10

Assurance Testing

Planning for Bayesian assurance testing involves determining a test plan that guarantees that a reliability-related quantity of interest meets or exceeds a specified requirement at a desired level of confidence. Within a Bayesian hierarchical framework, this chapter determines test plans for binomial, Poisson, and Weibull testing. Also, we develop Weibull assurance test plans using available data from an associated accelerated life testing program.

10.1 Introduction

This chapter focuses on developing a test plan for assuring (or demonstrating) that, at a desired level of confidence, a reliability-related quantity of interest meets or exceeds a specified requirement. For binomial testing, we test n devices either as a demand for successful operation or for a specified length of time and observe the total number of devices failing the test x. For example, a tester may try an emergency diesel generator (EDG) to see if it will start on demand, or a tester may place a sample of a particular nonwoven material under stress for a given length of time to see if it survives the test. The reliability-related quantity of interest for both examples is the probability π that an item survives the test, and the required binomial test plan consists of the total number of devices tested n as well as the maximum allowed number of failures c.

Although practitioners often use “assuring” and “demonstrating” synonymously, Meeker and Escobar (2004) distinguishes between reliability demonstration and reliability assurance testing. A traditional *reliability demonstration test* is essentially a *classical* (i.e., frequency based) hypothesis test, which uses only the data from the test to assess whether the reliability-related quantity of interest meets or exceeds the requirement. Consider how many modern systems, such as communication devices and transportation systems, are highly reliable. For these systems, reliability demonstration tests often require an impractical amount of testing. In response to this dilemma, Meeker and Escobar (2004) defines an alternative *reliability assurance test* as one that uses additional supplementary data and information to reduce the required amount of testing. The additional data and information may include appropriate reliability models, earlier test results on the same or similar devices, expert judgment regarding performance, knowledge of the environmental conditions under which the devices are used, benchmark design information on similar devices, prior knowledge of possible failure modes, etc. Because all of the Bayesian test plans considered in this chapter use such supplementary data and information, we refer to them as reliability assurance tests.

Life testing has many aspects in common with assurance testing. However, the primary goal in designing a life test tends to be quite different than assuring conformance to a specified reliability requirement. In designing such life tests, we often have as our primary goal improving the estimation precision of certain reliability-related quantities of interest. However, such differences notwithstanding, the basic ideas underlying life and assurance testing are similar, namely, to address such questions as “How many devices do I need to test?”, “How long do I need to test each device?”, or “What is the maximum number of failures permitted for a successful test?”

Because data from an assumed sampling distribution provide the basis for deciding whether the population of products being tested meets the specified requirement, there are two kinds of errors to make. A population of unreliable products (one that does not meet the requirement) may, in fact, pass the test, whereas a reliable population may fail it. This important acknowledgment makes us think about the (probabilistic) *risks* that we incur in conducting the test. The precise form of the risks is an important consideration in classical assurance tests and is an important consideration in developing Bayesian assurance tests as well. The *test criteria* are precise probabilistic statements regarding the risks we are willing to incur when developing a test plan. The following sections discuss several of the more popular criteria.

To begin our discussion of test criteria, suppose that π denotes some reliability-related quantity of interest such that large values of π are more desirable than small values. Note that reliability is one such quantity, while the mean and quantiles of a specified failure time distribution are others. It is common to base both classical and Bayesian test plans on two specified levels of π: π_0, an *acceptable reliability level* (ARL), and π_1, a *rejectable reliability level* (RRL), where $\pi_1 \leq \pi_0$. The literature sometimes refers to the region $\pi_1 \leq \pi \leq \pi_0$ as the *indifference region*. Although the precise definition of ARL and RRL differ between the classical and Bayesian test criteria, we use them in an equivalent way.

10.1.1 Classical Risk Criteria

It is quite common to use two criteria in determining classical test plans. The *producer's risk* is the probability of failing the test when $\pi = \pi_0$, whereas the *consumer's risk* is the probability of passing the test when $\pi = \pi_1$. Suppose that we specify a maximum value, α, of the producer's risk and a maximum value, β, of the consumer's risk. For binomial testing, these criteria become

$$\begin{aligned} Producer's\ Risk &= \mathbf{P}(Test\ Is\ Failed \,|\, \pi_0) \\ &= \mathbf{P}(y > c \,|\, \pi_0) \\ &= \sum_{y=c+1}^{n} \binom{n}{y} (1-\pi_0)^y \pi_0^{n-y} \le \alpha\,, \end{aligned} \tag{10.1}$$

and

$$\begin{aligned} Consumer's\ Risk &= \mathbf{P}(Test\ Is\ Passed \,|\, \pi_1) \\ &= \mathbf{P}(y \le c \,|\, \pi_1) \\ &= \sum_{y=0}^{c} \binom{n}{y} (1-\pi_1)^y \pi_1^{n-y} \le \beta\,, \end{aligned} \tag{10.2}$$

where $\pi_1 \le \pi_0, n$ is the number of test units, and c is the maximum number of failures allowed.

To choose a test plan for specified values of $(\alpha, \pi_0, \beta, \pi_1)$, we find the required binomial test plan (n, c) by simultaneously solving Eqs. 10.1 and 10.2. Numerous textbooks provide additional details of this purely classical approach, for example, see Tobias and Trindade (1995).

10.1.2 Average Risk Criteria

Easterling (1970) first proposed using average operating characteristics and corresponding risk criteria. These risk criteria are similar to the classical criteria in Sect. 10.1.1, except that now we condition on the events $\pi \ge \pi_0$ and $\pi \le \pi_1$, respectively. To do this requires a suitable prior distribution for π, as specified by $p(\pi)$. The *average producer's risk* is the probability of failing the test when $\pi \ge \pi_0$. Choosing a maximum allowable average producer's risk α, the binomial test plan (n, c) is

$$\begin{aligned} Average\ Producer's\ Risk &= \mathbf{P}(Test\ Is\ Failed \,|\, \pi \ge \pi_0) \\ &= \frac{\mathbf{P}(y > c, \pi \ge \pi_0)}{\mathbf{P}(\pi \ge \pi_0)} \\ &= \frac{\int_{\pi_0}^{1} \left[\sum_{y=c+1}^{n} \binom{n}{y} (1-\pi)^y \pi^{n-y}\right] p(\pi) \mathrm{d}\pi}{\int_{\pi_0}^{1} p(\pi) \mathrm{d}\pi} \\ &= \frac{\int_{\pi_0}^{1} \left[1 - \sum_{y=0}^{c} \binom{n}{y} (1-\pi)^y \pi^{n-y}\right] p(\pi) \mathrm{d}\pi}{\int_{\pi_0}^{1} p(\pi) \mathrm{d}\pi} \le \alpha\,. \end{aligned} \tag{10.3}$$

Likewise, the corresponding *average consumer's risk* is the probability of passing the test when $\pi \leq \pi_1$. Choosing a maximum allowable average consumer's risk β, the binomial test plan (n, c) is

$$\begin{aligned} Average\ Consumer's\ Risk &= \mathbf{P}(Test\ Is\ Passed \,|\, \pi \leq \pi_1) \\ &= \frac{\mathbf{P}(y \leq c, \pi \leq \pi_1)}{\mathbf{P}(\pi \leq \pi_1)} \\ &= \frac{\int_0^{\pi_1} \left[\sum_{y=0}^{c} \binom{n}{y} (1-\pi)^y \pi^{n-y}\right] p(\pi) \mathrm{d}\pi}{\int_0^{\pi_1} p(\pi) \mathrm{d}\pi} \leq \beta\,. \end{aligned} \tag{10.4}$$

Martz and Waller (1982) discusses the use of these risks and recommends taking care in applying these criteria. For example, the average consumer's risk may be a poor indication of what is likely really desired; namely, a maximum probability β that $\pi \leq \pi_1$ for a test that passes. The average consumer's risk given in Eq. 10.4 may be substantially larger than this desired maximum conditional probability β. The prior probability that $\pi \leq \pi_1$ may be quite small; however, if indeed $\pi \leq \pi_1$, the probability of passing the test may be large. In such a situation, using the average consumer's risk may be inappropriate, and therefore misleading.

10.1.3 Posterior Risk Criteria

We now consider fully Bayesian posterior risks that convey a completely different outlook from the corresponding classical or average risks. While the classical or average risks provide assurance that satisfactory devices will pass the test and that unsatisfactory devices will fail it, posterior risks provide precisely the assurance that practitioners often desire: if the test is passed, then the consumer desires a maximum probability β that $\pi \leq \pi_1$. On the other hand, if the test is failed, then the producer desires a maximum probability α that $\pi \geq \pi_0$. Unlike the average risks, these posterior risks are fully Bayesian in the sense that they are subjective probability statements about π.

