A Bayesian Sequential Life Test

V. D. Barnett

Department of Statistics
University of Newcastle upon Tyne
Newcastle upon Tyne, England

Detailed examination of the lifetimes of components and assemblies by means of the usual sequential probability ratio tests is often not feasible because of the prohibitive time or cost involved in such an examination. Situations commonly arise, however, where extraneous information is available (perhaps in the form of experience of similar situations or concerning the reputation of the supplier of the components or assemblies) which reflects on the current situation. Such information might be incorporated in a Bayesian analysis, but little work seems to have been done in this area. This paper presents a possible sequential life test incorporating prior information, applied to the simplest situation of components with exponentially distributed lifetimes, tested individually. The results for conventional sequential lifetests are used as a rough yardstick against which to measure the properties of the Bayesian lifetest discussed in the paper.

Key Words
Life Test
Sequential Life Test
Bayesian Life Test
Exponential Life Times
Poisson Random Walk
Curved Absorbing Boundaries

1. Introduction

The usual test procedures available for studying the lifetimes of components or assemblies are often impracticable due to the prohibitive cost or time involved in their application. (See e.g. Barnett and Ross, 1965). They also take no account of any information that the producer or consumer may have about the quality of a product or assembly before testing commences. What is needed is a procedure which is simple to operate and assess (often by non-statisticians), which incorporates any available prior information, and which produces a speedy decision consistent with desired statistical properties. These needs point to the use of a Bayesian sequential life test procedure, and one possible simple test of this type is discussed below.

The test described ignores any loss structure or costs of conducting the test and is based merely on the way in which consecutive observations of failures modify the prior information on the mean component lifetime to produce a

Received March 1970; revised Aug. 1970.

453
posterior distribution for this mean lifetime. Sampling continues until the posterior distribution is decisive on some appropriate criterion, at which stage it terminates and a decision is taken on the quality of the component or assembly. A more sophisticated approach using decision theory ideas is undoubtedly feasible. Certain qualitative comments on the loss structure can be made but realistic quantification of losses and experimentation costs in a given situation is difficult.

In this paper we shall ignore the general decision theoretic approach to life testing, and consider merely one way in which prior information on the product lifetime may be incorporated into a life test, without regard to losses and experimentation costs.

There seems to have been little study of the Bayesian approach to life testing. Primitive attempts have been made by Pinkham and St. James (1963) and Bonis (1966). These are fixed duration procedures for products with independent, exponentially distributed lifetimes. Their models assume just two possible values for the mean life time, as does indeed the usual sequential probability ratio test (SPRT). This restriction is not only artificial, but unnecessary in the Bayesian framework. Furthermore, we want to reach a decision as quickly as possible, and it seems sensible to employ a sequential procedure to achieve this aim.

2. THE BASIC DECISION CRITERION

Suppose we denote by $\theta$ the failure rate of the component under study. As the reciprocal of the mean life time, $\theta$ provides a measure of the quality of the component. A simple sequential test of the value of $\theta$ may be constructed from the following argument.

The prior information about $\theta$ which is available may be expressed in the form of a prior probability (density) function, $\pi(\theta)$. Observations of component failures modify our information about $\theta$. As time goes on the current posterior distribution, $\pi(\theta | t)$, of $\theta$ at time $t$ settles down and becomes less and less disperse. Assuming there is some critical value $\theta_c$ for the failure rate which the producer claims to be able to meet, and which satisfies the consumer, it seems appropriate to base decisions on the current posterior probability that $\theta > \theta_c$. If this becomes small enough the consumer will be satisfied with the product—if it becomes intolerably large the producer must accept that he has not met the specification. The test will continue until either of these conditions exist. See Figure 1: here the probability that $\theta < \theta_c$ has become sufficiently small at time $t_4$ to warrant rejecting the component.

