n}{j} F^j(t)[1 - F(t)]^{n-j}, & t \in [b_i, b_{i-1}), \quad i = k_0, k_0 - 1, \dots, l_0 + 1, \\ \sum_{j=l_0}^n \binom{n}{j} F^j(t)[1 - F(t)]^{n-j}, & t \geq b_{l_0}, \end{cases} \tag{6}$$

$$= \begin{cases} \frac{n!}{k_0!(n - k_0 - 1)!} \int_0^{F(t)} x^{k_0} (1 - x)^{n-k_0-1} dx, & t < b_{k_0}, \\ \frac{n!}{(i - 1)!(n - i)!} \int_0^{F(t)} x^{i-1} (1 - x)^{n-i} dx, & t \in [b_i, b_{i-1}), \\ & i = k_0, k_0 - 1, \dots, l_0 + 1, \\ \frac{n!}{(l_0 - 1)!(n - l_0)!} \int_0^{F(t)} x^{l_0-1} (1 - x)^{n-l_0} dx, & t \geq b_{l_0}. \end{cases} \tag{7}$$

*In the case when there exists  $i_0 \in \{1, \dots, n\}$  such that  $i_0 - 1 < b(t) \leq i_0$  for each  $t \geq 0$ , the distribution function  $F_{\tau_0}(t)$  of the stopping time  $\tau_0$  is of the form*

$$\begin{aligned} F_{\tau_0}(t) &= \sum_{j=i_0}^n \binom{n}{j} F^j(t)[1 - F(t)]^{n-j} \\ &= \frac{n!}{(i_0 - 1)!(n - i_0)!} \int_0^{F(t)} x^{i_0-1} (1 - x)^{n-i_0} dx, \end{aligned}$$

and the density function  $f_{\tau_0}(t)$  of the stopping time  $\tau_0$  is

$$f_{\tau_0}(t) = \frac{n!}{(i_0 - 1)!(n - i_0)!} [F(t)]^{i_0 - 1} [1 - F(t)]^{n - i_0} f(t).$$

In the case when  $b(t) > n - 1$  for each  $t \geq 0$ , the distribution function  $F_{\tau_0}(t)$  of the stopping time  $\tau_0$  is of the form

$$F_{\tau_0}(t) = [F(t)]^n,$$

and the density function  $f_{\tau_0}(t)$  of the stopping time  $\tau_0$  is

$$f_{\tau_0}(t) = n[F(t)]^{n-1} f(t).$$

*Proof* For  $t < b_{k_0}$  we have  $b(t) > k_0$ . Hence

$$P(\tau_0 \leq t) = P(K(t) \geq k_0 + 1) = P(X_{(k_0+1):n} \leq t).$$

For  $t \in [b_i, b_{i-1})$ ,  $i = k_0, k_0 - 1, \dots, l_0 + 1$ , it holds  $P(\tau_0 \leq t) = P(X_{i:n} \leq t)$ . For  $t \geq b_{l_0}$ , we have  $P(\tau_0 \leq t) = P(X_{l_0:n} \leq t)$ . Using the known formula for the distribution function of the appropriate order statistics, we obtain (6) and (7).  $\square$

*Remark 3* If  $\mathcal{D} = \emptyset$  then the stopping time  $\tau_0$  is a continuous type random variable; otherwise it is a mixed type random variable and  $\mathcal{D}$  is the set of discontinuities of the distribution function  $F_{\tau_0}$ . Moreover, for  $b_i \in \mathcal{D}$ ,  $P(\tau_0 = b_i) = \binom{n}{i} F^i(b_i) [1 - F(b_i)]^{n-i}$ .

**Corollary 1** In the case when  $\mathcal{D} \neq \emptyset$ , the expected value of the stopping time  $\tau_0$  is given by

$$\begin{aligned} E(\tau_0) &= \sum_{j=0}^{k_0} \binom{n}{j} \int_0^{b_{k_0}} F^j(t) [1 - F(t)]^{n-j} dt \\ &\quad + \sum_{i=k_0}^{l_0+1} \sum_{j=0}^{i-1} \binom{n}{j} \int_{b_i}^{b_{i-1}} F^j(t) [1 - F(t)]^{n-j} dt \\ &\quad + \sum_{j=0}^{l_0-1} \binom{n}{j} \int_{b_{l_0}}^{\infty} F^j(t) [1 - F(t)]^{n-j} dt \\ &= \int_0^{b_{k_0}} \left\{ 1 - \frac{n!}{k_0!(n - k_0 - 1)!} \int_0^{F(t)} x^{k_0} (1 - x)^{n - k_0 - 1} dx \right\} dx \end{aligned}$$

$$\begin{aligned}
 & + \sum_{i=k_0}^{l_0+1} \int_{b_i}^{b_{i-1}} \left\{ 1 - \frac{n!}{(i-1)!(n-i)!} \int_0^{F(t)} x^{i-1} (1-x)^{n-i} \right\} dx \\
 & + \int_{b_{l_0}}^{\infty} \left\{ 1 - \frac{n!}{(l_0-1)!(n-l_0)!} \int_0^{F(t)} x^{l_0-1} (1-x)^{n-l_0} \right\} dx. \tag{8}
 \end{aligned}$$