For a test that fails, the *posterior producer's risk* is the probability that $\pi \geq \pi_0$, or $\mathbf{P}(\pi \geq \pi_0 \,|\, Test\ Is\ Failed)$. Notice that this is simply the posterior probability that $\pi \geq \pi_0$ given that we have observed more than c failures. Using Bayes' Theorem, and assuming a maximum allowable posterior producer's risk α, an expression for the posterior producer's risk for the binomial test plan (n, c) is

$$\begin{aligned} Posterior\ Producer's\ Risk &= \mathbf{P}(\pi \geq \pi_0 \,|\, Test\ Is\ Failed) \\ &= \int_{\pi_0}^{1} p(\pi \,|\, y > c) \mathrm{d}\pi \\ &= \int_{\pi_0}^{1} \frac{f(y > c \,|\, \pi) p(\pi)}{\int_0^1 f(y > c \,|\, \pi) p(\pi) \mathrm{d}\pi} \, \mathrm{d}\pi \end{aligned} \tag{10.5}$$

$$= \frac{\int_{\pi_0}^{1} \left[\sum_{y=c+1}^{n} \binom{n}{y} (1-\pi)^y \pi^{n-y} \right] p(\pi) \mathrm{d}\pi}{\int_0^1 \left[\sum_{y=c+1}^{n} \binom{n}{y} (1-\pi)^y \pi^{n-y} \right] p(\pi) \mathrm{d}\pi}$$

$$= \frac{\int_{\pi_0}^{1} \left[1 - \sum_{y=0}^{c} \binom{n}{y} (1-\pi)^y \pi^{n-y} \right] p(\pi) \mathrm{d}\pi}{1 - \int_0^1 \left[\sum_{y=0}^{c} \binom{n}{y} (1-\pi)^y \pi^{n-y} \right] p(\pi) \mathrm{d}\pi} \le \alpha.$$

Similarly, given that the test is passed, the *posterior consumer's risk* is the probability that $\pi \le \pi_1$, or $\mathbf{P}(\pi \le \pi_1 \,|\, Test\ Is\ Passed)$. Notice that this is simply the posterior probability that $\pi \le \pi_1$ given that we have observed no more than c failures. Using Bayes' Theorem, and assuming a maximum allowable posterior consumer's risk β, an expression for the posterior consumer's risk for the binomial test plan (n, c) is

$$\begin{aligned} Posterior\ Consumer's\ Risk &= \mathbf{P}(\pi \le \pi_1 \,|\, Test\ Is\ Passed) \qquad (10.6) \\ &= \int_0^{\pi_1} p(\pi \,|\, y \le c) \mathrm{d}\pi \\ &= \int_0^{\pi_1} \frac{f(y \le c \,|\, \pi) p(\pi)}{\int_0^1 f(y \le c \,|\, \pi) p(\pi) \mathrm{d}\pi} \mathrm{d}\pi \\ &= \frac{\int_0^{\pi_1} \left[\sum_{y=0}^{c} \binom{n}{y} (1-\pi)^y \pi^{n-y} \right] p(\pi) \mathrm{d}\pi}{\int_0^1 \left[\sum_{y=0}^{c} \binom{n}{y} (1-\pi)^y \pi^{n-y} \right] p(\pi) \mathrm{d}\pi} \le \beta. \end{aligned}$$

Example 10.1 Binomial test plan for new modems. Consider finding a binomial test plan using the posterior consumer's risk criterion. Hart (1990) develops a reliability assurance test for a new modem, denoted by B, that is similar to an earlier modem, denoted by A. Modem A is currently in production and is very reliable. The major difference between the two modems is that B operates at a different frequency than A. Also, the same production line that builds A will produce B and both modems use most of the same components. Further, Hart (1990) reports that a binomial assurance test for modem A on 150 units yielded 6 failures.

One of the test objectives is to show that, after successful testing, the 0.1 quantile of the posterior reliability distribution for B is at least 0.938, the 0.1 quantile of A's posterior reliability distribution. Similar to Hart (1990), we use a $Beta[86.4, 3.6] = Beta[(0.6 \times 150)(144/150), (0.6 \times 150)(6/150)]$ prior distribution for π. This prior distribution arises from treating an A test as "worth" 60% of a B test or $90 = 0.6 \times 150$ total tests. Note that the 0.1 quantile of this prior distribution is 0.932, which is only slightly smaller than the requirement. Therefore, we anticipate that the test plan will require only a small sample of B modems.

A *minimum sample size* (or *zero-failure*) test plan is one in which we test n modems and state that the test is passed if there are no failures, that is,

$c = 0$. Given our $Beta(86.4, 3.6)$ prior distribution, $\pi_1 = 0.938, \beta = 0.10$, and $c = 0$, we find the desired Bayesian zero-failure test plan by solving Eq. 10.6 for sample size or number of tests n. Using Eq. 10.6 yields the expression

$$\begin{aligned}
\mathbf{P}(\pi \le 0.938 \,|\, Test\ Is\ Passed) \\
&= \frac{\int_0^{0.938} \binom{n}{0} (1-\pi)^0 \pi^n p(\pi) \mathrm{d}\pi}{\int_0^1 \binom{n}{0} (1-\pi)^0 \pi^n p(\pi) \mathrm{d}\pi} \\
&= \frac{\int_0^{0.938} \pi^n \frac{\Gamma(86.4+3.6)}{\Gamma(86.4)\Gamma(3.6)} \pi^{86.4-1} (1-\pi)^{3.6-1} \mathrm{d}\pi}{\int_0^1 \pi^n \frac{\Gamma(86.4+3.6)}{\Gamma(86.4)\Gamma(3.6)} \pi^{86.4-1} (1-\pi)^{3.6-1} \mathrm{d}\pi} \\
&= I(0.938; 86.4+n, 3.6) \le 0.10\,,
\end{aligned} \tag{10.7}$$

where $I(z; \alpha, \beta)$ denotes the *incomplete beta function ratio*. Upon evaluating the incomplete beta function ratio in Eq. 10.7 for increasing values of n, we find that $n = 9$ is the smallest integer that satisfies the inequality. Consequently, the plan consists of testing 9 B modems. If none fail, we can then claim that $\mathbf{P}(\pi \le 0.938 \,|\, No\ Failures\ in\ 9\ Tests) = 0.097 < 0.10$, as required. In this case, given no failures in the 9 B modem tests, the 0.1 quantile of the posterior distribution of π is 0.9384. Finally, the unconditional probability of passing the test is simply

$$\begin{aligned}
\mathbf{P}[Test\ Is\ Passed] &= \int_0^1 \binom{n}{0} (1-\pi)^0 \pi^n p(\pi) \mathrm{d}\pi \\
&= \frac{\Gamma(86.4+n)\Gamma(3.6)}{\Gamma(86.4+3.6+n)} \frac{\Gamma(86.4+3.6)}{\Gamma(86.4)\Gamma(3.6)} \\
&= \frac{\Gamma(95.4)\Gamma(90)}{\Gamma(99)\Gamma(86.4)} = 0.70.
\end{aligned}$$

10.2 Binomial Testing

Now consider both the average and posterior risks for the binomial sampling distribution within a hierarchical framework. Suppose that we have failure count data from $m > 1$ situations, such as m different plants. Let x_i denote the observed number of failures in a sample of size n_i for $i = 1, \ldots, m$, and let $\mathbf{x}$ represent all the observed failure count data. Then, conditional on the success probability π_i, assume that the X_i are conditionally independent and that $X_i \,|\, \pi_i \sim Binomial(n_i, 1 - \pi_i)$. Also assume that the π_i can be modeled hierarchically — specifically, that given δ and γ they are independent and identically distributed (i.i.d.) with a common $Beta(\delta, \gamma)$ distribution. Finally, we specify a prior distribution for the hyperparameters (δ, γ), denoted by $p(\delta, \gamma)$, which is a proper (but usually diffuse) joint distribution.

10.2.1 Binomial Posterior Consumer's and Producer's Risks

Given that there are observed data $X_i \sim Binomial(n_i, 1-\pi_i)$, suppose that we are interested in developing a binomial test plan (n, c) using the posterior risk criteria as our test criteria. Our test plan is for a situation "similar" to those previously observed, where we describe similarity by assuming that for the new situation, an item has probability π of surviving the test, where given δ and γ, π and the π_i are i.i.d. $Beta(\delta, \gamma)$.