To consider this procedure in more detail we must be more specific about the prior distribution of $\theta$. We assume that prior information about $\theta$ is available from the experience of both the consumer and producer, and that this may be quantified to form a prior distribution, $\pi(\theta)$, for $\theta$. How $\pi(\theta)$ is determined in practice is not clear; one possibility is that the combined objective and subjective experience of the two parties should be quantified to the extent of mutually agreeing one or two summary measures of $\theta$ a priori. These may then be applied to a convenient parameterised family of distributions to yield an appropriate
A BAYESIAN SEQUENTIAL LIFE TEST

FIGURE 1—Realisation of test leading to rejection at time $t_0$.

$\pi(\theta)$ (See Lindley, 1965). But in any given situation it is not obvious what constitutes a “convenient” family of prior distributions, except in the narrow mathematical sense. In this latter respect many Bayesian analyses, and in particular the tests described below, are facilitated if the conjugate family of prior distributions is used i.e. in the present problem the family of distributions 'closed with respect to sampling' from the life-time distribution. But mathematical convenience is not enough, and we need to show either that the choice of prior distribution does not critically effect the properties of the test, or that in the given situation prior knowledge can be adequately represented by a member of the conjugate family. Practical casework on the choice of prior distributions in the life testing situation would be most interesting. Indeed it is vital if test engineers etc., are to be expected to adopt Bayesian life testing procedures. When purely quantitative prior information is to be processed empirical Bayes procedures may be relevant. See also Winkler (1967) on the construction of prior distributions.

The subsequent discussion assumes that the prior distribution has been chosen from the conjugate family, but there is nothing implicit in the test structure which demands this.

To illustrate the test in more detail we go over to the case of components with exponentially distributed life times, where the components are tested individually. Both fixed duration (or fixed sample size), and sequential, tests for exponential life times, without prior information, have been discussed by Epstein and Sobel (1953, 1955). Closed sequential tests for this situation have been more recently described by Hoel (1968). These various results may be...
used as a yardstick against which to measure the properties of the Bayesian life tests described below.

It should be noted that there is nothing special about the exponential case. The same type of test may be used, and analogous results and comparisons derived, in any situation where the life-time distribution admits a conjugate family of prior distributions for which probabilities are tabulated or easily calculated.

3. DETAILS OF THE TEST IN THE EXPONENTIAL CASE

We suppose that items are tested separately and have lifetimes which are independent and follow an exponential distribution

\[ f(t) = \theta e^{-\theta t}, \quad t > 0, \]  

where \( \theta \) is the failure rate.

The parameter \( \theta \) is assumed to have a prior distribution closed under sampling from (1) so that

\[ \pi(\theta) = \frac{\lambda(\lambda \theta)^{r-1}e^{-\lambda \theta}}{\Gamma(r)} \]  

where the parameters \((r, \lambda)\) characterise the appropriate member of the conjugate family. That is, \( \theta \) has a priori a gamma distribution, \( \Gamma(r, \lambda) \).

If the test is allowed to run for a total time \( t \), during which \( n \) faults have occurred, the posterior distribution of \( \theta \) is easily seen to be \( \Gamma(r + n, \lambda + t) \).

The test criterion consists of continuing observation as long as

\[ 1 - \alpha > P(\theta > \theta_0) > \beta \]  

where \( \alpha, \beta \) are previously chosen small values, say 0.05 or 0.01. As soon as (3) is violated the test terminates with rejection of the product if \( P(\theta > \theta_0) > 1 - \alpha \) or acceptance if \( P(\theta > \theta_0) < \beta \). In terms of the current time, \( t \), and number of faults, \( n \), the “continue sampling” condition (3) has the form

\[ \beta_{r,n} = \frac{1}{2\theta_0} \chi^2(\alpha, r + t) > \lambda + t > \frac{1}{2\theta_0} \chi^2(\beta, r + t) = \alpha_{r,n}, \]  

where \( \chi^2(\xi) \) is the \( \xi \)-quantile of the \( \chi^2 \) distribution with \( r \) degrees of freedom. As soon as \( \lambda + t > \beta_{r,n} \) we accept, or \( \lambda + t < \alpha_{r,n} \) we reject, the product. As for the classical sequential test, the procedure is ‘semicontinuous’ (Epstein and Sobel, 1955) in general, in that acceptance may take place at any time, rejection only on the occurrence of a fault. Denoting by \((N, T)\) the number of events, and cumulative time, respectively, the test has the form of a Poisson process taking place between two curved absorbing barriers. The prior information acts in specifying the starting point of the process \((r, \lambda)\) and may be interpreted as the equivalent sample size and test duration in our prior knowledge of \( \theta \). See Figure 2.