Using the decomposition of the mixed type distribution of the optimal stopping time  $\tau_0$  in this case on the discrete part and on the continuous part, the expected value of the stopping time  $\tau_0$  can be written in the following form

$$\begin{aligned}
 E(\tau_0) &= \sum_{i=k_0}^{l_0} b_i \binom{n}{i} F^i(b_i) [1 - F(b_i)]^{n-i} \\
 &+ \frac{n!}{k_0!(n-k_0-1)!} \int_0^{b_{k_0}} t f(t) F^{k_0}(t) [1 - F(t)]^{n-k_0-1} dt \\
 &+ \sum_{i=k_0}^{l_0} \frac{n!}{(i-1)!(n-i)!} \int_{b_i}^{b_{i-1}} t f(t) F^{i-1}(t) [1 - F(t)]^{n-i} dt.
 \end{aligned}$$

In the case when there exists  $i_0 \in \{1, \dots, n\}$  such that  $i_0 - 1 < b(t) \leq i_0$  for each  $t \geq 0$

$$E(\tau_0) = \frac{n!}{(i_0-1)!(n-i_0)!} \int_0^{\infty} t f(t) F^{i_0-1}(t) [1 - F(t)]^{n-i_0} dt.$$

In the case when  $b(t) > n - 1$  for each  $t \geq 0$

$$E(\tau_0) = n \int_0^{\infty} t f(t) F^{n-1}(t) dt.$$

*Remark 4* The expected value  $E(\tau_0)$  rarely can be written in the closed form. But if  $F(t) = 1 - \exp\{-\beta t\}$ , i.e. when  $T_1, \dots, T_n$  are exponentially distributed and if exists  $i_0 \in \{1, \dots, n\}$  such that  $i_0 - 1 < b(t) \leq i_0$  for each  $t \geq 0$ , or  $b(t) > n - 1$  for each  $t \geq 0$ , then applying the Renyi's representation (see for example [Govindarajulu 2007](#), Corollary 2.6.2.1) we obtain

$$E(\tau_0) = \begin{cases} \beta^{-1} \sum_{i=1}^n \frac{1}{n-i+1}, & \text{when } b(t) > n - 1 \text{ for each } t \geq 0, \\ \beta^{-1} \sum_{i=1}^{i_0} \frac{1}{n-i+1}, & \text{when } i_0 - 1 < b(t) \leq i_0 \text{ for each } t \geq 0. \end{cases}$$



### 4 The distribution and the expected value of the number of observations at the moment of the optimal stopping

**Theorem 2** *In the case when  $\mathcal{D} \neq \emptyset$ , the distribution of the number of observations at the moment of the optimal stopping time  $\tau_0$  is of the form*

$$P(K(\tau_0) = i) = \begin{cases} \sum_{j=k_0+1}^n \binom{n}{j} F^j(b_{k_0})[1 - F(b_{k_0})]^{n-j}, & i = k_0 + 1, \\ \binom{n}{i} \{F^i(b_i)[1 - F(b_i)]^{n-i} \\ + i \int_{b_i}^{b_{i-1}} f(t)F^{i-1}(t)[1 - F(t)]^{n-i} dt\}, & i = k_0, \dots, l_0 \neq 0, \\ (1 - F(b_{l_0}))^n, & i = l_0 = 0, \\ 0, & \text{otherwise.} \end{cases} \tag{9}$$

*In the case when there exists  $i_0 \in \{1, \dots, n\}$  such that  $i_0 - 1 < b(t) \leq i_0$  for each  $t \geq 0$ ,  $P(K(\tau_0) = i_0) = 1$ .*

*In the case when  $b(t) > n - 1$  for each  $t \geq 0$ ,  $P(K(\tau_0) = n) = 1$ .*

*Proof* Observe that  $P(K(\tau_0) = k_0 + 1) = P(X_{(k_0+1):n} < b_{k_0})$  and

$$P(K(\tau_0) = i) = P(b_i \leq \tau_0 < b_{i-1}) = P((X_{i:n} < b_{i:n}) \wedge (X_{(i+1):n} > b_{i:n})) + P(b_i \leq X_{i:n} < b_{i-1:n})$$

for  $i = k_0, \dots, l_0 \neq 0$ . Note also that  $P(K(\tau_0) = l_0 = 0) = P(X_{1:n} > b_{l_0})$ .

