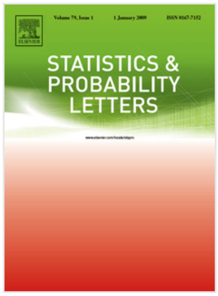
Accepted Manuscript



Bayesian sequential tests of the initial size of a linear pure death process

I.B.J. Goudie

PII: S0167-7152(15)00067-X
DOI: http://dx.doi.org/10.1016/j.spl.2015.02.017
Reference: STAPRO 7233

To appear in: *Statistics and Probability Letters*

Received date: 31 October 2014
Revised date: 16 February 2015
Accepted date: 17 February 2015

Please cite this article as: Goudie, I.B.J., Bayesian sequential tests of the initial size of a linear pure death process. *Statistics and Probability Letters* (2015), http://dx.doi.org/10.1016/j.spl.2015.02.017

Bayesian sequential tests of the initial size of a linear pure death process

I. B. J. Goudie[1]

School of Mathematics and Statistics,
University of St Andrews, St Andrews, Scotland KY16 9SS

Abstract

We provide a recursive algorithm for determining the sampling plans of invariant Bayesian sequential tests of the initial size of a linear pure death process of unknown rate. These tests compare favourably with the corresponding truncated sequential probability ratio tests.

Keywords: censored sampling; exponential order statistics; invariance; Jelinski-Moranda model; truncated sequential probability ratio test.
2010 MSC: 62L10, 62N05, 68M15

1. Introduction

The death-times $T_1, \ldots, T_j$ from a linear pure death process (LPDP) with unknown rate parameter $\lambda > 0$ are observed sequentially. We seek Bayesian sequential tests of the unknown initial size n of the process. The observations obtained can equivalently be regarded as a type II censored sample from an exponential distribution with unknown mean $\lambda^{-1} < \infty$.

Sequential testing of the size n dates back to Hoel (1968) who assumed that the observations were order statistics from a known distribution. The case of exponential observations, with unknown mean, was considered by Goudie (1985), who used a truncated sequential probability ratio test (TSPRT). The LPDP serves as a basic model in software reliability, where it is known as the Jelinski-Moranda model. Sequential Bayesian approaches to estimating the number of faults and to choosing the stopping time were

Email address: ig@st-andrews.ac.uk (I. B. J. Goudie)
[1]Tel: +44 (0)1334 463705 Fax: +44 (0)1334 463748

given by Zacks (2009). Washburn (2006) considered estimation of reliability for a generalisation of this model. The cognate problem of determining when all faults have been detected has also received attention (Fang et al., 2003). In ecology, the LPDP arises as a removal model in continuous time or as describing the times to first capture in a continuous-time behavioural capture-recapture model (cf. Hwang and Chao, 2002).

2. The invariant statistics

As the times $T_1 < \ldots < T_j$ can be viewed as the smallest j order statistics of an exponential sample, their joint probability density function (p.d.f.) is

$$j! \binom{n}{j} \lambda^j \exp\left[-\lambda\left\{(n-j)t_j + \sum_1^j t_i\right\}\right] \qquad 0 < t_1 < \ldots < t_j.$$

If $X_1 \equiv 0$ and $X_i = i - [(T_1 + \ldots + T_i)/T_i]$ for $i = 2, \ldots, n$, then (X_j, T_j) is minimal sufficient for (n, λ). It follows from Goudie and Goldie (1981) that X_j is a maximal invariant under scale transformations and has p.d.f.

$$f(x_j|\, n) = j! \binom{n}{j} (n - x_j)^{-j}\, h(x_j) I_{(0,j-1)}(x_j),$$

where $I_{(c,d)}$ is the indicator function for the interval (c, d) and

$$h(x) = \frac{1}{(j-2)!} \sum_{s=0}^{j-2} (-1)^s \binom{j-1}{s} (x-s)^{j-2} I_{(s,j-1)}(x),$$

which is the p.d.f. of an Irwin-Hall distribution, namely the distribution of the sum of $j-1$ independent uniform random variables on the interval $(0, 1)$.