We should note that when we take a final decision to “accept” or “reject” we not only know that \( P(\theta > \theta_0) \) has become particularly small, or large, respectively; but also have a full measure of the “final precision” of the test.
This arises in the form of the posterior distribution of $\theta$ which is $\Gamma(t + n, \lambda + t)$, where $t$ is the time at which the decision is taken and $n$ the number of faults observed. This facility is not available with the usual type of non-Bayesian test.

To assess the properties of the test we need to study the behaviour of a Poisson process enclosed between two absorbing barriers, in order to evaluate operating characteristic (OC) and average sample number (ASN) functions for comparison with existing sequential life testing procedures. Results exist for the case of a single curved absorbing barrier (Daniels, 1963; Whittle, 1961, etc.) but there would seem to be little hope of obtaining explicit, or even useful approximate, results for this two barrier problem.\footnote{Recent results by Durbin (1971) provide a useful algorithm suitable for calculating the OC and ASN functions.}

But it is questionable what statistical interpretation can be placed on OC curves or ASN functions in this Bayesian framework even if they were available. These functions will vary with the prior information, that is with the starting point $(r, \lambda)$. The conditional OC curve or ASN function \textit{for a particular} $(r, \lambda)$
seems to have limited relevance for comparison of this test with non-Bayesian tests. However, conservative comparison of this test with the SPRT may be obtained by assuming that we have no prior information on \( \theta \), and evaluating the corresponding OC curve and ASN function. For this we might use the customarily accepted form for the nil-prior of a positive parameter

\[
\pi(\theta) = \frac{1}{\theta}, \quad \theta > 0.
\]

This is equivalent to starting the Poisson process at \((0, 0)\). There is initially no obvious reason why we should ever violate the condition (3), particularly if \( \theta = \theta_0 \). Whether or not a terminal decision need necessarily arise in this case, depends on whether or not ultimate absorption is certain for the corresponding Poisson process. Compensating for the drift in the process, we have asymptotically a drift-free random walk with absorbing barriers at \( \pm k \sqrt{n} \).

This is within the limit of the configurations of absorbing boundary for which ultimate absorption is certain (Whittle, 1965). So a terminal decision will in fact be reached, even when \( \theta = \theta_0 \), but we may well expect this decision to take a long time on occasions. The numerical results in section 5 quantify this, and compare the current test (when there is no prior information) with the SPRT. These results were obtained by simulation.

Whilst the length of the test will be reduced on average when we are not in a state of prior ignorance about \( \theta \), corresponding to the Poisson process starting away from the origin at \((r, \lambda)\) (see Figure 2), it would certainly be sensible to explore the possibility of closed procedures to avoid the occasional very long test duration. This has been studied for the classical SPRT by Armitage (1957), Anderson (1960), Schneiderman & Armitage (1962a, b), Hoel (1968) etc., with some profit. For the Bayesian test above there are various approaches that are possible. If the test goes on for a long time this means that the distribution of \( \theta \) is becoming quite localised in the region of \( \theta_0 \). In this case we should surely wish to accept the product, so we might add the condition that we accept if the variance of the posterior distribution becomes sufficiently small before (4) is violated. With an appropriate specification of "sufficiently small" this has the effect of closing the "continue sampling" region. No detailed study has been made of this proposal. Instead, we consider in the next section an alternative modification which leads to a closed test with a more direct practical interpretation, and which includes the basic test as a special case.