Using the known formula for the density function of the appropriate order statistics or formulas (6), (7), we obtain (9). □

**Corollary 2** *In the case when  $\mathcal{D} \neq \emptyset$ , the expected value of the number of observations at the moment of the optimal stopping time is of the form*

$$E(K(\tau_0)) = (k_0 + 1) \sum_{j=k_0+1}^n \binom{n}{j} F^j(b_{k_0})[1 - F(b_{k_0})]^{n-j} + \sum_{i=k_0}^{l_0} i \binom{n}{i} \{F^i(b_i)[1 - F(b_i)]^{n-i} + i \int_{b_i}^{b_{i-1}} f(t)F^{i-1}(t)[1 - F(t)]^{n-i} dt\}. \tag{10}$$

*In the case when there exists  $i_0 \in \{1, \dots, n\}$  such that  $i_0 - 1 < b(t) \leq i_0$  for each  $t \geq 0$ , then*

$$E(K(\tau_0)) = i_0.$$

*In the case  $b(t) > n - 1$  for each  $t \geq 0$ , we have  $E(K(\tau_0)) = n$ .*

## 5 The value of the risk of the optimal sequential procedures

Taking into account Theorems 1 and 2 we obtain the following corollary.

**Corollary 3** *In the case when  $\mathcal{D} \neq \emptyset$ , the risk  $\mathcal{V}_h(\tau_0)$  of the sequential procedures considered is of the form*

$$\begin{aligned}
 \mathcal{V}_h(\tau_0) = & h(k_0 + 1) \sum_{j=k_0+1}^n \binom{n}{j} F^j(b_{k_0}) [1 - F(b_{k_0})]^{n-j} \\
 & + \sum_{i=k_0}^{l_0+1} h(i) \binom{n}{i} \{F^i(b_i) [1 - F(b_i)]^{n-i} \\
 & + i \int_{b_i}^{b_{i-1}} f(t) F^{i-1}(t) [1 - F(t)]^{n-i} f(t) dt\} \\
 & + \mathbf{1}_{\{l_0 > 0\}} h(l_0) \binom{n}{l_0} \{F^{l_0}(b_{l_0}) [1 - F(b_{l_0})]^{n-l_0} \\
 & + l_0 \int_{b_{l_0}}^{\infty} f(t) F^{l_0-1}(t) [1 - F(t)]^{n-l_0} dt\} + \mathbf{1}_{\{l_0=0\}} h(0) [1 - F(b_0)]^n \\
 & + \sum_{i=k_0}^{l_0} c(b_i) \binom{n}{i} F^i(b_i) [1 - F(b_i)]^{n-i} \\
 & + \frac{n!}{k_0!(n - k_0 - 1)!} \int_0^{b_{k_0}} c(t) f(t) F^{k_0}(t) [1 - F(t)]^{n-k_0-1} dt \\
 & + \sum_{i=k_0}^{l_0} \frac{n!}{(i - 1)!(n - i)!} \int_{b_i}^{b_{i+1}} c(t) f(t) F^{i-1}(t) [1 - F(t)]^{n-i} dt. \quad (11)
 \end{aligned}$$

*In the case when there exists  $i_0 \in \{1, \dots, n\}$  such that  $i_0 - 1 < b(t) \leq i_0$  for each  $t \geq 0$ , then*

$$\mathcal{V}_h(\tau_0) = h(i_0) + \frac{n!}{(i_0 - 1)!(n - i_0)!} \int_0^{\infty} c(t) f(t) F^{i_0-1}(t) [1 - F(t)]^{n-i_0} dt.$$

*In the case  $b(t) > n - 1$  for each  $t \geq 0$ , we have*

$$\mathcal{V}_h(\tau_0) = h(n) + n \int_0^{\infty} c(t) f(t) F^{n-1}(t) dt.$$

### 6 A special case—estimating the mean of a normal distribution with known variance

In this section we give the distribution of the optimal stopping time and of the number of observations at the moment of stopping in the Bayes sequential procedure considered by Starr, Wardrop and Woodroffe (Starr et al. 1976). We also compare the risk of the optimal strategy with the risk of the adaptive strategy proposed in the paper of Starr et al. (1976). In Sect. 4 of Starr et al. (1976) the problem of estimating the mean  $\vartheta$  of a normal distribution when the observations become available at random times is considered. A Bayesian approach by placing a prior distribution over  $\vartheta$  is applied. The normal prior distribution with mean  $m_0$  and variance  $\sigma_0^2$  is assumed. Thus, conditionally, given  $\vartheta$ ,  $Y_1, \dots, Y_n$  are independent normally distributed random variables with unknown mean  $\vartheta$  and known variance  $\sigma^2$ . The random variable  $Y_i$  is observed at time  $t_i, i = 1, \dots, n$ , where  $t_1, \dots, t_n$  are the order statistics of random variables  $T_1, \dots, T_n$  which are independent of  $Y_1, \dots, Y_n$ . They assumed that if the process of observation is stopped at time  $t$  and an estimator  $d$  is used then the loss incurred is