In this paper we derive tests of the hypothesis $H_0 : n = N_0$ against the alternative $H_1 : n = N_1$, where $N_1 > N_0$. We restrict attention to procedures based on the sequence $X_2, \ldots, X_j$ of invariant statistics. The tests derived are thus suitable for a Bayesian with a knowledge of λ too weak for deriving useful information from the component T_j of the sufficient statistic (X_j, T_j).

Before experimentation, the prior probability that hypothesis H_v is true is π_v $(v = 0, 1)$, where $\pi_0 + \pi_1 = 1$. Having observed $X_2, \ldots, X_j$, the likelihood ratio in favour of H_1 reduces to $L_j = f(x_j|\, N_1)/f(x_j|\, N_0)$, since $f(x_2, \ldots, x_j|n)$ is the product of $f(x_j|\, n)$ and $f(x_2, \ldots, x_{j-1}|x_j)$, where, by the results of section 7 below, the latter factor does not depend on n. Note that we use f for various sampling distributions: the same functional form is only implied when the arguments are the same. Thus, after observing X_j, the posterior probability p_j^0 that H_0 is true is given by Bayes' Theorem as $p_j^0 = \pi_0/(\pi_0 + \pi_1 L_j)$.

3. Cost structure

Errors of the first and second kind are assumed to result in non-negative stopping losses of K_0 and K_1 respectively. For $j = 2, \ldots, N_0$, the expected cost of stopping after observing X_j is thus $\min(K_0\, p_j^0,\ K_1\, p_j^1)$. In the $(j,\, x_j)$ plane, we are thus indifferent between the two hypotheses at points $(j,\, m_j)$ where m_j is the value of X_j satisfying $K_0\, p_j^0 = K_1\, p_j^1$. As p_j^0 is an increasing function of x_j, stopping after j observations implies choosing H_0 if $x_j > m_j$ or H_1 if $x_j < m_j$. For $j = 2, \ldots, N_0$, the cost of the j^{th} observation x_j is taken to be $c_j > 0$. We assume that the sequence $\{c_j\}$ is non-decreasing, which will usually be realistic as the expected waiting time between observations increases. We ignore the cost of observing X_1 as it does not affect the sampling plans generated: if it were sufficiently high, however, it would be cheaper to take an immediate decision than to take any observations.

Under the TSPRT, a sample path can lie in the continuation region (CR) after N_0 observations, and it is then advantageous to wait to see if another observation occurs. We permit comparable action here, and charge c_{N_0+1} for waiting from time T_{N_0} to time $(N_0 - x_{N_0})T_{N_0}/(N_0 - A_{N_0+1})$ for a suitable constant $A_{N_0+1} > x_{N_0}$. If we wait for this time, and an $(N_0+1)^{\text{th}}$ observation exists but is not seen, elementary algebra shows that $X_{N_0+1} > A_{N_0+1}$. This cost structure, however, precludes comparison within the decision theoretic framework of sampling plans with different values of A_{N_0+1}, implying that A_{N_0+1} has to be chosen arbitrarily or by some other method than balancing costs. Our optimal sampling plans are thus optimal within the class of policies based on the invariant statistics and with a given termination point. We will index the sampling plans not by the value of A_{N_0+1} but by the corresponding posterior probability π^* on H_0 at the time that sampling terminates if no $(N_0+1)^{\text{th}}$ observation is seen.

4. The shape of the optimal continuation region

Once A_{N_0+1} is fixed, a standard result (cf. DeGroot, 2004, p. 307) shows that, from any point $(j,\, x_j)$ for $j \le N_0$, the cost of awaiting one further observation and then acting optimally is a concave function of p_j^0. As the expected stopping cost is $\min(K_0\, p_j^0,\ K_1\, p_j^1)$ and p_j^0 is an increasing function of x_j, we see that sampling continues after j observations ($2 \le j \le N_0$) if and only if $R_j < x_j < A_j$, where $\{A_j\}$ and $\{R_j\}$ are sequences of constants such that $0 \le R_j \le m_j \le A_j \le j-1$. We also set $A_1 = R_1 = 0$. For $j \ge 2$, when sampling terminates, we accept H_0 if $x_j \ge A_j$ or reject H_0 if $x_j \le R_j$.