4. A Modified Test

If the sequences \( \{\alpha_n\} \) and \( \{\beta_n\} \) specifying the reject-and accept-boundaries are multiplied by constants \( k_0 \) and \( k_1 \) respectively then for any \( k_0 > 1 \) and \( k_1 < 1 \) these modified boundaries will intersect whatever the values of \( \alpha \) and \( \beta \). Should these modified boundaries prove sensible on practical grounds in defining a sequential test procedure, the continue sampling region will be closed and the procedure will lead fairly rapidly to a terminal decision even when \( k_0 \) and \( k_1 \) are close to 1, and even if we have no prior knowledge about \( \theta \).

We find that this procedure does in fact correspond to a quite reasonable
practical criterion for acceptance or rejection of the product. Recognising that
we must seek a compromise between the conflicting interests of the consumer
and producer, we set up a compromise region for \( \theta \), say \( \theta_1 < \theta < \theta_0 \). This region
is used in the following way in setting up a decision procedure.

*The producer says:* “If the probability that \( \theta > \theta_0 \) is sufficiently small you
(the consumer) must surely accept the product”.

*The consumer replies:* “Alright, but if the probability that \( \theta < \theta_1 \) becomes
particularly small you must allow me to reject the product”.

It is possible that both these conditions are satisfied simultaneously. To
avoid ambiguity we give precedence to the first of the conditions in this situation.

More specifically the test now proceeds as follows. Two mutually acceptable
small constants \( \alpha \) and \( \beta \) are chosen and observations are taken as long as

\[
P(\theta > \theta_0) > \beta \quad \text{and} \quad P(\theta < \theta_1) > \alpha. \tag{6}
\]

As soon as (6) is violated the test terminates with acceptance if \( P(\theta > \theta_0) < \beta \)
or, when \( P(\theta > \theta_0) > \beta \), with rejection if \( P(\theta < \theta_1) < \alpha \). Thus if the prior
distribution is \( \Gamma(r, \lambda) \) sampling is continued as long as

\[
\frac{1}{2\theta_0} \chi^2_{2(\alpha + \epsilon)} > \lambda + \epsilon > \frac{1}{2\theta_1} \chi^2_{2(\beta + \epsilon)} \tag{7}
\]

that is, see (4), as long as

\[
\beta_{r+\epsilon} > \lambda + \epsilon > k\alpha_{r+\epsilon}, \tag{8}
\]

where \( k = \frac{\theta_0}{\theta_1} \geq 1 \). As soon as \( \lambda + \epsilon > \beta_{r+\epsilon} \) we accept the product, otherwise
if \( \lambda + \epsilon < k\alpha_{r+\epsilon} \) we reject it.

The decision boundaries \( \beta_{r+\epsilon} \) and \( k\alpha_{r+\epsilon} \) now meet, for any \( k > 1 \), and the
procedure must terminate after a finite number of faults. Figure 2 shows this
effect (although, in practice, \( \theta_0 \) for the modified test will not coincide with \( \theta_0 \)
for the basic test but is likely to be larger because of the compromise nature
of the test criterion). In fact there will be a maximum number of faults that
are possible before a decision is taken, given as \( N^* \) where

\[
X_{2\alpha}^2(1 - \beta) = kX_{2\beta}^2(\epsilon). \tag{9}
\]

This maximum \( N^* \) includes, of course, the equivalent number of faults, \( r \), in
the prior distribution of \( \theta \), so that the actual maximum number of faults will
be \( n^* = N^* - r \). Needless to say, the average number of faults to produce a
decision will be considerably less than this number. Using Fisher’s approximation
to the percentage points of the \( \chi^2 \) distribution, \( N^* \) has the approximate value

\[
N^* = \frac{1}{4} \left( \left( \frac{z_{1-\alpha} - \sqrt{k} z_\epsilon}{\sqrt{k} - 1} \right)^2 + 1 \right), \tag{10}
\]

where \( z_\epsilon \) is the \( \xi \)-quantile of the standardised normal distribution.