Recall that both the posterior producer's risk and posterior consumer's risk specify criteria on the posterior distribution for the binomial probability of success π. Since there are now observed data $\mathbf{x}$, we condition on that data and now use $p(\pi \,|\, \mathbf{x})$ in place of $p(\pi)$ in Eqs. 10.5 and 10.6. In particular, for a binomial test plan (n, c), an expression for the posterior producer's risk is

$$\begin{aligned} &\mathbf{P}(\pi \geq \pi_0 \mid \textit{Test Is Failed}, \mathbf{x}) \\ &\quad = \frac{\int_{\pi_0}^{1} \left[1 - \sum_{y=0}^{c} \binom{n}{y} (1-\pi)^y \pi^{n-y}\right] p(\pi \,|\, \mathbf{x}) \mathrm{d}\pi}{1 - \int_0^1 \left[\sum_{y=0}^{c} \binom{n}{y} (1-\pi)^y \pi^{n-y}\right] p(\pi \,|\, \mathbf{x}) \mathrm{d}\pi} . \end{aligned} \tag{10.8}$$

Similarly, the posterior consumer's risk is

$$\begin{aligned} &\mathbf{P}(\pi \leq \pi_0 \mid \textit{Test Is Passed}, \mathbf{x}) \\ &\quad = \frac{\int_{0}^{\pi_1} \left[\sum_{y=0}^{c} \binom{n}{y} (1-\pi)^y \pi^{n-y}\right] p(\pi \,|\, \mathbf{x}) \mathrm{d}\pi}{\int_0^1 \left[\sum_{y=0}^{c} \binom{n}{y} (1-\pi)^y \pi^{n-y}\right] p(\pi \,|\, \mathbf{x}) \mathrm{d}\pi} . \end{aligned} \tag{10.9}$$

Now we must determine how to evaluate these criteria. There are two equivalent ways to approach the problem. First, notice that both of the posterior risk criteria are probability statements about the posterior distribution of π given different data. For a given choice of (n, c), the posterior producer's risk can be calculated by using Markov chain Monte Carlo (MCMC) to find the posterior distribution of π given $\mathbf{x}$ and $y > c$ and then calculating the proportion of posterior draws with $\pi \geq \pi_0$. Similarly, for a given choice of (n, c), the posterior consumer's risk can be calculated by using MCMC to find the posterior distribution of π given $\mathbf{x}$ and $y \leq c$ and then calculating the proportion of posterior draws with $\pi \leq \pi_1$. Notice, however, that for each choice of (n, c), this requires using MCMC to calculate two posterior distributions for π.

Suppose instead that we condition only the data $X_i \sim Binomial(n_i, 1-\pi_i)$ with $\pi_i \sim Beta(\delta, \gamma)$ and use MCMC to obtain the posterior predictive distribution $p(\pi \,|\, \mathbf{x})$. We can obtain draws from the posterior predictive distribution $p(\pi \,|\, \mathbf{x})$ using the N posterior draws for (δ, γ) by drawing $\pi^{(j)} \sim Beta(\delta^{(j)}, \gamma^{(j)})$, and then use these samples to evaluate Eqs. 10.8 and 10.9 using Monte Carlo integration. In general, to evaluate $\mathrm{E}[g(x)] = \int g(x)p(x)dx$, obtain a random sample $x_1, \ldots, x_N$ from $p(x)$ and approximate the expectation as $\frac{1}{N}\sum_{i=1}^{N} g(x_i)$.

We evaluate the posterior producer's risk as

$$\mathbf{P}[\pi \geq \pi_1 \,|\, \mathit{Test\ Is\ Failed}, \mathbf{x}] \approx \frac{\frac{1}{N}\sum_{j=1}^{N}\left[1-\sum_{y=0}^{c}\binom{n}{y}(1-\pi^{(j)})^{y}(\pi^{(j)})^{n-y}\right] I(\pi^{(j)} \geq \pi_0)}{1-\frac{1}{N}\sum_{j=1}^{N}\left[\sum_{y=0}^{c}\binom{n}{y}(1-\pi^{(j)})^{y}(\pi^{(j)})^{n-y}\right]},$$

and the posterior consumer's risk as

$$\mathbf{P}[\pi \leq \pi_1 \,|\, \mathit{Test\ Is\ Passed}, \mathbf{x}] \approx \frac{\sum_{j=1}^{N}\left[\sum_{y=0}^{c}\binom{n}{y}(1-\pi^{(j)})^{y}(\pi^{(j)})^{n-y}\right] I(\pi^{(j)} \leq \pi_1)}{\sum_{j=1}^{N}\left[\sum_{y=0}^{c}\binom{n}{y}(1-\pi^{(j)})^{y}(\pi^{(j)})^{n-y}\right]}.$$

The expression for the unconditional probability of passing the test is

$$\mathbf{P}(\mathit{Test\ Is\ Passed} \,|\, \mathbf{x}) \approx \frac{1}{N}\sum_{j=1}^{N}\left[\sum_{y=0}^{c}\binom{n}{y}(1-\pi^{(j)})^{y}(\pi^{(j)})^{n-y}\right]. \tag{10.10}$$

Let

$$b^{(j)}(y) = \frac{\binom{n}{y}B(n-y+\delta^{(j)}, y+\gamma^{(j)})}{B(\delta^{(j)}, \gamma^{(j)})},$$

where $B(\alpha, \beta)$ is the beta function. We can also evaluate the posterior producer's risk using the posterior draws $\delta^{(j)}, \gamma^{(j)} \,|\, \mathbf{x}$ as

$$\mathbf{P}[\pi \geq \pi_0 \,|\, \mathit{Test\ Is\ Failed}, \mathbf{x}] \approx \frac{\sum_{j=1}^{N}\left[1-I(\pi_0;\delta^{(j)},\gamma^{(j)})-\sum_{y=0}^{c} b^{(j)}(y)[1-I(\pi_0; n-y+\delta^{(j)}, y+\gamma^{(j)})]\right]}{\sum_{j=1}^{N}\left[1-\sum_{y=0}^{c} b^{(j)}(y)\right]},$$

and the posterior consumer's risk as

$$\mathbf{P}[\pi \leq \pi_1 \,|\, \mathit{Test\ Is\ Passed}, \mathbf{x}] \approx \frac{\sum_{j=1}^{N}\left[\sum_{y=0}^{c} b^{(j)}(y) I(\pi_1; n-y+\delta^{(j)}, y+\gamma^{(j)})\right]}{\sum_{j=1}^{N}\sum_{y=0}^{c} b^{(j)}(y)},$$

where $I(z;\alpha,\beta)$ is the incomplete beta function ratio. An additional expression for the unconditional probability of passing the test is

$$\begin{aligned}\mathbf{P}(\mathit{Test\ Is\ Passed} \,|\, \mathbf{x}) &\approx \frac{1}{N}\sum_{j=1}^{N}\sum_{y=0}^{c} b^{(j)}(y)\\ &\approx \frac{1}{N}\sum_{j=1}^{N}\left[\sum_{y=0}^{c}\frac{\binom{n}{y}B(n-y+\delta^{(j)}, y+\gamma^{(j)})}{B(\delta^{(j)}, \gamma^{(j)})}\right].\end{aligned}$$

To obtain a test plan, simultaneously solve the pair of inequalities given by

$$\mathbf{P}[\pi \geq \pi_0 \,|\, Test\ Is\ Failed, \mathbf{x}] \leq \alpha \tag{10.11}$$

and

$$\mathbf{P}[\pi \leq \pi_1 \,|\, Test\ Is\ Passed, \mathbf{x}] \leq \beta, \tag{10.12}$$

for the pair of integers (n, c), where $0 \leq c < n$, and where α and β are the desired maximum posterior producer's and consumer's risks.

We can find such test plans because Eqs. 10.11 and 10.12 have opposite effects. For fixed c, as n increases, $\mathbf{P}[\pi \leq \pi_1 \,|\, Test\ Is\ Passed, \mathbf{x}]$ decreases, whereas $\mathbf{P}[\pi \geq \pi_0 \,|\, Test\ Is\ Failed, \mathbf{x}]$ increases. On the other hand, for fixed n, as c increases, the opposite is true. Consequently, we can use the algorithm in Fig. 10.1 to find the required test plan.

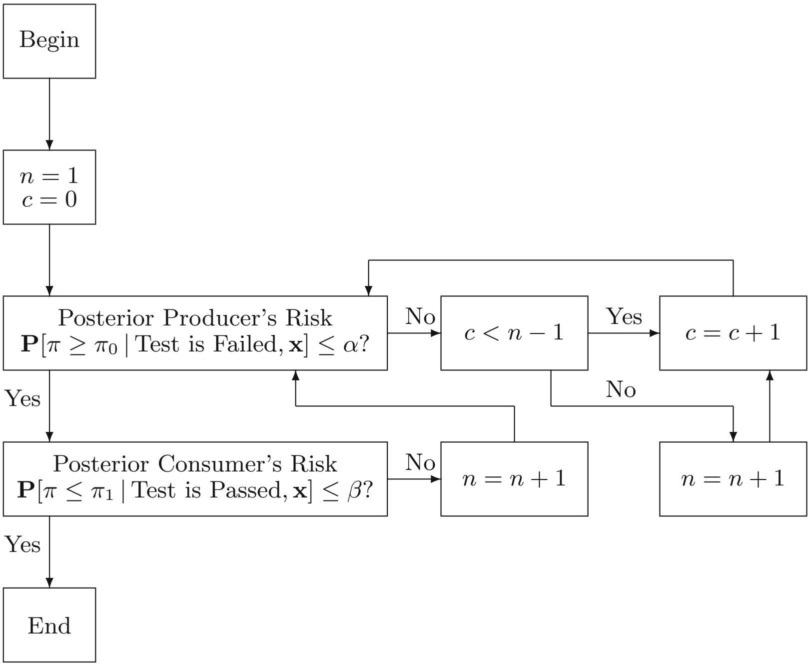

Fig. 10.1. An algorithm for finding Bayesian hierarchical binomial test plans.

