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

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.
Bayesian sequential tests of the initial size of a linear pure death process

I. B. J. Goudie

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
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

\[
j! \left( \frac{n}{j} \right) \lambda^j \exp \left[ -\lambda \left( (n-j)t_j + \sum_{i=1}^{j} t_i \right) \right] 0 < t_1 < \ldots < t_j.
\]

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

\[
f(x_j | n) = j! \left( \frac{n}{j} \right) (n-x_j)^{-j} h(x_j) I_{(0,j)}(x_j),
\]

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 \( \lambda \) 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 \( \pi_v \) \((v = 0,1)\), 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 \( f(x_2, \ldots, x_{j-1} | n) \) 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 \((N_0 + 1)\)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 \( \pi^* \) on \( H_0 \) at the time that sampling terminates if no \((N_0 + 1)\)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 \( j \leq N_0 \), the cost of awaiting one further
observation and then acting optimally is a concave function of \( p_0^j \). As the expected stopping cost is \( \min(K_0 p_0^j, K_1 p_1^j) \) and \( p_0^j \) is an increasing function of \( x_j \), we see that sampling continues after \( j \) observations \( (2 \leq j \leq N_0) \) if and only if \( R_j < x_j < A_j \), where \( \{A_j\} \) and \( \{R_j\} \) are sequences of constants such that \( 0 \leq R_j \leq m_j \leq A_j \leq j - 1 \). We also set \( A_1 = R_1 = 0 \). For \( j \geq 2 \), when sampling terminates, we accept \( H_0 \) if \( x_j \geq A_j \) or reject \( H_0 \) if \( x_j \leq R_j \).

It is easy to verify that \( X_{j+1} > X_j (j \geq 1) \), 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 \( K_1 p_{j+1}^j \) given \( X_j = x_j \) is just \( K_1 p_1^j \). A similar argument shows that it would be inappropriate to set \( A_{N_0+1} \leq m_{N_0} \).

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 \( k \geq j_r \), the lower boundary cannot be crossed before the \( (k+1) \)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 \( j \geq 2 \), 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

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

A simple example is given in Figure 1. We also set \( M = \max\{j : 2 \leq j \leq N_0, A_j \neq R_j\} \). 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 \( (N_0 + 1) \)th observation, we have \( M = N_0 \).
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 \(G_j^v(x_j, s) = \{(N_v - x_j)/(j + s - x_j)\}/p_j^v\), we prove in sections 8 and 9 that, for \(x_j \in [r_{j,k}, r_{j,k+1})\) with \(b(j) \leq k \leq k(j)\) and \(2 \leq j \leq M\),

\[
Z_j(x_j) = C_{j,k} + \bar{K}_0 \sum_{s=0}^{k-j} \sum_{v=0}^{k} \left( -\frac{N_v - j}{s} \right) D_{j+s,k}^v G_j^v(x_j, s),
\]

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

The coefficients \(D^v_{j,k}\) are determined recursively. The initial values are given in section 6, while for \(1 \leq j < M\) we use the equations

\[
D^v_{j,k} - \delta_1^v K_0 U_j^1(A_j+1) = \sum_{q=k+1}^{k+j+1} W^v_{j+1,q},
\]

where, for \(j \geq 2\),

\[
S^v_{j,k}(x) = \left( C_{j,k} + \delta_0^v \bar{K}_0 \right) U^1_{j-1}(x) + \sum_{s=0}^{k-1} \left( \frac{N_v - j + 1}{s + 1} \right) D^v_{j+s,k} V^s_{j-1}(x),
\]

Fig. 1. An example showing the grid \(\{r_{j,k}\}\) defined on the \((j, x_j)\) plane.
\[ W_{j,k}^v = S_{j,k}^v (r_{j,k}) - S_{j,k}^v (r_{j,k+1}), \quad (4) \]

and, for \( j \geq 1 \),
\[ U_v^j(x) = \left\{ \frac{(j-x)}{(N_v-x)} \right\}^j, \quad V_s^j(x) = \left\{ \frac{(j-x)}{(j+1+s-x)} \right\}^j. \]

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 \((j,x)\) plane is given by
\[ Z_1(0) = c_2 + \delta^1_p K_0 p_0^1 + \sum_{v=0}^{1} D_{1,k}^v G_1^v(0,0), \quad (5) \]
where the coefficients \( D_{1,k}^v \) 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 \( D_{M,M}^v \) (\( v = 0,1 \)) 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 \( j = 1 \) and \( k = N_0 \), and then test as follows whether it is ever worthwhile to await an \((N_0+1)\)th observation. The expected cost at \((N_0,m_{N_0})\) of an immediate decision is given by \( K_0 p_{N_0} = K_1 p_{N_0}^1 \), with the probabilities evaluated at \( x_{N_0} = m_{N_0} \). On the other hand, the expected cost of awaiting an \((N_0+1)\)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
\[ D_{N_0,N_0}^0 = 0, \quad D_{N_0,N_0}^1 = K_1 U_{N_0}^1(A_{N_0+1}). \quad (6) \]

If \( Z_{N_0}(m_{N_0}) < K_0 p_{N_0}^0 \), 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

$$D_{j,v}^v = (-1)^{v+1} K_v U_j^v(m_{j+1}) \quad v = 0, 1.$$  

Thus let $M$ be the largest $j$ for which $Z_j(m_j)$ is less than the value of $K_0 p_j^0$ 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 $D_{j,k}^v$ 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 $Z_j(x_j) = K_0 p_j^0$. 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 $Z_j(x_j) = K_1 p_j^1$.