Let us consider, through an example, what is implied by this modified decision
criterion. For ease of understanding it is simpler to argue in terms of \( \mu = 1/\theta \), rather than failure rate. Suppose the consumer wants a product with
m.t.h.f. 15,000 hours. He must compromise, and accept the product if it is
shown that \( p(\mu < 12,000) < 0.05 \), say. The producer, in return, must agree to the rejection of the product if it is earlier shown that \( p(\mu > 18,000) < 0.05 \) say. Here \( \theta_0 = (12,000)^{-1}, \theta_1 = (18,000)^{-1}, \alpha = \beta = 0.05; \) so that \( k = 1.5 \).

Admittedly this approach requires both parties to compromise with their real aims, in setting up the compromise region, \((\theta_1, \theta_2)\). In return they gain the advantage of a much more speedy decision. The compromise region accommodates the conflicting interests of the two parties and tries to reconcile them as far as possible within the aims of producing a test of manageable duration. As \( \theta_0 \rightarrow \theta_1 \), the policy is surely incontrovertible, but the decision region is open and a decision may take an intolerably long time. With \( \theta_0 \neq \theta_1 \), a decision may arise quite quickly and the decision criterion is reasonably objective and fair to both parties. A possible criticism is that such a procedure may lead to "over-design" by the producer, which conflicts with value engineering principles. But this, and other objections seem to be no more valid here than in setting up the necessary "indifference region" in the SPRT.

The compromise region, and indifference region for the corresponding SPRT, will not be expected to coincide. The basic philosophy is quite different in the two cases. Typically the compromise region will be smaller than the indifference region. Also it may be felt necessary to trade compromise for precision in reducing \( \alpha \) and \( \beta \) when \((\theta_1, \theta_2)\) is widened.

As before, the information about \( \theta \) at the moment a decision is taken is simply given by the posterior distribution of \( \theta \), viz \( \Gamma(r + n, \lambda + t) \).

5. Numerical Comparison of the Bayesian Test and the SPRT

The real test of the proposed procedure is in how it behaves in practice. In this section some numerical details of the performance characteristics of the procedure are presented, and quantitative comparisons made between the Bayesian sequential life test and an equivalent SPRT. To assess the statistical properties of the Bayesian test we need to study the probabilistic behaviour of the basic Poisson process. No explicit results appear to be possible, and the results below are obtained by simulation. Corresponding results for the SPRT are obtained by calculation from the usual approximation formulae given in the literature. (Exact results for the ASN will tend to be lower, but the few exact results available suggest that errors are likely to be small; rarely as much as 10%.) All results are accurate to the number of decimal places given. The basic test of section 3 is not considered separately, but as the limiting form of the modified test when \( k = 1 \).

Comparison of the two types of test is made difficult because of the basic difference in the nature of the tests. In the first place the OC and ASN functions, whilst providing a natural measure of the behaviour of the SPRT, have less obvious interpretations for the Bayesian test. However, these functions provide some sort of yardstick for comparison and in the absence of any better measures must be used for this purpose. It makes sense to consider OC and ASN functions only for the least favourable Bayesian test, with prior ignorance about \( \theta \). This unfairly penalises the Bayesian test, since it is designed precisely for situations where we know something about \( \theta \) a priori; and ASN functions must inevitably
overestimate those typical of the real life situation. But more fundamental than this is the difficulty of defining "equivalent" tests. The compromise and indifference regions are designed for different purposes, and strictly should not be equated for comparison of the two tests. Also the interpretation of $\alpha$ and $\beta$ is quite different to that of the probabilities of the two types of error which specify the SPRT.

For detailed discussion we consider the cases $\theta = 1$ for $k = 1, 1.5, 2.0, 2.5, 3.0$ (i.e. $\theta = 1, 1.5, 2.0, 2.5, 3.0$) and $(\alpha, \beta) = (0.05, 0.05), (0.01, 0.01)$ on the assumption that there is no prior information available on $\theta$. The complete ASN functions for the Bayesian tests are presented graphically in the lower portions of Figures 3 and 4; the upper portions of these figures show the corresponding OC functions for the various values of $k$, when $\alpha = \beta = 0.05$ and $\alpha = \beta = 0.01$, respectively.