$$\mathcal{L}_t(\vartheta, d) = (d - \vartheta)^2 + ct,$$

where  $c > 0$ . Thus, it is of the form given by (1) with  $\mathcal{L}(\vartheta, d) = (d - \vartheta)^2, c_A = 0$ , and  $c(t) = ct$ . In the case of the squared error loss, for any stopping time  $t$ , the Bayes estimator of the unknown parameter  $\vartheta$  is the mean of the posterior distribution, i.e. in the case considered it is given by

$$d(t) = \left( \sigma_0^{-2}m_0 + \sigma^{-2} \sum_{i=1}^{K(t)} Y_i \right) / \left[ \sigma_0^{-2} + K(t)\sigma^{-2} \right]$$

(see Starr et al. 1976, Lemma 4.1). The conditional expected loss, given  $\mathcal{F}_s, s \leq t$ , associated with the Bayes decision  $d$  is of the form

$$\sigma^2[\varepsilon + K(t)]^{-1} + ct,$$

where  $\varepsilon = \sigma^2/\sigma_0^2$ , i.e. it is of the form (3) with  $h(K(t)) = \sigma^2[\varepsilon + K(t)]^{-1}$ . The function  $h$  satisfies the assumption of Lemma 1, and if the function  $c/\rho(t)$  is non-decreasing, then one can show that the optimal stopping time  $\tau_0$  has the form of (4) with

$$b(t) = \frac{1}{2} \left[ -2\varepsilon - 1 - \sigma^2\rho(t)/c + \sqrt{(2\varepsilon + 1 + \sigma^2\rho(t)/c)^2 + 4n\sigma^2\rho(t)/c - 4\varepsilon(\varepsilon + 1)} \right]. \tag{12}$$

The distributions of the optimal stopping time  $\tau_0$  and the number of observations  $K(\tau_0)$  at the moment of stopping can be obtained from Theorems 1 and 2, respectively.

Let  $T_1, \dots, T_n$  be distributed according to the Weibull distribution  $\mathcal{W}(\alpha, \beta^{1/\alpha})$ , with density function

$$f(t) = \alpha\beta t^{\alpha-1} \exp\{-\beta t^\alpha\} \mathbf{1}_{(0,\infty)}(t),$$

where  $0 < \alpha \leq 1, \beta > 0$ , and the corresponding failure rate rate function

$$\rho(t) = \alpha\beta t^{\alpha-1}. \tag{13}$$

In this model, the function  $c/\rho(t)$  is non-decreasing, and  $b(t)$  is defined by (12) with  $\rho(t)$  given by (13).

When  $\alpha = 1$ , i.e.  $T_1, \dots, T_n$  are distributed according to the exponential distribution, then

$$b(t) = \frac{1}{2} \left[ -2\varepsilon - 1 - \sigma^2\beta/c + \sqrt{(2\varepsilon + 1 + \sigma^2\beta/c)^2 + 4n\sigma^2\beta/c - 4\varepsilon(\varepsilon + 1)} \right]$$

and is independent of  $t$ .

Applying Corollaries 1, 2 and 3 one can easily obtain, respectively, the expected value of the stopping time  $\tau_0$ , the expected value of the number of observations at the moment of stopping and the risk of the procedure considered.

In the subsequent part of this section we will formulate the result on the exact distribution (Theorem 3) of the adaptive stopping time determining the adaptive sequential procedure considered by Starr et al. (1976). We also give the explicit formulas for expectation of the adaptive stopping time, for the distribution of the number of observations at the moment of stopping and for the sequential risk corresponding to the adaptive procedure (Corollary 4, Theorem 4 and Corollary 6, respectively). Then, using the formulas derived we compare numerically in Table 1 the behavior of the adaptive sequential procedure with the optimal procedure.

Let  $\delta = (\tau, d(\tau))$  be a sequential procedure for which  $\tau \geq t_1$  and

$$d(\tau) = \sum_{i=1}^{K(\tau)} Y_i / K(\tau). \tag{14}$$

The conditional expected loss  $\mathfrak{L}_h(\tau)$  associated with the decision function  $d(\tau)$  given by (14) is of the form

$$\mathfrak{L}_h(\tau) = \frac{\sigma^2}{K(\tau)} + c\tau,$$

i.e. it is of the form of (3) with  $h(K(t)) = \sigma^2/K(t)$ . By Lemmas 1 and 2 the optimal stopping time is defined by

$$\begin{aligned} \tau_0 &= \inf\{t \geq 0 : \sigma^2[n - K(t)]/K(t)[K(t) + 1] \leq c/\rho(t)\} \\ &= \inf\{t \geq 0 : (K(t) \geq b(t)) \vee (K(t) = n)\}, \end{aligned}$$