Example 10.2 Hierarchical binomial test plan for EDGs. Martz et al. (1996) analyzes failure count data from EDGs in 63 U.S. commercial nuclear power plants to assess their reliability of accepting the electrical load and run (*load-run*) on demand. See Table 10.1, which summarizes the data collected over the four-year period 1988-1991.

Table 10.1. EDG load-run demand data (x failures out of n demands) (Martz et al., 1996)

Plant	x	n	Plant	x	n	Plant	x	n
1	11	854	22	0	238	43	5	216
2	2	373	23	1	370	44	0	252
3	5	618	24	2	302	45	1	419
4	5	157	25	3	152	46	0	136
5	3	542	26	2	294	47	3	185
6	2	202	27	0	101	48	7	382
7	0	65	28	0	283	49	6	304
8	2	166	29	2	117	50	2	130
9	4	574	30	0	115	51	1	121
10	2	201	31	1	196	52	2	295
11	8	388	32	0	310	53	1	289
12	2	287	33	4	134	54	0	181
13	3	431	34	2	132	55	2	150
14	5	358	35	0	242	56	0	334
15	14	1120	36	4	132	57	2	263
16	1	321	37	4	320	58	0	92
17	3	225	38	0	996	59	2	466
18	8	433	39	1	253	60	2	387
19	7	468	40	13	704	61	1	183
20	1	218	41	2	151	62	5	278
21	9	317	42	2	385	63	0	212

We model the failure counts X_i as conditionally independent given π_i with $Binomial(n_i, 1-\pi_i)$ distributions, where π_i is the ith plant reliability. Because the plants have all been built and operated to the same Nuclear Regulatory Commission (NRC)-controlled safety standards, model the 63 $\pi_i \mid \delta, \gamma$ as i.i.d. with a common $Beta(\delta, \gamma)$ distribution. In our analysis of these data, we use independent and diffuse $InverseGamma(0.1, 0.1)$ prior distributions for δ and γ. See Table 10.2, which summarizes the marginal posterior distributions for δ and γ and the predictive distribution for π. The hierarchical binomial model fits the load-run demand data well (see Exercise 4.13).

Using the posterior risk criteria, suppose now that we want to find the test plan having the posterior consumer's and producer's risks $\beta = 0.05$ for $\pi_1 = 0.985$, and $\alpha = 0.10$ for $\pi_0 = 0.999$, respectively.

Using the algorithm in Fig. 10.1, we obtain the required test plan $n = 42$ and $c = 0$. For this test plan, the actual posterior consumer's risk $\mathbf{P}[\pi \leq 0.985 \mid Test\ Is\ Passed, \mathbf{x}] = 0.0014$ and the actual posterior producer's risk $\mathbf{P}[\pi \geq 0.999 \mid Test\ Is\ Failed, \mathbf{x}] = 0.0992$. Therefore, at the new plant, load and run the EDGs 42 times. (Note that we are also assuming here that

Table 10.2. Posterior distribution summaries for the hyperparameters δ and γ and the predictive distribution for π

			Quantiles				
Parameter	Mean	Std Dev	0.025	0.05	0.50	0.95	0.975
δ	286.4	135.6	121.5	136.9	253.2	557.4	652.8
γ	2.710	1.243	1.226	1.360	2.402	5.171	6.204
π	0.9903	0.0067	0.9735	0.9778	0.9917	0.9981	0.9987

all the EDGs at the new plant have the same load-run demand reliability.) If there are no failures, then we can assure an EDG load-run demand reliability of 0.985 with a 0.95 probability. On the other hand, if the test is failed, we have the assurance that the EDG load-run demand reliability is not greater than 0.999 with a 0.90 probability. Finally, from Eq. 10.10, the unconditional probability of passing this test is approximately 0.69.

10.2.2 Hybrid Risk Criterion

Fitzgerald et al. (1999) uses the posterior consumer's risk criterion in Eq. 10.6 and the average producer's risk criterion in Eq. 10.3 to determine alternative binomial test plans. These "hybrid" criteria are appropriate in situations where we are interested in a single specified level of reliability, which Fitzgerald et al. (1999) refers to as the *target reliability level*. Fitzgerald et al. (1999) reasons that, in many applications of reliability assurance testing, the consumer and producer are one and the same. Consequently, in such situations, there may be less adversarial tension in determining appropriate risk criteria, and therefore, little motivation to specify two different (and acceptable) levels of reliability (π_0 and π_1) for the required test plan.

Likewise, we use this hybrid approach and seek a Bayesian binomial test plan for which there is a probability of at most β that $\pi \leq \pi^*$ for a passing test, and simultaneously, there is a probability of at most α that a test is failed given $\pi \geq \pi^*$.

Notationally, we want to find the test plan (n, c) that satisfies the hybrid risk criteria given by

$$\mathbf{P}[\mathit{Test\ Is\ Failed} \,|\, \pi \geq \pi^*, \mathbf{x}] \leq \alpha \tag{10.13}$$

and

$$\mathbf{P}[\pi \leq \pi^* \,|\, \mathit{Test\ Is\ Passed}, \mathbf{x}] \leq \beta. \tag{10.14}$$

For test plan (n, c), we can approximate the average producer's risk by

$$\mathbf{P}[\mathit{Test\ Is\ Failed} \,|\, \pi \geq \pi^*, \mathbf{x}] \approx \tag{10.15}$$

$$\frac{\sum_{j=1}^{N} \left[1 - I(\pi^*; \delta^{(j)}, \gamma^{(j)}) - \sum_{y=0}^{c} b^{(j)}(y)[1 - I(\pi^*; n - y + \delta^{(j)}, y + \gamma^{(j)})]\right]}{\sum_{j=1}^{N} \left[1 - I(\pi^*; \delta^{(j)}, \gamma^{(j)})\right]}.$$

We find the desired Bayesian test plan by employing the algorithm in Fig. 10.1, where Eq. 10.15 replaces the producer's risk inequality statement used in the algorithm.

Example 10.3 Hierarchical binomial test plan using hybrid criteria. Again consider Example 10.2. Consider finding a Bayesian test plan having the posterior consumer's and average producer's risks $\beta = 0.05$ and $\alpha = 0.05$ for $\pi^* = 0.98$, respectively. Using the algorithm in Fig. 10.1 with the modification described above, we obtain the required test plan $n = 71$ and $c = 2$. For this test plan, the actual posterior consumer's risk is $\mathbf{P}[\pi \leq 0.98 \,|\, Test\ Is\ Passed, \mathbf{x}] = 0.0497$ and the actual average producer's risk $\mathbf{P}[Test\ Is\ Failed \,|\, \pi \geq 0.98, \mathbf{x}] = 0.0333$. Because we now only consider a single target reliability level, this test plan requires significantly more testing than that found in Example 10.2, in which two different reliability values were specified. From Eq. 10.10, the unconditional probability of passing this test is approximately 0.951.

10.3 Poisson Testing

We now move from binomial testing within a hierarchical framework to Poisson testing. Recall from Chap. 6 that a homogeneous Poisson process generates a sequence of events for which the times between successive failures (the *inter-failure times*) are independently and identically $Exponential(\lambda)$ distributed. For a particular Poisson testing situation, let us develop a Bayesian test plan for assuring that the *failure rate* λ does not exceed a specified requirement. From Poisson process theory, the number of failures X occurring in fixed total test or operating time T has a $Poisson(\lambda T)$ distribution. Let T represent the total operating or exposure time of the devices during which we assume that devices are either repaired or replaced when they fail.

Suppose that the plan tests n devices for the length of time t_0, replacing devices as they fail, so that the total operating time is $T = nt_0$. In determining a Bayesian test plan, we seek (T, c), where c is the maximum allowed number of failures. The test is passed if no more than c failures occur in total test time T. Note that any combination of n and t_0 satisfying $T = nt_0$ provides an acceptable test plan.