It is easy to verify that $X_{j+1} > X_j$ $(j \ge 1)$, implying that sample paths have positive gradient. So we assume that $R_{j+1} > R_j$ whenever $R_j > 0$ $(j = 2, \ldots, N_0 - 1)$. If this were not so, there would exist integers $i_1 < i_2 < i_3$ for which it is possible to stop and accept H_1 after i_1 or i_3 observations, but impossible to do so after i_2 observations. As the sequence $\{c_j\}$ of sampling costs is non-decreasing, this would not appear to be reasonable.

It is easy to show that the upper boundary has a positive gradient. If it were otherwise, there would be points $(j,\, x_j)$ from which, with probability one, exactly one more observation would be awaited before sampling ceased. This would be more expensive than stopping immediately, since (cf. Table 1 below) the predicted value of $K_1\, p_{j+1}^1$ given $X_j = x_j$ is just $K_1\, p_j^1$. A similar argument shows that it would be inappropriate to set $A_{N_0+1} \le m_{N_0}$.

As the sample paths have positive gradient, from any point $(j,\, x_j)$, the value of x_j relative to the points $\{R_k\}$ determines where the lower boundary may be crossed. In fact, setting $j_r = \max\{j : R_j = 0\}$, we note that, if $x_j \in (R_k, R_{k+1})$, where $k \ge j_r$, the lower boundary cannot be crossed before the $(k+1)^{\text{th}}$ observation. For each j, the interval (R_j, A_j) of values of x_j in the CR can be sub-divided by any elements of the sequence $\{R_k : k > j\}$ lying in that interval. To index the highest and lowest sub-divisions in the CR, for $j \ge 2$, we set $k(j) = \max\{k : R_k < A_j\}$ and $b(j) = \max(j, j_r)$. The ordinates of the grid of points $\{(j, r_{j,k})\}$ are then given by

$$r_{j,k} = \begin{cases} R_k & b(j) \le k \le k(j); \\ A_j & k = k(j) + 1. \end{cases}$$

A simple example is given in Figure 1. We also set $M = \max\{j : 2 \le j \le N_0, A_j \ne R_j\}$. Thus $M + 1$ is the maximum number of observations that we see or await. If, for some sample paths, the policy requires us to await an $(N_0 + 1)^{\text{th}}$ observation, we have $M = N_0$.

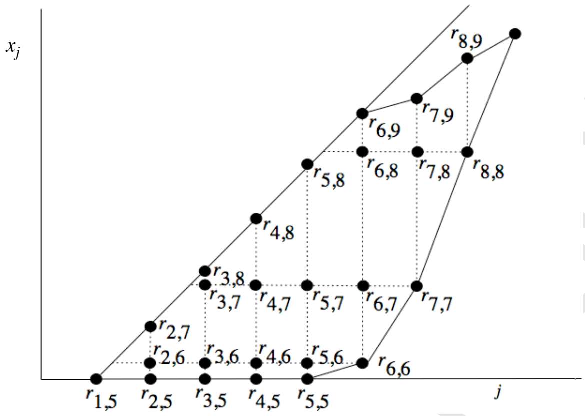


Fig. 1. *An example showing the grid* $\{r_{j,k}\}$ *defined on the* (j, x_j) *plane.*