**Table 1** Comparisons of the values  $V_1$  versus the values  $V_0$  with respect to the number  $n$  of devices

| $n$ | No. | $\alpha$ | $\beta$ | $\sigma$ | $E(\tau_0)$ | $E(\tau_1)$ | $E(K(\tau_0))$ | $E(K(\tau_1))$ | $V_0$   | $V_1$   | $re(V_1)$ |
|-----|-----|----------|---------|----------|-------------|-------------|----------------|----------------|---------|---------|-----------|
| 10  | 1   | 0.6      | 2       | 2        | 0.97189     | 4.03823     | 8.92768        | 19.92640       | 0.54568 | 0.80711 | 0.47910   |
|     | 2   | 0.6      | 1.5     | 1        | 0.54243     | 1.23559     | 6.64962        | 18.08730       | 0.20551 | 0.24561 | 0.19510   |
|     | 3   | 0.6      | 0.8     | 0.5      | 0.36508     | 0.60682     | 3.56333        | 5.31706        | 0.10821 | 0.11883 | 0.09817   |
|     | 4   | 0.8      | 1.5     | 1        | 0.69008     | 1.21705     | 6.99505        | 18.43171       | 0.21198 | 0.24166 | 0.13998   |
|     | 5   | 0.8      | 1       | 1        | 0.83433     | 1.39427     | 6.00794        | 17.29956       | 0.24991 | 0.27600 | 0.10440   |
|     | 6   | 0.8      | 0.5     | 1        | 1.36679     | 1.85946     | 4.89965        | 12.50963       | 0.34170 | 0.36565 | 0.07009   |
| 50  | 1   | 0.6      | 2       | 2        | 0.38454     | 0.93800     | 34.00451       | 92.82936       | 0.15615 | 0.18742 | 0.20020   |
|     | 2   | 0.6      | 1.5     | 1        | 0.20409     | 0.36206     | 22.01281       | 72.11303       | 0.06591 | 0.07222 | 0.09573   |
|     | 3   | 0.6      | 0.8     | 0.5      | 0.13110     | 0.19828     | 10.45786       | 13.10988       | 0.03714 | 0.03938 | 0.06006   |
|     | 4   | 0.8      | 1.5     | 1        | 0.28404     | 0.39560     | 21.22672       | 75.47273       | 0.07554 | 0.07884 | 0.04372   |
|     | 5   | 0.8      | 1       | 1        | 0.35915     | 0.47731     | 17.91782       | 67.10003       | 0.09176 | 0.09503 | 0.03561   |
|     | 6   | 0.8      | 0.5     | 1        | 0.53131     | 0.67220     | 13.05305       | 21.24129       | 0.12980 | 0.13352 | 0.02860   |
| 100 | 1   | 0.6      | 2       | 2        | 0.25565     | 0.53611     | 58.76521       | 174.87552      | 0.09366 | 0.10714 | 0.14394   |
|     | 2   | 0.6      | 1.5     | 1        | 0.13301     | 0.22136     | 36.07010       | 123.25676      | 0.04106 | 0.04420 | 0.07649   |
|     | 3   | 0.6      | 0.8     | 0.5      | 0.08445     | 0.12488     | 16.51022       | 20.53721       | 0.02364 | 0.02486 | 0.05181   |
|     | 4   | 0.8      | 1.5     | 1        | 0.19488     | 0.25508     | 33.41304       | 138.86232      | 0.04943 | 0.05089 | 0.02966   |
|     | 5   | 0.8      | 1       | 1        | 0.24502     | 0.31117     | 27.77191       | 116.09244      | 0.06053 | 0.06204 | 0.02502   |
|     | 6   | 0.8      | 0.5     | 1        | 0.36192     | 0.44353     | 19.89749       | 24.35918       | 0.08648 | 0.08829 | 0.02094   |

The value  $re(V_1)$  denotes the relative error of  $V_1$  with respect to  $V_0$

with the stopping boundary

$$b(t) = \frac{1}{2} \left[ - (1 + \sigma^2 \rho(t)/c) + \sqrt{(1 + \sigma^2 \rho(t)/c)^2 + 4n\sigma^2 \rho(t)/c} \right] > 0. \tag{15}$$

The adaptive stopping time proposed in the paper of [Starr et al. \(1976\)](#) is of the form

$$\tau_1 = \inf\{t \geq 0 : tK(t) \geq \sigma^2/c\} = \inf\{t \geq 0 : K(t) \geq \sigma^2/ct\}. \tag{16}$$

Let us notice that the stopping time  $\tau_1$  requires knowledge of neither  $n$  nor  $F$  for its implementation.