As with binomial testing, let us assume in Poisson testing that there are data available from $m > 1$ situations. Let x_i denote the observed number of failures in total operating time T_i for the ith situation, and let $\mathbf{x}$ represent all the observed failure data. Then, conditioning on λ_i, $X_i \sim Poisson(\lambda_i T_i)$, where the X_i are conditionally independent. We model the λ_i hierarchically, assuming they are i.i.d. $Gamma(\eta, \kappa)$, given η and κ, and specify a prior distribution for the hyperparameters (η, κ), denoted by $p(\eta, \kappa)$.

We can write an expression for the posterior producer's risk for the Poisson test plan (T, c), where λ_0 is the rejectable failure rate and there is a maximum allowable posterior producer's risk α:

$$\begin{aligned} \mathit{Posterior\ Producer's\ Risk} &= \mathbf{P}(\lambda \leq \lambda_0 \,|\, \mathit{Test\ Is\ Failed}) \\ &= \int_0^{\lambda_0} p(\lambda \,|\, y > c)\mathrm{d}\lambda \\ &= \int_0^{\lambda_0} \frac{f(y > c \,|\, \lambda)p(\lambda)}{\int_0^\infty f(y > c \,|\, \lambda)p(\lambda)\mathrm{d}\lambda}\, \mathrm{d}\lambda \\ &= \frac{\int_0^{\lambda_0} \left[1 - \sum_{y=0}^{c} \frac{(\lambda T)^y \exp(-\lambda T)}{y!}\right] p(\lambda)\mathrm{d}\lambda}{\int_0^{\infty} \left[1 - \sum_{y=0}^{c} \frac{(\lambda T)^y \exp(-\lambda T)}{y!}\right] p(\lambda)\mathrm{d}\lambda} \leq \alpha. \end{aligned} \tag{10.16}$$

Similarly, given that the test is passed, we can write an expression for the posterior consumer's risk for the Poisson test plan (T, c), with acceptable failure rate λ_1 and maximum allowable posterior consumer's risk β, as

$$\begin{aligned} \mathit{Posterior\ Consumer's\ Risk} &= \mathbf{P}(\lambda \geq \lambda_1 \,|\, \mathit{Test\ Is\ Passed}) \\ &= \int_{\lambda_1}^{\infty} p(\lambda \,|\, y \leq c)\mathrm{d}\lambda \\ &= \int_{\lambda_1}^{\infty} \frac{f(y \leq c \,|\, \lambda)p(\lambda)}{\int_0^\infty f(y \leq c \,|\, \lambda)p(\lambda)\mathrm{d}\lambda}\, \mathrm{d}\lambda \\ &= \frac{\int_{\lambda_1}^{\infty} \left[\sum_{y=0}^{c} \frac{(\lambda T)^y \exp(-\lambda T)}{y!}\right] p(\lambda)\mathrm{d}\lambda}{\int_0^{\infty} \left[\sum_{y=0}^{c} \frac{(\lambda T)^y \exp(-\lambda T)}{y!}\right] p(\lambda)\mathrm{d}\lambda} \leq \beta. \end{aligned} \tag{10.17}$$

Since there are available data $\mathbf{x}$, we use the posterior distribution for $\lambda, p(\lambda \,|\, \mathbf{x})$, in Eqs. 10.16 and 10.17 to construct our test plan for a new situation. We can calculate the posterior producer's risk either using posterior predictive draws $\lambda^{(j)}$ or using posterior draws $(\eta^{(j)}, \kappa^{(j)})$. Let

$$g^{(j)}(y) = \frac{(\kappa^{(j)})^{\eta^{(j)}} T^y}{y! \Gamma(\eta^{(j)})(T + \kappa^{(j)})^{y + \eta^{(j)}}}.$$

$$\begin{aligned} &\mathbf{P}[\lambda \leq \lambda_0 \,|\, \mathit{Test\ Is\ Failed}, \mathbf{x}] \\ &\quad \approx \frac{\sum_{j=1}^{N} \left[1 - \sum_{y=0}^{c} \frac{(\lambda^{(j)} T)^y \exp(-\lambda^{(j)} T)}{y!}\right] I(\lambda^{(j)} \leq \lambda_0)}{\sum_{j=1}^{N} \left[1 - \sum_{y=0}^{c} \frac{(\lambda^{(j)} T)^y \exp(-\lambda^{(j)} T)}{y!}\right]} \\ &\quad \approx \frac{\sum_{j=1}^{N} \left[\gamma(\eta^{(j)}, \kappa^{(j)}\lambda_0)/\Gamma(\eta^{(j)}) - \sum_{y=0}^{c} g^{(j)}(y)\gamma[y + \eta^{(j)}, (T + \kappa^{(j)})\lambda_0]\right]}{\sum_{j=1}^{N} \left[1 - \sum_{y=0}^{c} g^{(j)}(y)\Gamma(y + \eta^{(j)})\right]}, \end{aligned}$$

where $\gamma(q, z)$ denotes the *lower incomplete gamma function*.

The expressions for the posterior consumer's risk are

$$\begin{aligned}
&\mathbf{P}[\lambda \geq \lambda_1 \,|\, \mathit{Test\ Is\ Passed}, \mathbf{x}] \\
&\quad \approx \frac{\sum_{j=1}^{N} \left[\sum_{y=0}^{c} \frac{(\lambda^{(j)}T)^y \exp(-\lambda^{(j)}T)}{y!}\right] I(\lambda^{(j)} \geq \lambda_1)}{\sum_{j=1}^{N} \left[\sum_{y=0}^{c} \frac{(\lambda^{(j)}T)^y \exp(-\lambda^{(j)}T)}{y!}\right]} \\
&\quad \approx \frac{\sum_{j=1}^{N} \left[\sum_{y=0}^{c} g^{(j)}(y)\{\Gamma(y+\eta^{(j)}) - \gamma[y+\eta^{(j)}, (T_0+\kappa^{(j)})\lambda_1]\}\right]}{\sum_{j=1}^{N} \left[\sum_{y=0}^{c} g^{(j)}(y)\Gamma(y+\eta^{(j)})\right]}.
\end{aligned}$$

We can also write the unconditional probability of passing the test as

$$\begin{aligned}
\mathbf{P}[\mathit{Test\ Is\ Passed}|\mathbf{x}] &\approx \frac{1}{N}\sum_{j=1}^{N}\left[\sum_{y=0}^{c}\frac{(\lambda^{(j)}T)^y \exp(-\lambda^{(j)}T)}{y!}\right] \\
&\approx \frac{1}{N}\sum_{j=1}^{N}\sum_{y=0}^{c} g^{(j)}(y)\Gamma(y+\eta^{(j)}) \\
&\approx \frac{1}{N}\sum_{j=1}^{N}\left[\sum_{y=0}^{c}\frac{(\kappa^{(j)})^{\eta^{(j)}} T^y \Gamma(y+\eta^{(j)})}{y!\Gamma(\eta^{(j)})(T+\kappa^{(j)})^{y+\eta^{(j)}}}\right].
\end{aligned}$$

To obtain a test plan, simultaneously solve the following pair of nonlinear inequalities:

$$\mathbf{P}[\lambda \leq \lambda_0 \,|\, \mathit{Test\ Is\ Failed}, \mathbf{x}] \leq \alpha \tag{10.18}$$

and

$$\mathbf{P}[\lambda \geq \lambda_1 \,|\, \mathit{Test\ Is\ Passed}, \mathbf{x}] \leq \beta, \tag{10.19}$$

where $\lambda_0 \leq \lambda_1$.

Because T is continuous, we can hold either of the risks in Eqs. 10.18 and 10.19 at its precise value. Holding the posterior producer's risk at exactly α, use the algorithm in Fig. 10.2 to obtain the desired test plan. On the other hand, holding the posterior consumer's risk at precisely β, simply reverse the two main steps in the procedure.

Example 10.4 Hierarchical Poisson test plan for pumps. Gaver and O'Muircheartaigh (1987) provides the data shown in Table 10.3 on the number of pump failures x observed in t thousands of operating hours for $m = 10$ different systems at the Farley 1 U.S. commercial nuclear power plant.

Note that we have listed the data in increasing order of the corresponding maximum likelihood estimates (MLEs) $\widehat{\lambda}$.