In Table 1 we single out two quantities as summary measures of the performance of the Bayesian tests: these are the maximum number of items on test, $r + n^*$, and the maximum expected number of items on test, $\text{Max} E(r + n)$.

Note that for the basic test of section 3 (i.e. $k = 1$) both $r + n^*$, and more particularly $\text{Max} E(r + n)$, are infinite in spite of a terminal decision being inevitable.

We see immediately from the numerical results that it would not be reasonable to compare the two tests by considering a SPRT where the indifference region is the same as the compromise region $(\theta_1, \theta_0)$ for the Bayesian test, and where the risks are the same as the $(\alpha, \beta)$ for the Bayesian test. Looking at Figures 3 and 4, it is clear that the Bayesian tests have very low power in this direct comparison with the SPRT. For example, the Bayesian test with $k = 2.5$ and $\alpha = \beta = 0.05$ has operating characteristics 0.95 and 0.05 at $\theta = 0.5$ and $\theta = 4.0$, (call these $O_{25}$ and $O_5$), respectively. So for legitimate comparison we should relate this Bayesian test to the (impracticable) SPRT with $k = 8.0$ and $\alpha = \beta = 0.05$. In fact, the most discriminating Bayesian test with $\alpha = \beta = 0.05$ is the degenerate one with $k = 1$; even this only corresponds to a SPRT with $k = 0.5/0.5 = 4$ when $\alpha = \beta = 0.05$.

This suggests the following rough concept of "equivalence" for comparing the Bayesian test with the SPRT. A SPRT is equivalent to a Bayesian test if

<table>
<thead>
<tr>
<th>$k$</th>
<th>$1$</th>
<th>$1.5$</th>
<th>$2.0$</th>
<th>$2.5$</th>
<th>$3.0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha = \beta =$</td>
<td>all</td>
<td>0.01</td>
<td>0.05</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>Bayesian Test: $r + n^* = N^*$</td>
<td>$\infty$</td>
<td>132.9</td>
<td>66.6</td>
<td>46.2</td>
<td>23.2</td>
</tr>
<tr>
<td>Bayesian Test: $\text{Max} E(r + n)$</td>
<td>$\infty$</td>
<td>29.0</td>
<td>7.3</td>
<td>12.2</td>
<td>4.1</td>
</tr>
</tbody>
</table>

TABLE 1

Maximum numbers and maximum expected numbers of items on test for the Bayesian test.
it has the same $0_5$ and $0_9$ (or $0_{99}$ and $0_1$) values in its OC function. So given a Bayesian test with prescribed $(\alpha, \beta)$ and compromise region $(\theta_1, \theta_2)$ we calculate the $0_5$ and $0_9$ (or $0_{99}$ and $0_1$) points. We then take as an equivalent SPRT the one with risks $(0.05, 0.05)$ and indifference region $(0_{95}, 0_5)$ or with risks $(0.01, 0.01)$ and indifference region $(0_{99}, 0_1)$. For these equivalent SPRT's the $k$-values will be $0_5/0_{95}$ or $0_1/0_{99}$, respectively. We can now compare equivalent Bayesian tests and SPRT's in terms of their ASN functions.

To enter the range of tests of reasonable discrimination we need to reduce $\alpha$ and $\beta$ in the Bayesian test. Thus if $\alpha = \beta = 0.01$ we find that

<table>
<thead>
<tr>
<th>$k$</th>
<th>$0_{95}$</th>
<th>$0_5$</th>
<th>$0_5/0_{95}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>0.98</td>
<td>3.58</td>
<td>3.7</td>
</tr>
<tr>
<td>2.5</td>
<td>0.95</td>
<td>3.00</td>
<td>3.2</td>
</tr>
<tr>
<td>2.0</td>
<td>0.90</td>
<td>2.43</td>
<td>2.7</td>
</tr>
<tr>
<td>1.5</td>
<td>0.80</td>
<td>1.83</td>
<td>2.1</td>
</tr>
<tr>
<td>1.2</td>
<td>0.83</td>
<td>1.42</td>
<td>1.7</td>
</tr>
</tbody>
</table>

So for a fair comparison we relate the following pairs of Bayesian tests and SPRT's—the so called “equivalent tests”.