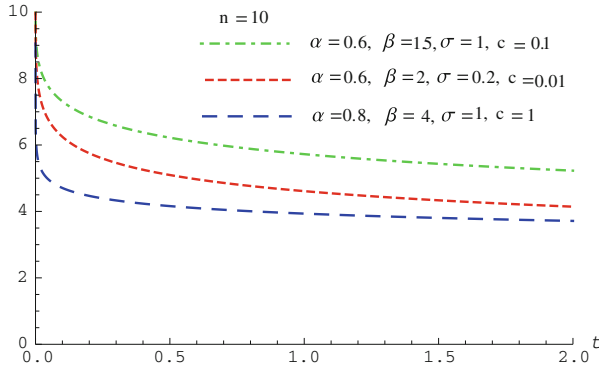
Figure 1 contains plots of the boundaries  $b(t)$  given by (15) with  $\rho(t)$  defined by (13) for some values of  $\alpha, \beta, \sigma, c$ , and  $n = 10$ . Figure 2 shows three sample paths of the process stopped at the boundary  $b(t; n, \alpha, \beta, \sigma, c)$  for  $n = 10, \alpha = 0.6, \beta = 2, \sigma = 0.2, c = 0.1$ .

Denote

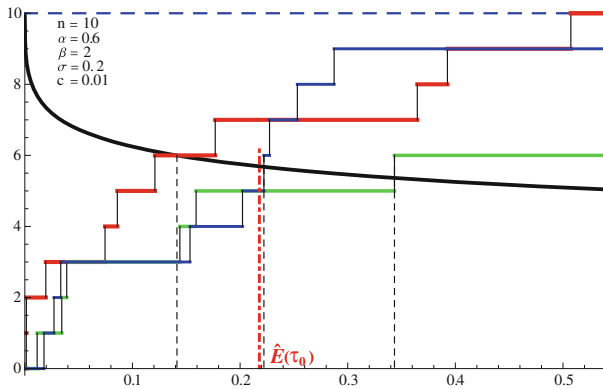
$$V_0(n, F) = E(\sigma^2/K(\tau_0) + c\tau_0),$$

and

$$V_1(n, F) = E(\sigma^2/K(\tau_1) + c\tau_1),$$



**Fig. 1** Plots of boundaries for some values of the model parameters



**Fig. 2** Three sample paths of the process stopped at the boundary  $b(t; n, \alpha, \beta, \sigma, c)$  for  $n = 10, \alpha = 0.6, \beta = 2, \sigma = 0.2, c = 0.01$ . The value  $\hat{E}(\tau_0)$  denotes the estimate of the expectation value of the stopping time  $\tau_0$

where the expected values are calculated under the assumption that  $T_1, \dots, T_n$  are distributed according to the distribution function  $F$ .

In the paper of [Starr et al. \(1976\)](#) it has been shown that  $V_1(n, F) \sim V_0(n, F)$  as  $n \rightarrow \infty$ , provided that the following additional condition is satisfied: for some  $m \geq 1$ ,

$$\int_0^\infty [1 - F(t)]^m dt < \infty. \tag{17}$$

Of course, if (17) fails for all  $m$ , then  $V_0(n, F) = \infty$  for all  $n$ .

We will compare  $V_1(n, F)$  with  $V_0(n, F)$  for some values of  $n$  and Weibull  $\mathcal{W}(\alpha, \beta^{1/\alpha})$  distribution functions  $F$  with some  $0 < \alpha < 1$  and  $\beta > 0$ .

The value  $V_0(n, F)$  is the risk of the optimal sequential procedure  $(\tau_0, d(\tau_0))$ , and can be obtained applying Corollary 3. By formula (12), and the definition of  $b_i, i = 0, \dots, n$ , given by (5), in the case considered we have  $b_n = 0, b_0 = \infty$ , and

$$b_i = \left[ \frac{ci(i+1)}{\alpha\beta\sigma^2(n-i)} \right]^{1/(\alpha-1)}, \quad i = 1, \dots, n-1.$$