We model the failures as conditionally independent given their individual failure rates λ_i with $\mathit{Poisson}(\lambda_i t_i)$ distributions. Given η and κ, we model the λ_i as i.i.d. $\mathit{Gamma}(\eta, \kappa)$ and use independent and diffuse

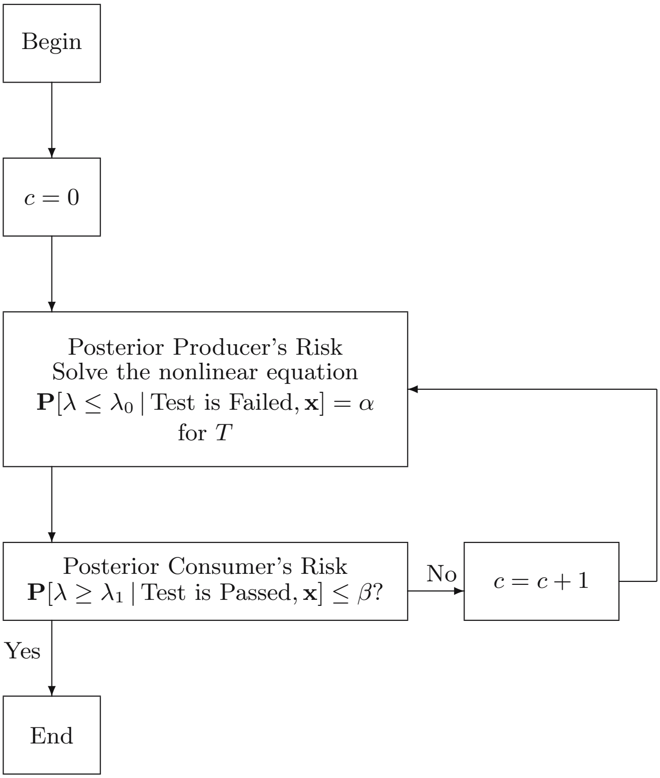

Fig. 10.2. An algorithm for finding Bayesian Poisson test plans.

Table 10.3. Pump failure count data from Farley 1 U.S. nuclear power plant (number of failures x in t thousands of operating hours) (Gaver and O'Muircheartaigh, 1987)

System	x_i (failures)	t_i (thousand hours)	$\widehat{\lambda}$ (MLE)
1	5	94.320	5.3 x 10^{-2}
2	1	15.720	6.4 x 10^{-2}
3	5	62.880	8.0 x 10^{-2}
4	14	125.760	11.1 x 10^{-2}
5	3	5.240	57.3 x 10^{-2}
6	19	31.440	60.4 x 10^{-2}
7	1	1.048	95.4 x 10^{-2}
8	1	1.048	95.4 x 10^{-2}
9	4	2.096	191.0 x 10^{-2}
10	22	10.480	209.9 x 10^{-2}

$InverseGamma(0.001, 0.001)$ prior distributions for η and κ. Table 10.4 summarizes the marginal posterior distributions of η and κ as well as the predictive distribution of λ. The hierarchical Poisson model fits the pump failure count data well (see Exercise 4.17).

Table 10.4. Posterior distribution summaries for the gamma distribution hyperparameters (η, κ) given $\mathbf{x}$ and of the predictive distribution for λ for pump example

			Quantiles				
Parameter	Mean	Std Dev	0.025	0.05	0.50	0.95	0.975
η	0.7981	0.3635	0.2995	0.3419	0.7272	1.4840	1.6950
κ	1.284	0.855	0.229	0.306	1.087	2.971	3.460
λ	0.796	1.306	0.002	0.007	0.390	2.939	4.037

Using the posterior risk criteria in Eqs. 10.18 and 10.19, suppose that we want to find the Poisson test plan with risk parameters $\lambda_0 = 0.2, \alpha = 0.05, \lambda_1 = 0.7$, and $\beta = 0.05$. Using the algorithm in Fig. 10.2 with $N = 10,000$ joint posterior draws of (η, κ) given $\mathbf{x}$, we find the required test plan $T_0 = 5.43$ and $c = 1$. The actual risks for this test plan are $\mathbf{P}[\lambda \geq 0.7 \,|\, Test\ Is\ Passed, \mathbf{x}] = 0.0171$ and $\mathbf{P}[\lambda \leq 0.2 \,|\, Test\ Is\ Failed, \mathbf{x}] = 0.0500$, and the unconditional probability of passing this test is approximately 0.44. To implement this test plan for new systems, accumulate 5,430 hours of pump operating time with repair (or replacement) of the failed pumps. If no more than one failure occurs, then the test is passed, otherwise, the test is failed.

Although it is inappropriate for Example 10.4, in many cases we are free to accumulate the required total test time T by choosing any desired combination of n devices and time on test t_0 satisfying $T = nt_0$. To illustrate this tradeoff, Fig. 10.3 shows selected combinations of the number of test devices n and required test time t_0 satisfying $nt_0 = 5,430$ hours. For example, if $n = 5$, then test each of these pumps (with repair or replacement) for $t_0 = 1,000$ hours each.

10.4 Weibull Testing

The previous sections on binomial and Poisson testing considered attribute test data, which capture the survival/nonsurvival of each device on test. This section focuses on lifetime data, where testers record the actual failure times. We assume that the failure times t follow a $Weibull(\lambda, \beta)$ distribution with scale parameter λ and shape parameter β, with probability density function

$$f(t|\lambda, \beta) = \lambda\beta t^{\beta-1} \exp(-\lambda t^{\beta}), \;\; t > 0, \, \lambda > 0, \, \beta > 0. \tag{10.20}$$

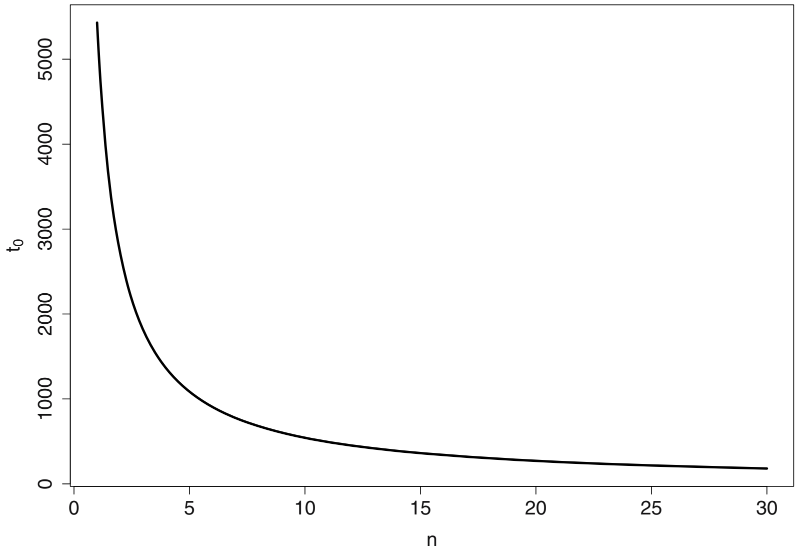

Fig. 10.3. The required test time t_0 in hours versus the number of test devices n for a Poisson test plan.

Suppose that we would like to use the posterior risk criteria to develop a Weibull test plan (n, t_0, c), where we put n units on test for t_0 time units and the test passes if no more than c units fail. To define the risk criteria, we specify requirements on reliability at time t_*, $R(t_*)$. Let

$$\begin{aligned} m(y) &= (1 - \exp[-\lambda t_0^\beta])^y \exp[-(n-y)\lambda t_0^\beta]\,, \\ k_0 &= -\log(\pi_0) t_*^{-\beta} \quad, \text{ and} \\ k_1 &= -\log(\pi_1) t_*^{-\beta}\,. \end{aligned}$$

For a Weibull test plan (n, t_0, c), we calculate the posterior producer's risk as

$$\begin{aligned} &\mathbf{P}(R(t_*) \geq \pi_0 \,|\, \textit{Test Is Failed}) \qquad (10.21) \\ &\quad = \mathbf{P}(\exp(-\lambda t_*^\beta) \geq \pi_0 \,|\, \textit{Test Is Failed}) \\ &\quad = \mathbf{P}(\lambda \leq -\log(\pi_0) t_*^{-\beta} \,|\, \textit{Test Is Failed}) \\ &\quad = \int_0^\infty \int_0^{k_0} f(\lambda, \beta \,|\, \textit{Test Is Failed}) \; \mathrm{d}\lambda \mathrm{d}\beta \\ &\quad = \int_0^\infty \int_0^{k_0} \frac{\mathbf{P}(\textit{Test Is Failed} \,|\, \lambda, \beta) p(\lambda, \beta)}{\int_0^\infty \int_0^\infty \mathbf{P}(\textit{Test Is Failed} \,|\, \lambda, \beta) p(\lambda, \beta) \; \mathrm{d}\lambda \mathrm{d}\beta} \; \mathrm{d}\lambda \mathrm{d}\beta \end{aligned}$$

$$= \frac{\int_0^\infty \int_0^{k_0} \left[1 - \sum_{y=0}^{c} m(y)\right] p(\lambda, \beta)\ \mathrm{d}\lambda \mathrm{d}\beta}{\int_0^\infty \int_0^\infty \left[1 - \sum_{y=0}^{c} m(y)\right] p(\lambda, \beta)\ \mathrm{d}\lambda \mathrm{d}\beta} \,.$$