Step 7. Calculate $S_{j,k}^v(r_{j,k}), S_{j,k}^v(r_{j,k+1})$ and $W_{j,k}^v$ 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 $Z_i(0)$.

7. Distributional results and integral formulae

The joint p.d.f. of $X_i, \ldots, X_j$, 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

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

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

$$f(x_{j+1} | x_j; N_v) = \frac{j (N_v - j) G_j^v(x_j, 0) (j - x_{j+1})^{j-1}}{p_j^v (N_v - x_{j+1})^{j+1} I_{(x_j,j)}(x_{j+1})} .$$  

(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                         | \(G^v_j(x_j, 0) \left[ U^v_j(c) - U^v_j(d) \right] / p^v_j\) |
| \((c, d)\)   | \(G^v_j(x_j + s, j) / p^v_j + 1\) | \((N_v - j) G^v_j(x_j, 0) \left[ V^v_j(c) - V^v_j(d) \right] / \{ (s + 1) p^v_j \}\) |

--

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

\[
 f_p(x_{j+1} | x_j) = p^0_j f(x_{j+1} | x_j; N_0) + p^1_j f(x_{j+1} | x_j; N_1).
\]

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, \(p^v_{j+1} f_p(x_{j+1} | x_j) = p^v_j f(x_{j+1} | x_j; N_v)\). In the final three lines of the table, we set \(c = x_j\), and note that \(G^v_j(x_j, 0)U^v_j(x_j) = p^v_j\) and \(G^v_j(x_j, 0)V^v_j(x_j) = G^v_j(x_j, s + 1)\).

8. Initial values for the backward induction

Suppose an \((N_0 + 1)\)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
\[
Z_{N_0}(x_{N_0}) = c_{N_0+1} + K_1 p_1^1 P[X_{N_0+1} > A_{N_0+1} | X_{N_0} = x_{N_0}; N_1]
\]
\[
= c_{N_0+1} + K_1 G_{N_0}^1(x_{N_0}, 0) U_{N_0}^1(A_{N_0+1})
\]
\[
= c_{N_0+1} + \sum_{v=0}^1 W_{N_0}^v(x_{N_0}, 0),
\]
where the coefficients \(D_{N_0,N_0}^v\) for \(v = 0, 1\) are given by (6). As in Step 1 of section 6, if \(Z_{N_0}(m_{N_0}) < K_0 p_0^0\), 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 \(K_1 p_{j+1}^1 I_{(m_{j+1}, j]}(x_{j+1}) + K_0 p_0^0 I_{(x_{j}, m_{j+1})}(x_{j+1})\). 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
\[
Z_j(x_j) = c_{j+1} + K_0 p_j^0 + \sum_{v=0}^1 D_{x_j}^v G_{x_j}^v(x_j, 0).
\]
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 \(Z_j(m_j) < K_0 p_j^0\).

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 \(b(j) \leq k \leq k(j)\), where \(j\) is such that \(M \geq j \geq i + 1 > 1\). 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
\[
\left[ -\sum_{v=0}^1 G_v^i(x_i, 0) \left\{ \left( C_{i+1,k}^i + \delta_v^p K_0^i \right) U_v^i(x) + \sum_{s=0}^{k-i-1} \left( \frac{N_u - i}{s+1} \right) D_{i+1,k}^v V_s^i(x) \right\} \right]_{x=c}^{d}.
\]
When \((c, d) = (r_{i+1,k}, r_{i+1,k+1})\), by equations (3) and (4), this reduces to
\[
\sum_{v=0}^1 W_{i+1,k}^v G_v^i(x_i, 0).
\]
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

$$C_{i+1,k} + K_0 p_i^0 - \sum_{v=0}^{1} \left[ S_{i+1,k}^{v}(r_{i+1,k+1}) G_i^v(x_i, 0) - \sum_{s'=1}^{k-1} \binom{N_x - i}{s'} D_{i+1,s}^v G_i^v(x_i, s') \right]. \quad (12)$$

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

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

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

$$c_{i+1} + K_1 U_1^i (A_{i+1}) G_i^1(x_i, 0) + \sum_{q=k+1}^{k+1} \sum_{v=0}^{1} W_1^v G_i^v(x_i, 0). \quad (14)$$

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 $K_0 [p_i^0 - G_i^0(x_i, 0) U_0^i (R_{i+1})]$ 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 16th observation was seen, we terminated sampling when $X_{16}$ equaled 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 $\pi^*$, chose $c$ and $\pi_0$ to match the error probabilities of the TSPRT.

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

<table>
<thead>
<tr>
<th></th>
<th>TSPRT</th>
<th>Bayesian tests</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A = 0.0809$</td>
<td>$\pi^*$</td>
<td>0.925</td>
</tr>
<tr>
<td>$R = 6.84$</td>
<td>$c$</td>
<td>0.0661</td>
</tr>
<tr>
<td></td>
<td>$\pi_0$</td>
<td>0.425</td>
</tr>
</tbody>
</table>

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 $\pi^*$, though at the expense of some increase in the ATT. For small $\lambda$, 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