5. The minimal continuation loss

Central to the Bayesian analysis is the loss that will be incurred at any stage by taking another observation and then pursuing the optimal policy. We denote this loss, when starting from $(j,\, x_j)$, after j observations, by $Z_j(x_j)$. Setting $G_j^v(x_j, s) = \{(N_v - x_j)/(j+s-x_j)\}^j p_j^v$, we prove in sections 8 and 9 that, for $x_j \in [r_{j,k},\ r_{j,k+1})$ with $b(j) \le k \le k(j)$ and $2 \le j \le M$,

$$Z_j(x_j) = C_{j,k} + \tilde{K}_0\, p_j^0 + \sum_{v=0}^{1} \sum_{s=0}^{k-j} \binom{N_v - j}{s} D_{j+s,k}^v\, G_j^v(x_j, s), \qquad (1)$$

where $C_{j,k} = c_{j+1} + \ldots + c_{k+1}$ and $\tilde{K}_0 \ = K_0\left(1 - \delta_k^{N_0}\right)$, where, as throughout this paper, the lower case delta is a Kronecker delta.

The coefficients $D_{j,k}^v$ are determined recursively. The initial values are given in section 6, while for $1 \le j < M$ we use the equations

$$D_{j,k}^v - \delta_1^v K_1 U_j^1(A_{j+1}) - \sum_{q=k+1}^{k(j+1)} W_{j+1,q}^v = \begin{cases} -\,\delta_0^v\, K_0\, U_j^0(R_{j+1}) & k = j > 1, \\ -\,\delta_0^v\, \delta_{j_r}^1\, K_0\, U_1^0(R_2) & k = j = 1, \\ (\delta_1^j - 1) S_{j+1,k}^v\,(r_{j+1,k+1}) & k > j, \end{cases} \qquad (2)$$

where, for $j \ge 2$,

$$S_{j,k}^v(x) = \left(C_{j,k} + \delta_0^v\, \tilde{K}_0\right) U_{j-1}^v(x) + \sum_{s=0}^{k-j} \binom{N_v - j + 1}{s+1} D_{j+s,k}^v\, V_{j-1}^s(x), \qquad (3)$$

$$W^v_{j,k} = S^v_{j,k}\left(r_{j,k}\right) - S^v_{j,k}\left(r_{j,k+1}\right), \tag{4}$$

and, for $j \geq 1$,

$$U^v_j(x) = \{(j-x)/(N_v - x)\}^j, \qquad V^s_j(x) = \{(j-x)/(j+1+s-x)\}^j.$$

In equation (2) we use the convention that the sum is omitted if the lower limit $k+1$ exceeds the upper limit $k(j+1)$. It also follows from (1) that the total expected cost of pursuing the optimal policy from the point $(1,0)$ in the $(j,\, x_j)$ plane is given by

$$Z_1(0) = c_2 + \delta^1_{j_r} K_0 p^0_1 + \sum_{v=0}^{1} D^v_{1,k}\, G^v_1(0,0), \tag{5}$$

where the coefficients $D^v_{1,k}$ are given by (2) with $k = b(2) - 1$.

6. Implementation of the algorithm

The steps given below specify how to determine the initial values $D^v_{M,M}$ $(v = 0,1)$ for the recursion. They also give the order of evaluation for the boundary points of the CR and the coefficients used in the algorithm. When, in Step 4, solution of an equation for a lower boundary point R_j is required, one should first check whether, as $x_j \downarrow 0$, the expected cost of stopping exceeds that of continuing to sample. If it does then $R_j = 0$. If it does not, the equation for the boundary point has a solution in $(0, m_j)$. This can be found by the Newton-Raphson method, using numerical differentiation to obtain the values of the relevant derivative. Similar comments apply to Step 6, where solution of an equation for an upper boundary point is indicated.