Therefore  $k_0 = n - 1, l_0 = 1$ , and

$$V_0(n, F) = \sigma^2 \sum_{i=1}^{n-1} \binom{n}{i} \left\{ \frac{1}{i} F^i(b_i)[1 - F(b_i)]^{n-i} + \int_{b_i}^{b_{i-1}} F^{i-1}(t)[1 - F(t)]^{n-i} f(t) dt \right\} + \frac{\sigma^2}{n} F^n(b_{n-1}) + cE(\tau_0),$$

with  $F(t) = [1 - \exp(-\beta t^\alpha)]\mathbf{1}_{(0, \infty)}(t)$ , and  $E(\tau_0)$  given by (8). Additionally, applying Corollary 1, we obtain that in the case considered

$$E(\tau_0) = \sum_{i=1}^{n-1} b_i \binom{n}{i} F^i(b_i)[1 - F(b_i)]^{n-i} + n \int_0^{b_{n-1}} tf(t)F^{n-1}(t)dt + \sum_{i=1}^{n-1} \frac{n!}{(i-1)!(n-i)!} \int_{b_i}^{b_{i-1}} tf(t)F^{i-1}(t)[1 - F(t)]^{n-i} dt,$$

and applying Corollary 2 we have

$$E(K(\tau_0)) = nF^n(b_{n-1}) + \sum_{i=1}^{n-1} i \binom{n}{i} \left\{ F^i(b_i)[1 - F(b_i)]^{n-i} + i \int_{b_i}^{b_{i-1}} f(t)F^{i-1}(t)[1 - F(t)]^{n-i} dt \right\}.$$

To obtain  $V_1(n, F)$  we have to derive the distributions of the stopping time  $\tau_1$  and the number of observations  $K(\tau_1)$  at the moment of stopping.

Denote

$$\tilde{b}_i = \frac{\sigma^2}{ci}, \quad i = 1, \dots, n, \tag{18}$$

and  $\tilde{b}_0 = \infty$ .

**Theorem 3** *The distribution function  $F_{\tau_1}(t)$  of the adaptive stopping time  $\tau_1$  given by (16) is of the form*

$$F_{\tau_1}(t) = \begin{cases} 0, & t < \tilde{b}_n, \\ \sum_{j=i}^n \binom{n}{j} F^j(t)[1 - F(t)]^{n-j} = \frac{n!}{(i-1)!(n-i)!} \int_0^{F(t)} x^{i-1}(1-x)^{n-i} dx, & t \in [\tilde{b}_i, \tilde{b}_{i-1}), \quad i = 1, \dots, n. \end{cases}$$

**Corollary 4** *The expected value of the adaptive stopping time  $\tau_1$  is given by*

$$\begin{aligned}
 E(\tau_1) &= \sum_{i=1}^n \int_{\tilde{b}_i}^{\tilde{b}_{i-1}} \sum_{j=0}^{i-1} \binom{n}{j} F^j(t) [1 - F(t)]^{n-j} dt \\
 &= \sum_{i=1}^n \int_{\tilde{b}_i}^{\tilde{b}_{i-1}} \left[ 1 - \frac{n!}{(i-1)!(n-i)!} \int_0^{F(t)} x^{i-1} (1-x)^{n-i} dx \right] dt. \quad (19)
 \end{aligned}$$

Using the decomposition of the mixed type distribution of the adaptive stopping time  $\tau_1$  on the discrete part and on the continuous part, the expected value of the stopping time  $\tau_1$  can be written in the following form

$$\begin{aligned}
 E(\tau_1) &= \sum_{i=1}^n \tilde{b}_i \binom{n}{i} F^i(\tilde{b}_i) [1 - F(\tilde{b}_i)]^{n-i} \\
 &\quad + \sum_{i=1}^n \frac{n!}{(i-1)!(n-i)!} \int_{\tilde{b}_i}^{\tilde{b}_{i-1}} t f(t) F^{i-1}(t) [1 - F(t)]^{n-i} dt.
 \end{aligned}$$

**Theorem 4** *The distribution of the number of observations at the moment of the adaptive stopping time  $\tau_1$  is of the form*

$$\begin{aligned}
 P(K(\tau_1) = i) &= \begin{cases} 0, & i = 0, \\ \binom{n}{i} \left\{ F^i(\tilde{b}_i) [1 - F(\tilde{b}_i)]^{n-i} + i \int_{\tilde{b}_i}^{\tilde{b}_{i-1}} F^{i-1}(t) [1 - F(t)]^{n-i} f(t) dt \right\}, & i = 1, \dots, n. \end{cases}
 \end{aligned}$$

**Corollary 5** *The expected value of the number of observations at the moment of the adaptive stopping time  $\tau_1$  is of the form*

$$\begin{aligned}
 E(K(\tau_1)) &= \sum_{i=1}^n i \binom{n}{i} \left\{ F^i(\tilde{b}_i) [1 - F(\tilde{b}_i)]^{n-i} \right. \\
 &\quad \left. + i \int_{\tilde{b}_i}^{\tilde{b}_{i-1}} F^{i-1}(t) [1 - F(t)]^{n-i} f(t) dt \right\}. \quad (20)
 \end{aligned}$$



**Corollary 6** *The risk  $V_1(n, F)$  of the procedure  $(\tau_1, d(\tau_1))$  is of the form*

$$V_1(n, F) = \sigma^2 \sum_{i=1}^n \binom{n}{i} \left\{ \frac{1}{i} F^i(\tilde{b}_i) [1 - F(\tilde{b}_i)]^{n-i} + \int_{\tilde{b}_i}^{\tilde{b}_{i-1}} F^{i-1}(t) [1 - F(t)]^{n-i} f(t) dt \right\} + cE(\tau_1),$$

where  $E(\tau_1)$  is given by (19), and  $\tilde{b}_i, i = 0, \dots, n$ , are given by (18).