We calculate the posterior consumer's risk as

$$\begin{aligned}
&\mathbf{P}(R(t_*) \leq \pi_1 \,|\, \textit{Test Is Passed}) \qquad (10.22)\\
&\quad = \mathbf{P}(\exp(-\lambda t_*^{\beta}) \leq \pi_1 \,|\, \textit{Test Is Passed})\\
&\quad = \mathbf{P}(\lambda \geq -\log(\pi_1) t_*^{-\beta} \,|\, \textit{Test Is Passed})\\
&\quad = \int_0^\infty \int_{k_1}^\infty f(\lambda, \beta \,|\, \textit{Test Is Passed})\ \mathrm{d}\lambda \mathrm{d}\beta\\
&\quad = \int_0^\infty \int_{k_1}^\infty \frac{\mathbf{P}(\textit{Test Is Passed} \,|\, \lambda, \beta) p(\lambda, \beta)}{\int_0^\infty \int_0^\infty \mathbf{P}(\textit{Test Is Passed} \,|\, \lambda, \beta) p(\lambda, \beta)\ \mathrm{d}\lambda \mathrm{d}\beta}\ \mathrm{d}\lambda \mathrm{d}\beta\\
&\quad = \frac{\int_0^\infty \int_{k_1}^\infty \left[\sum_{y=0}^{c} m(y))\right] p(\lambda, \beta)\ \mathrm{d}\lambda \mathrm{d}\beta}{\int_0^\infty \int_0^\infty \left[\sum_{y=0}^{c} m(y)\right] p(\lambda, \beta)\ \mathrm{d}\lambda \mathrm{d}\beta} \,.
\end{aligned}$$

10.4.1 Single Weibull Population Testing

One description of a failure time distribution is its reliable life. The *reliable life*, t_R, for specified R, is the time beyond which $100 \times R\%$ of the population will survive. In other words, t_R is the $(1 - R)$th quantile of the failure time distribution. For the Weibull distribution described by Eq. 10.20, $t_R = \lambda^{-1/\beta}[-\log(R)]^{1/\beta}$.

Among a variety of testing schemes, we focus here on a minimum sample size (or zero-failure) test plan $(n, t_0, c = 0)$. For a zero-failure test plan, test n devices each for a length of time t_0, and the test is passed if we observe no failures. To use such a test plan, we must determine appropriate values for both n and t_0. Meeker and Escobar (1998) considers such classical test plans in situations where the Weibull shape parameter β is known. We relax this restriction here by considering plans within a Bayesian hierarchical framework. In turn, there are two such cases to study: (1) an assurance test plan based on available data from a single Weibull population, and (2) an assurance test plan based on available data from a Weibull accelerated life test program.

Consider developing a test criterion to assure that $t_R > t_{R^*}$. For example, a manufacturer may want to assure that 99% of a certain expensive electronic product will survive a one-year warranty period; in this case, $R = 0.99$ and $t_{R^*} = 8,760$ hours. We use the posterior risk criterion $P[t_R > t_{R^*} \,|\, \textit{Test Is Passed}] \geq 1 - \alpha$. If the test is passed, we would like a high probability $(1 - \alpha)$ that $t_R > t_{R^*}$ — a high probability that the 0.99 quantile of the lifetime of the electronic products is greater than 8,760 hours. This leads to the expression

$$\mathbf{P}(t_R > t_{R^*} \,|\, \textit{Test Is Passed}) \qquad (10.23)$$

$$
\begin{aligned}
&= \mathbf{P}(\lambda^{-1/\beta}[-\log(R)]^{1/\beta} > t_{R^*} \,|\, \textit{Test Is Passed}) \\
&= \mathbf{P}(\lambda < -\log(R) t_{R^*}^{-\beta} \,|\, \textit{Test Is Passed}) \\
&= \int_0^\infty \int_0^{-\log(R)t_{R^*}^{-\beta}} f(\lambda, \beta \,|\, \textit{Test Is Passed}) \, \mathrm{d}\lambda \mathrm{d}\beta \\
&= \int_0^\infty \int_0^{-\log(R)t_{R^*}^{-\beta}} \frac{\mathbf{P}(\textit{Test Is Passed} \,|\, \lambda, \beta) p(\lambda, \beta)}{\int_0^\infty \int_0^\infty \mathbf{P}(\textit{Test Is Passed} \,|\, \lambda, \beta) p(\lambda, \beta) \mathrm{d}\lambda \mathrm{d}\beta} \, \mathrm{d}\lambda \mathrm{d}\beta \\
&= \frac{\int_0^\infty \int_0^{-\log(R)t_{R^*}^{-\beta}} \exp(-n\lambda t_0^\beta) p(\lambda, \beta) \mathrm{d}\lambda \mathrm{d}\beta}{\int_0^\infty \int_0^\infty \exp(-n\lambda t_0^\beta) p(\lambda, \beta) \mathrm{d}\lambda \mathrm{d}\beta} \geq 1 - \alpha \, .
\end{aligned}
$$

Notice the similarities of this risk formulation to the posterior consumer's risk. As with the Poisson test plan, for a fixed n, we can solve for t_0 to meet the desired risk criteria. With the chosen $c = 0$, this also specifies the level of posterior producer's risk.

Suppose now that we have failure time data from $m > 1$ situations. Note that some of the available failure time data may be censored. Let $t_{ij}, i = 1, \ldots, m, j = 1, \ldots, n_i$, denote the observed failure or censoring time for the jth device in the ith situation, and let $\mathbf{t}$ denote all the observed failure time data. Given λ_i and β, model the T_{ij} as conditionally independent with $T_{ij} \sim \textit{Weibull}(\lambda_i, \beta)$. We assume a common shape parameter, because in practice, the failure times of similar devices often (but not always) exhibit the same general Weibull shape because they share common intrinsic failure mechanisms. To complete the model, let us use the following prior distributions:

$$
\begin{aligned}
\lambda_i &\sim \textit{Gamma}(\eta, \kappa), \quad i = 1, \ldots, m, \\
(\eta, \kappa) &\sim p(\eta, \kappa), \text{ and} \\
\beta &\sim p(\beta),
\end{aligned}
$$

with known hyperparameters for $p(\eta, \kappa)$ and $p(\beta)$.

We want to develop a zero-failure test plan for a new situation where we assume the failure time data will be distributed $\textit{Weibull}(\lambda, \beta)$, with $\lambda \sim \Gamma(\eta, \kappa)$. Conditioning on the observed data $\mathbf{t}$ in Eq. 10.23,

$$
\begin{aligned}
&\mathbf{P}(t_R > t_{R^*} \,|\, \textit{Test Is Passed}, \mathbf{t}) \\
&\quad = \frac{\int_0^\infty \int_0^{-\log(R)t_{R^*}^{-\beta}} \exp(-n\lambda t_0^\beta) p(\lambda, \beta \,|\, \mathbf{t}) \mathrm{d}\lambda \mathrm{d}\beta}{\int_0^\infty \int_0^\infty \exp(-n\lambda t_0^\beta) p(\lambda, \beta \,|\, \mathbf{t}) \mathrm{d}\lambda \mathrm{d}\beta} \, .
\end{aligned} \tag{10.24}
$$

Assuming that we have $j = 1, \ldots, N$ MCMC draws from the posterior distributions (given $\mathbf{t}$) of η, κ, and β and N draws from the predictive distribution of $\lambda, \lambda^{(j)} \sim \Gamma(\eta^{(j)}, \kappa^{(j)})$, we can calculate our criterion as follows:

$$
\mathbf{P}(t_R > t_{R^*} \,|\, \textit{Test Is Passed}, \mathbf{t}) \tag{10.25}
$$

$$
\begin{aligned}
&= \frac{\int_0^\infty \int_0^{-\log(R)t_{R^*}^{-\beta}} \exp(-n\lambda t_0^\beta) p(\lambda, \beta \mid \mathbf{t}) \mathrm{d}\lambda \mathrm{d}\beta}{\int_0^\infty \int_0^\infty \exp(-n\lambda t_0^\beta) p(\lambda, \beta \mid \mathbf{t}) \mathrm{d}\lambda \mathrm{d}\beta} \\
&\approx \frac{\sum_{j=1}^N \exp(-n\lambda^{(j)} t_0^{\beta^{(j)}}) I[\lambda^{(j)} \le -\log(R) t_{R^*}^{-\beta^{(j)}}]}{\sum_{j=1}^N \exp(-n\lambda^{(j)} t_0^{\beta^{(j)}})} \\
&\approx \frac{\sum_{j=1}^N \frac{(\kappa^{(j)})^{\eta^{(j)}} \gamma[\eta^{(j)}, [-\log(R)](\kappa^{(j)} + n t_0^{\beta^{(j)}})/t_{R*}^{\beta^{(j)}}]}{(\kappa^{(j)} + n t_0^{\beta^{(j)}})^{\eta^{(j)}} \Gamma(\eta^{(j)})}}{\sum_{j=1}^N \frac{(\kappa^{(j)})^{\eta^{(j)}}}{(\kappa^{(j)} + n t_0^{\beta^{(j)}})^{\eta^{(j)}}}},
\end{aligned}
$$

where $\gamma(q, z)$ denotes the lower incomplete gamma function.