Step 1. Set $j_r = 1$ and $k = N_0$, and then test as follows whether it is ever worthwhile to await an $(N_0+1)^{\text{th}}$ observation. The expected cost at (N_0, m_{N_0}) of an immediate decision is given by $K_0\, p^0_{N_0} = K_1\, p^1_{N_0}$, with the probabilities evaluated at $x_{N_0} = m_{N_0}$. On the other hand, the expected cost of awaiting an $(N_0+1)^{\text{th}}$ observation for a time not exceeding that at which x_{N_0+1} would equal A_{N_0+1} is given by $Z_{N_0}(m_{N_0})$ with

$$D^0_{N_0,N_0} = 0, \qquad D^1_{N_0,N_0} = K_1\, U^1_{N_0}(A_{N_0+1}). \tag{6}$$

If $Z_{N_0}(m_{N_0}) < K_0\, p^0_{N_0}$, set $M = N_0 = j = b(j) = k(j) = k$ and go to Step 4.

Step 2. Here we determine the maximum number $M+1$ of observations that will be seen or awaited, knowing in this case that $M < N_0$. To find M, evaluate $Z_j(m_j)$ with $k = j$ for successively smaller values of j, starting with $j = N_0 - 1$. For each j, the required coefficients are

$$D^v_{j,j} = (-1)^{v+1} K_v \, U^v_j(m_{j+1}) \qquad\qquad v = 0, 1. \tag{7}$$

Thus let M be the largest j for which $Z_j(m_j)$ is less than the value of $K_0 \, p^0_j$ at $x_j = m_j$. Set $A_{M+1} = R_{M+1} = m_{M+1}$, and with $j = k = M = b(j) = k(j)$, go to Step 4.

Step 3. Calculate the coefficients $D^v_{j,k}$ for $v = 0, 1$. If $k > b(j)$, go to Step 5. If $R_{j+1} = 0$, set $R_j = 0$ and go to Step 5.

Step 4. Find the value R_j of x_j that satisfies $Z_j(x_j) = K_0 \, p^0_j$. If $j_r = 1$ and $R_j = 0$, set $j_r = j$. If $j = M$, go to Step 6.

Step 5. By comparing the expected losses from stopping and from continuing to sample, determine whether or not the point (j, R_{k+1}) lies in the CR. If it is in the CR, go to Step 7.

Step 6. Find the value A_j of x_j that satisfies $Z_j(x_j) = K_1 \, p^1_j$.

Step 7. Calculate $S^v_{j,k}(r_{j,k}), S^v_{j,k}(r_{j,k+1})$ and $W^v_{j,k}$ for $v = 0, 1$. If the point (j, R_{k+1}) lies in the CR, increase k by one and return to Step 3. Otherwise, set $k(j) = k$, decrease j by one and put $k = b(j)$. If the new value of j exceeds one, return to Step 3. Otherwise, using (5), evaluate $Z_1(0)$.

7. Distributional results and integral formulae

The joint p.d.f. of $\alpha_2, \ldots, \alpha_j$, where $\alpha_i = i - X_i$ $(i = 2, \ldots, j)$ is known (Goudie, 1985). It follows that, for $v = 0, 1$, the joint p.d.f. $f(x_2, \ldots, x_j; N_v)$ of $X_2, \ldots, X_j$ is

$$(j-1)!\, j! \binom{N_v}{j} (N_v - x_j)^{-j} \prod_{i=1}^{j-1} \left[\left\{ (i - x_{i+1})^{i-1} / (i - x_i)^i \right\} I_{(x_i, i)}(x_{i+1}) \right].$$

Thus, for $j \geq 2$, the conditional p.d.f. of X_{j+1} given $X_2, \ldots, X_j$ depends only on X_j and equals

$$f(x_{j+1} | \, x_j; N_v) = \frac{j\,(N_v - j)\, G^v_j(x_j, 0)\, (j - x_{j+1})^{j-1}}{p^v_j \, (N_v - x_{j+1})^{j+1}} I_{(x_j, j)}(x_{j+1}). \tag{8}$$

Table 1. *Integrals over subsets of* (x_j, j)