The full ASN functions for these equivalent SPRT's are shown as dotted lines on Figures 3 and 4. From these functions, Table 3 singles out for comparison the maximum expected numbers of items on test, $\text{Max } E(r + n)$, for the Bayesian tests, and the maximum expected number of faults, $\text{Max } E(R)$, for the equivalent SPRT's.

The results of Table 3 place the Bayesian sequential test in perspective. We see that for equivalent tests, the Bayesian test has a maximum expected number of items on test in all cases greater than the maximum expected number of faults for the SPRT. But several points must be recognised in this comparison.

i) The superiority of the SPRT in the maximum of the ASN function amounts to no more than just a few observations. Furthermore the ASN function for the Bayesian test has been calculated in terms of the number of items on test, rather than the number of faults; the former is typically about one more than the latter, reducing the real superiority of the SPRT by this amount. Consideration of

---

**Table 2**

<table>
<thead>
<tr>
<th>Equivalent Tests</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bayesian Test</strong></td>
</tr>
<tr>
<td>$k = 1.5; \alpha = \beta = 0.05$</td>
</tr>
<tr>
<td>$k = 3.0; \alpha = \beta = 0.01$</td>
</tr>
<tr>
<td>$k = 2.5; \alpha = \beta = 0.01$</td>
</tr>
<tr>
<td>$k = 2.0; \alpha = \beta = 0.01$</td>
</tr>
<tr>
<td>$k = 1.5; \alpha = \beta = 0.01$</td>
</tr>
</tbody>
</table>
Figure 3—Performance characteristics of tests when $\alpha = \beta = 0.05$.
Figure 4—Performance characteristics of tests when $\alpha = \beta = 0.01$. 

---equivalent SPRT
items on test rather than faults appears to be more in tune with the economic structure of the test, in that costs arise because items are placed on test rather than because they fail. This change of emphasis makes little practical difference in the types of test discussed here, but could be important when several items are placed on test simultaneously. (See Epstein and Sobel, 1955). The total duration of the test is also important from the cost point of view; depending on the nature of the components being tested, the test may be item- or duration-dominated. Results on test duration have been obtained for the Bayesian test, corresponding to those given above for numbers of items on test.

ii) There is no suggestion that the Bayesian test should be used in preference to the SPRT where there is no prior information about \( \theta \). The Bayesian test was motivated by the need for a simple sequential test, with easy interpretation of its outcome, for cases where some prior information on \( \theta \) does exist. The results of Table 3, in conjunction with remark (i) above, make the proposed Bayesian test attractive on this score. The small differences in Max \( E(r + n) \) and Max \( E(R) \) show that relatively little prior information (in terms of equivalent numbers of items on test a priori) is needed to redress the balance. In practice the Bayesian test may often be superior (see iii); particularly if we can accept what Bonis (1966) says:

'... the probability of doing good work should be estimable (a priori). After all some electronics plants have been in business for over 18 years now, and other types of firm for longer periods of time. Certainly it is ridiculous for a Statistician to assume that by the time the first article is made and passed for inspection, that it has (at most) only a 50–50 chance, say, of meeting specifications. (Pinkham & St. James, 1963, suggest) that in the absence of better information, 90% of a good design group’s products will meet design goals early in the product life cycle: this certainly seems a reasonable requirement for experienced firms.'

But the difficulty remains of knowing how to mathematically express such a conviction (should we accept it) in the form of a prior distribution for \( \theta \).

iii) In the absence of any more objective criterion, the Bayesian test has been assessed in terms of how the maximum of its ASN function compares with
that of the equivalent SPRT. This is only one aspect of comparison that might be made between the two types of test. A more comprehensive comparison can be made by considering the complete ASN functions (See Figures 3, 4).