Table 1 contains the comparisons of the values  $V_1$  versus the values  $V_0$  with respect to the number  $n$  of devices, and the values of the relative errors of  $V_1$  with respect to  $V_0$  for some values of  $\alpha, \beta, \sigma$  and  $c = 0.1$ . The values of  $E(\tau_0), E(\tau_1), E(K(\tau_0))$  and  $E(K(\tau_1))$  have been calculated using formulas (8), (19), (10) and (20), respectively. We observe that for the same choice of the model parameters the relative errors  $\text{re}(V_1) = (V_1 - V_0)/V_0$  become significantly smaller as the number  $n$  of devices increases. As mentioned above, it was shown by Starr et al. (1976) that, under condition (17),  $V_1 \sim V_0$  as  $n \rightarrow \infty$ .

**Open Access** This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

## References

- Anscombe F (1952) Large sample theory of sequential estimation. Proc Camb Phil Soc 48:600–607
- Bhattacharya PK, Mallik A (1973) Asymptotic normality of the stopping times of some sequential procedures. Ann Stat 1(6):1203–1211
- Chow YS, Robbins H (1965) On the asymptotic theory of fixed-width sequential confidence intervals for the mean. Ann Math Stat 36:457–462
- Ghosh M, Mukhopadhyay N (1975) Asymptotic normality of stopping times in sequential analysis. Unpublished manuscript, Indian Statistical Institute, Calcutta, India
- Govindarajulu Z (2007) Nonparametric inference. World Scientific, Singapore
- Gut A (2012) Anscombe's theorem 60 years later. Seq Anal Design Methods Appl 31(3):368–396
- Jokiel-Rokita A (2006) The Bayes sequential estimation of a normal mean from delayed observations. Appl Math 33(3–4):275–282
- Jokiel-Rokita A, Magiera R (1999) Estimation with delayed observations for the multinomial distribution. Statistics 32:353–367
- Jokiel-Rokita A, Magiera R (2010) Estimation procedures with delayed observations. J Stat Plan Inference 140(4):992–1002
- Magiera R (1982) Estimation with delayed observations. Appl Math 17(2):249–258
- Mukhopadhyay N, Zacks S (2007) Bounded risk estimation of linear combinations of the location and scale parameters in exponential distributions under two-stage sampling. J Stat Plan Inference 137:3672–3686
- Mukhopadhyay N, Chattopadhyay B (2012) A tribute to Frank Anscombe and random central limit theorem from 1952. Seq Anal Design Methods Appl 31(3):265–277
- Pham-Gia T (1998) Distribution of stopping time in Bayesian sequential sampling. Aust N Z J Stat 40(2):221–227
- Robbins H (1959) Sequential estimation of the mean of a normal population. In: Grenander U (ed) Probability and statistics (Harald Cramer Volume). Almqvist & Wiksell, Uppsala, Sweden, pp 235–245
- Siegmund D (1968) On the asymptotic normality of one-sided stopping times. Ann Math Stat 39(5):1493–1497

- Starr N (1966a) On the asymptotic efficiency of a sequential procedure for estimating the mean. *Ann Math Stat* 37:1173–1185
- Starr N (1966b) The performance of sequential procedure for the fixed-width interval estimation of the mean. *Ann Math Stat* 37(1):36–50
- Starr N, Woodroffe M (1972) Further remarks on sequential estimation: the exponential case. *Ann Math Stat* 43(4):1147–1154
- Starr N, Wardrop R, Woodroffe M (1976) Estimating a mean from delayed observations. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* 35:103–113
- Woodroffe M (1977) Second order approximations for sequential point and interval estimations. *Ann Stat* 5(5):984–995
- Zacks S (2005) Some recent results on the distributions of stopping times of compound poisson processes with linear boundaries. *J Stat Plan Inference* 130:95–109
- Zacks S (2009) The exact distributions of the stopping times and their functionals in two-stage and sequential fixed-width confidence intervals of the exponential parameter. *Seq Anal* 28(1):69–81
- Zacks S, Mukhopadhyay N (2006) Exact risks of sequential point estimators of the exponential parameter. *Seq Anal* 25(2):203–226
- Zacks S, Mukhopadhyay N (2007) Distributions of sequential and two-stage stopping times for fixed-width confidence intervals in Bernoulli trials: applications in reliability. *Seq Anal* 26(4):425–441