Subset	Function	Integral with respect to $f(x_{j+1}\| x_j; N_v)$
(c,d)	1	$G_j^v(x_j,0)\left[U_j^v(c)-U_j^v(d)\right]/p_j^v$
(c,d)	$G_{j+1}^v(x_{j+1},s)/p_{j+1}^v$	$(N_v-j)\,G_j^v(x_j,0)\left[V_j^s(c)-V_j^s(d)\right]/\left\{(s+1)p_j^v\right\}$

Subset	Function	Integral with respect to $f_p(x_{j+1}\| x_j)$
(c,d)	p_{j+1}^v	$G_j^v(x_j,0)\left[U_j^v(c)-U_j^v(d)\right]$
(c,d)	1	$\sum_{v=0}^1 G_j^v(x_j,0)\left[U_j^v(c)-U_j^v(d)\right]$
(c,d)	$G_{j+1}^v(x_{j+1},s)$	$(N_v-j)G_j^v(x_j,0)\left[V_j^s(c)-V_j^s(d)\right]/(s+1)$
(x_j,d)	p_{j+1}^v	$p_j^v-G_j^v(x_j,0)U_j^v(d)$
(x_j,d)	1	$1-\sum_{v=0}^1 G_j^v(x_j,0)U_j^v(d)$
(x_j,d)	$G_{j+1}^v(x_{j+1},s)$	$(N_v-j)\left[G_j^v(x_j,s+1)-G_j^v(x_j,0)V_j^s(d)\right]/(s+1)$

It may be verified that this equation is also valid for $j=1$, for which it gives the unconditional p.d.f. of X_2. The p.d.f. of the predictive distribution of X_{j+1} given X_j can thus be obtained using

$$f_p(x_{j+1}|\,x_j) = p_j^0 f(x_{j+1}|\,x_j; N_0) + p_j^1 f(x_{j+1}|\,x_j; N_1).$$

The integral formulae in Table 1 can now be verified. In deriving the integrals over (c,d) with respect to the predictive p.d.f. note that, by Bayes' Theorem, $p_{j+1}^v f_p(x_{j+1}|\,x_j) = p_j^v f(x_{j+1}|\,x_j; N_v)$. In the final three lines of the table, we set $c = x_j$, and note that $G_j^v(x_j,0)U_j^v(x_j) = p_j^v$ and $G_j^v(x_j,0)V_j^s(x_j) = G_j^v(x_j,s+1)$.

8. Initial values for the backward induction

Suppose an $(N_0+1)^{\text{th}}$ observation is awaited. If it is observed, the expected terminal loss in choosing H_1 is zero. If, on the other hand, no such observation has been seen by the time at which X_{N_0+1} would equal A_{N_0+1}, the expected terminal loss in choosing H_0 at this time is the product of K_1 and the posterior probability on H_1. Weighting these outcomes by their predictive probabilities given X_{N_0}, and also including the sampling cost, it follows by Bayes' Theorem and Table 1 that

$$\begin{aligned} Z_{N_0}(x_{N_0}) &= c_{N_0+1} + K_1 p^1_{N_0} P\left[X_{N_0+1} > A_{N_0+1} \mid X_{N_0} = x_{N_0}; N_1\right] \\ &= c_{N_0+1} + K_1 G^1_{N_0}(x_{N_0}, 0)\, U^1_{N_0}(A_{N_0+1}) \\ &= c_{N_0+1} + \sum\nolimits_{v=0}^{1} D^v_{N_0,N_0} G^v_{N_0}(x_{N_0}, 0), \end{aligned} \tag{9}$$

where the coefficients $D^v_{N_0,N_0}$ for $v = 0, 1$ are given by (6). As in Step 1 of section 6, if $Z_{N_0}(m_{N_0}) < K_0\, p^0_{N_0}$, evaluated at m_{N_0}, we have $M = N_0$. In this case, equation (9) holds for x_{N_0} in some interval $[R_{N_0}, A_{N_0})$, where the end-points are determined in Steps 4 and 6 of the algorithm. We then have $b(N_0) = k(N_0) = N_0$, and so equation (9) agrees with the general expression (1).