We notice that, even when no prior information on \( \theta \) is available, there is a range of values of \( \theta \) for which the Bayesian test has smaller ASN than the SPRT. This happens for small \( \theta \). Thus the Bayesian test has the very desirable property that a good product will be more rapidly identified than on the equivalent SPRT. It seems less important that we might take a little longer to reject a bad product. The restricted SPRT’s have a similar property, see Hoel (1968).

Also, the choice of the SPRT for comparison with the Bayesian test is somewhat arbitrary. The whole spirit of the Bayesian test is to regard \( \theta \) as continuously varying, not restricted to just two values as in the SPRT. Furthermore, to enquire about the probabilities of acceptance or rejection when \( \theta \) has a specific value is artificial in the Bayesian context. Individual values of \( \theta \) vary in importance in accord with the prior distribution of \( \theta \).

However, in spite of the limitations in comparing the Bayesian test with the SPRT, some guidance is needed for the engineer in constructing a Bayesian test. The values of \( k \), \( \alpha \) and \( \beta \), obviously cannot be interpreted as the analogous quantities in the SPRT, and one possible guideline would be a table of the values of \( k \), \( \alpha \) and \( \beta \) to achieve required \( \theta_0 \) and \( \theta_{55} \). But other possibilities might be considered which are not so rigidly tied to the conventional measures of test performance such as the OC curve. In this connection we note that different combinations of \( k \), \( \alpha \) and \( \beta \), produce the same \( \theta_0 \) and \( \theta_{55} \); hence for example, the somewhat strange reversal in Max \( E(r + n) \) in Table 3 for \( k = 1.5 \), \( \alpha = \beta = 0.05 \) and \( k = 3.0 \), \( \alpha = \beta = 0.01 \). This is due to the increasing peakedness of the ASN function as \( \alpha \) and \( \beta \) increase, for fixed \( \theta_0/\theta_{55} \). Indeed, had we constructed a Bayesian test with \( \alpha = \beta = 0.01 \) to compare with the SPRT having \( \theta_0/\theta_{55} = 4.5 \), the difference between Max \( E(r + n) \) and Max \( E(R) \) for this case would certainly have been greatly reduced. Some extra criterion for choice of Bayesian test, other than the value of \( \theta_0/\theta_{55} \), is obviously needed. It would be of interest in this connection to set up plausible loss functions and compare the two types of test in terms of their expected losses. In a sense this is the dual of the comparison criterion used above; it applies a Bayesian performance characteristic to the non-Bayesian SPRT, rather than a non-Bayesian one to the Bayesian test.

6. Summary and Conclusions

a) Prompted by the need for a quick, easily applied, life-test which allows for the incorporation of prior information about the product, a Bayesian test is proposed with decision criteria based on the posterior probabilities of the product being “good” or “bad”.

b) The open nature of the decision region, and corresponding degeneracy of the ASN function for products of intermediate quality, makes this test of little practical value. A modified family of tests are constructed, including the
original one as a limiting case, using the idea of a *compromise region* for the quality of the product.

c) The modified tests have closed decision regions, so must terminate after a finite number of faults. Compared with "equivalent" SPRT's in terms of the maxima of their ASN functions the modified tests are only slightly inferior to the SPRT's when there is no prior information on product quality. Quite modest prior information renders them comparable.

d) The modified Bayesian tests appear to meet the requirements of a), but more work is needed in studying the propriety of the proposed conjugate family of prior distributions and on how an appropriate member of this family should be chosen in practice (recognising that prior information may often contain subjective value judgements). Work is also needed in relating the parameters of the Bayesian tests to meaningful practical performance characteristics of the tests.

7. Acknowledgement

My interest in this topic arose from conversations with Z. A. Lomnicki, who was also working on the problem of incorporating prior information in reliability and life-testing problems. I am grateful to him, in particular, for drawing my attention to references in the literature on earlier work in this area. I am grateful, also, to F. Downton and D. M. G. Wishart for helpful discussion, and to the referees for their comments.

References