For $j < N_0$, the expected cost of stopping after $j+1$ observations is $K_1 p^1_{j+1} I_{(m_{j+1},j)}(x_{j+1}) + K_0 p^0_{j+1} I_{(x_j, m_{j+1})}(x_{j+1})$. If, after seeing j observations, we take the predictive expectation of this stopping cost, and add the cost of the next observation, we obtain, using Table 1 and equation (7), that

$$Z_j(x_j) = c_{j+1} + K_0 p^0_j + \sum\nolimits_{v=0}^{1} D^v_{j,j} G^v_j(x_j, 0). \tag{10}$$

So, when $M < N_0$, we can, as in Step 2, evaluate $Z_j(m_j)$ for successively smaller values of j until we find the value M of j for which $Z_j(m_j) < K_0\, p^0_j$. Equation (10) then holds for $x_M \in [R_M, A_M)$, where this interval is again found as in Steps 4 and 6 of the algorithm. We then have $b(M) = k(M) = M$, and so equation (10) also agrees with expression (1).

9. The backward inductive proof

Assume that the algorithm holds for $b(j) \le k \le k(j)$, where j is such that $M \ge j \ge i+1 > 1$. Using equation (1) and Table 1, it then follows that the integral of $Z_{i+1}(x_{i+1})$, with respect to the predictive p.d.f. $f_p(x_{i+1}|\, x_i)$, over the interval $(c, d) \subseteq (r_{i+1,k},\ r_{i+1,k+1})$ is

$$\left[-\sum_{v=0}^{1} G^v_i(x_i, 0) \left\{ \left(C_{i+1,k} + \delta^v_0 \tilde{K}_0 \right) U^v_i(x) + \sum_{s=0}^{k-i-1} \binom{N_v - i}{s+1} D^v_{i+s+1,k} V^s_i(x) \right\} \right]_{x=c}^{d}.$$

When $(c, d) = (r_{i+1,k},\ r_{i+1,k+1})$, by equations (3) and (4), this reduces to

$$\sum\nolimits_{v=0}^{1} W^v_{i+1,k} G^v_i(x_i, 0). \tag{11}$$

Similarly, using (3) and Table 1 and writing $s' = s + 1$, the corresponding integral of $Z_{i+1}(x_{i+1})$ over the interval $(x_i,\ r_{i+1,k+1})$, where $x_i > r_{i+1,k}$, is

$$C_{i+1,k} + \tilde{K}_0 p_i^0 - \sum_{v=0}^{1} \left[S_{i+1,k}^v (r_{i+1,k+1}) G_i^v(x_i, 0) - \sum_{s'=1}^{k-i} \binom{N_v - i}{s'} D_{i+s',k}^v G_i^v(x_i, s') \right]. \quad (12)$$

For $x_i \in [r_{i,k}, r_{i,k+1})$, where $k > i$, the minimal continuation loss $Z_i(x_i)$ is

$$c_{i+1} + \int_{x_i}^{A_{i+1}} Z_{i+1}(x_{i+1}) f_p(x_{i+1}|\, x_i) dx_{i+1} + \int_{A_{i+1}}^{i} K_1 p_{i+1}^1 f_p(x_{i+1}|\, x_i) dx_{i+1}. \quad (13)$$

The part of the first integral over $(x_i, r_{i+1,k+1})$ is given by (12). The sum of the remaining part, if any, and the other two terms is given by (11) and Table 1 as

$$c_{i+1} + K_1 U_i^1(A_{i+1}) G_i^1(x_i, 0) + \sum_{q=k+1}^{k(i+1)} \sum_{v=0}^{1} W_{i+1,q}^v G_i^v(x_i, 0). \quad (14)$$

The sum of (12) and (14) thus gives an expression for $Z_i(x_i)$. It then follows, using equation (2), that this expression agrees with that provided by (1).

When $k = i$, the lower limit of the first integral in (13) becomes $r_{i+1,k+1} = R_{i+1}$, implying that the contribution from (12) no longer arises. There is now, however, an additional term, namely the integral of $K_0 p_{i+1}^0$ over the interval (x_i, R_{i+1}) with respect to the predictive p.d.f. of x_{i+1} given x_i. By Table 1 this additional term equals $K_0 \left[p_i^0 - G_i^0(x_i, 0) U_i^0(R_{i+1}) \right]$ and adding it to (14) gives an expression for $Z_i(x_i)$, for $x_i \in [R_i,\ r_{i,i+1})$, which, using equation (2), also agrees with that provided by (1).

10. Numerical Examples

In Table 2, for $N_0 = 15$ and $N_1 = 30$, we compare the above tests with the TSPRT, in which sampling continues if $A < f(x_j|\, N_1)/f(x_j|\, N_0) < R$ for $j \leq N_0$. If no 16^{th} observation was seen, we terminated sampling when X_{16} equalled the upper boundary point for X_{15}. The constants A and R were taken to be 0.0809 and 6.84 respectively in order to achieve type I and type II error probabilities of 0.05.

For the Bayesian tests, we took the sampling costs as $c_j = c/(N_0 - j + 1)$ for $j \leq N_0$, making them proportional to the expected waiting times under H_0. We also set $c_{N_0+1} = c$ and $K_0 = K_1 = 1$, and, for each value of π^*, chose c and π_0 to match the error probabilities of the TSPRT.

Table 2. *Properties of sequential tests of* $N_0 = 15$ *against* $N_1 = 30$

	TSPRT		Bayesian tests		
	$A = 0.0809$	π^*	0.925	0.95	0.975
	$R = 6.84$	c	0.0661	0.0746	0.0825
		π_0	0.425	0.453	0.477
ASN when $n = 15$	15.02		14.76	14.64	14.51
ASN when $n = 30$	15.08		15.01	14.98	14.95
ATT when $n = 15$	2.91		2.79	2.76	2.77
ATT when $n = 30$	0.71		0.71	0.71	0.71

For each of the tests in Table 2, the average sample number (ASN) and the average time to termination (ATT) can be evaluated (Goudie, 1985). For the ASN, we assume an observation that is awaited is counted as an observation, whether or not it occurs. For the ATT, if the upper boundary is reached, we assume that sampling terminates immediately without waiting for the observation to occur.

These results confirm that for this problem the TSPRT is not the optimal design for minimising either the ASN or the ATT for fixed error rates. The Bayesian tests show small reductions in the ASN, particularly under H_0. Further slight reductions in the ASN can be made by further increasing π^*, though at the expense of some increase in the ATT. For small λ, however, the most useful aspect here is the reduction of up to around 5% in the ATT under H_0, in addition to the saving in the ATT achieved by the TSPRT compared to the fixed sample size test. The Bayesian sampling plans thus also have merit when judged by frequentist criteria.

References

DeGroot, M. (2004), *Optimal statistical decisions*, Wiley, Hoboken.

Fang, X., Watson, R., Yan, W. and Yip, P. (2003), 'A procedure for complete fault detection with a removal process', *Journal of Statistical Planning and Inference* **117**, 1–14.

Goudie, I. B. J. (1985), 'Sequential testing of the size of an exponential sample', *Sequential Analysis* **4**, 179–198.

Goudie, I. B. J. and Goldie, C. (1981), 'Initial size estimation for the linear pure death process', *Biometrika* **68**, 543–550.

Hoel, D. (1968), 'Sequential testing of sample size', *Technometrics* **10**, 331–341.

Hwang, W.-H. and Chao, A. (2002), 'Continuous-time capture-recapture models with covariates', *Statistica Sinica* **12**, 1115–1131.

Washburn, A. (2006), 'A sequential Bayesian generalization of the Jelinski-Moranda software reliability model', *Naval Research Logistics* **53**, 354–362.

Zacks, S. (2009), *Stage-wise adaptive designs*, Wiley, Hoboken.