Note that we may base the choice of number of test devices n on other considerations, such as cost. It may also be interesting and useful to see how t_0 functionally depends on n, which we can examine by varying n over an appropriate range, solving for t_0, and plotting the results.

Example 10.5 Hierarchical Weibull test plan for pressure vessels. Gerstle and Kunz (1983) provides the failure times (in hours) for pressure vessels that were wrapped in Kevlar-49 fibers and subsequently tested at four different stresses: 23.4, 25.5, 27.6, and 29.7 megapascals (MPa). Crowder et al. (1991) analyzes these data assuming a constant Weibull shape parameter. The fibers came from eight different spools (numbered 1–8) of material, and both studies conclude that there is a significant spool effect. In this example, consider only the failure time data obtained at 23.4 MPa. See Table 10.5, which displays the failure time data at all the stresses; an asterisk indicates a time- or Type I-censored observation.

Because any difference in the reliability of the spools is primarily due to uncontrollable random manufacturing process variability (or noise), let us model the pressure vessel failure times corresponding to each spool as conditionally independent with $Weibull(\lambda_i, \beta)$ distributions. Given η and κ, the λ_i are i.i.d. $Gamma(\eta, \kappa)$. In our analysis of these data, we use independent and diffuse $InverseGamma(0.01, 0.01)$ prior distributions for η and κ. Also, we use an independent $Exponential(1.0)$ prior distribution for β; the motivation for this prior distribution is the analysis results of Crowder et al. (1991), which suggests values of β near 1.0. See Table 10.6, which summarizes the marginal posterior distributions for η, κ, and β given $\mathbf{t}$. The hierarchical Weibull model fits these data well (see Exercise 4.20). Table 10.6 also summarizes the predictive distribution of λ.

Now suppose that we want to find a Bayesian minimum sample size test plan at a stress of 23.4 MPa for $t_{R^*} = 2,000$ hours, $R = 0.9$, and $\alpha = 0.05$. By letting $n = 1, 2, \ldots, 30$ and solving Eq. 10.25 for the corresponding test length t_0, we obtain the graph shown in Fig. 10.4.

For example, suppose that we decide to test $n = 10$ pressure vessels all wrapped from a particular spool of Kevlar-49 fibers. Figure 10.4 indicates

Table 10.5. Failure times of Kevlar-49-wrapped pressure vessels at four stress levels (An asterisk indicates a time-censored or Type I-censored observation) (Gerstle and Kunz, 1983)

Stress (MPa)	Spool	Failure Time (hours)
29.7	1	444.4 755.2 952.2 1108.2
29.7	2	2.2 8.5 9.1 10.2 22.1 55.4 111.4 158.7
29.7	3	12.5 14.6 18.7 101.0
29.7	4	254.1 1148.5 1569.3 1750.6 1802.1
29.7	5	8.3 13.3 87.5 243.9
29.7	6	6.7 15.0 144.0
29.7	7	4.0 4.0 4.6 6.1 7.9 14.0 45.9 61.2
29.7	8	98.2 590.4 638.2
27.6	1	453.4 664.5 930.4 1755.5
27.6	2	71.2 199.1 403.7 432.2 514.1 544.9 694.1
27.6	3	19.1 24.3 69.8 136.0
27.6	4	876.7 1275.6 1536.8 6177.5
27.6	5	
27.6	6	514.2 541.6 1254.9
27.6	7	
27.6	8	554.2 2046.2
25.5	1	11487.3 14032.0 31008.0
25.5	2	1134.2 1824.3 1920.1 2383.0 3708.9 5556.0
25.5	3	1087.7 2442.5
25.5	4	13501.3 29808.0
25.5	5	11727.1
25.5	6	225.2 6271.1 7996.0
25.5	7	503.6
25.5	8	2974.6 4908.9 7332.0 7918.7 9240.3 9973.0
23.4	1	41000* 41000* 41000* 41000*
23.4	2	14400.0
23.4	3	8616.0
23.4	4	41000* 41000* 41000* 41000*
23.4	5	9120.0 20231.0 35880.0
23.4	6	7320.0 16104.0 20233.0
23.4	7	4000.0 5376.0
23.4	8	41000* 41000* 41000*

Table 10.6. Posterior distribution summaries for η, κ, and β given $\mathbf{t}$ and of the predictive distribution for λ

			Quantiles		
Parameter	Mean	Std Dev	0.025	0.50	0.975
β	2.255	0.643	1.128	2.211	3.773
η	0.2100	0.1730	0.0529	0.1666	0.6113
κ	2.313E+13	3.105E+14	7.346E+3	1.121E+8	6.463E+13
λ	5.630E−6	1.221E−4	7.304E−25	3.832E−11	1.273E−5

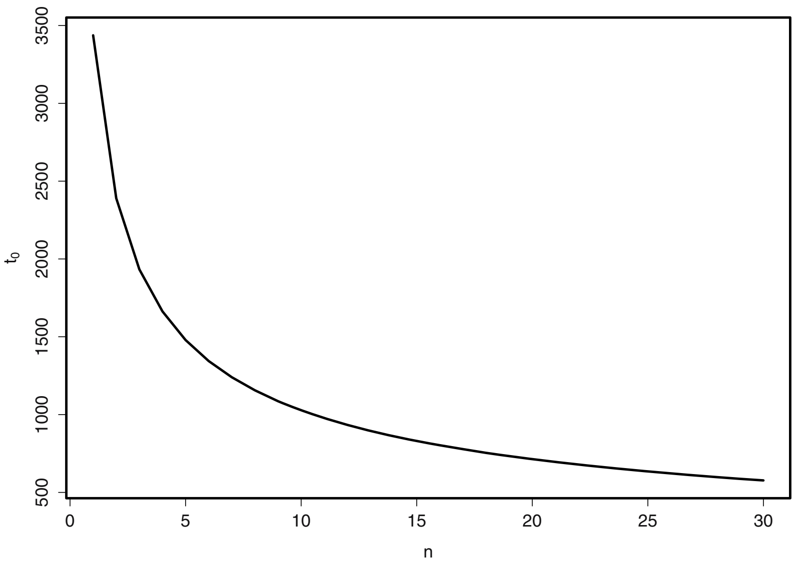

Fig. 10.4. The required Weibull test duration t_0 versus the number of test devices n for the pressure vessels example.

required testing of each of these pressure vessels for approximately $t_0 = 1,122$ hours at a stress of 23.4 MPa. If none of these fail, we can then claim, with 0.95 probability, that at least 90% of the pressure vessels wrapped from this spool will survive 2,000 hours at this stress.

10.4.2 Combined Weibull Accelerated/Assurance Testing

Now consider a Bayesian test plan based on data from a Weibull accelerated test program. Specifically, we consider again the failure time data in Table 10.5. Let $t_{ij}, i = 1, \ldots, m, j = 1, \ldots, n_i$, denote either the observed failure or censoring time for the jth device from the ith situation, with $\mathbf{t}$ denoting all of the observed failure time data. Let s_{ij} be the value of the stress s under which we obtained t_{ij}.

Conditional on λ_{ij} and β, we model the T_{ij} conditionally independent with a $Weibull(\lambda_{ij}, \beta)$ distribution. Let us define a model for λ_{ij} as

$$\lambda_{ij} = \exp(\gamma_0) s_{ij}^{\gamma_1} \omega_i \, , \tag{10.26}$$

where $\omega_i > 0$ is the random effect associated with the ith spool, and γ_0 and γ_1 are two regression parameters used to model the relationship between λ_{ij}

and s_{ij}. Recall that we presented similar regression models in Chap. 7 (see also León et al. (2007)).

We specify prior distributions $\omega_i \mid \eta, \kappa \sim Gamma(\eta, \kappa)$, $(\eta, \kappa) \sim p(\eta, \kappa)$, $\beta \sim p(\beta)$, $\gamma_0 \sim p(\gamma_0)$, $\gamma_1 \sim p(\gamma_1)$ with known hyperparameters for $p(\eta, \kappa)$, $p(\beta)$, $p(\gamma_0)$, and $p(\gamma_1)$.

To develop the Bayesian zero-failure test for a new spool, we assume the plan consists of testing n samples for time t_0 at stress s_0. We use MCMC to obtain posterior draws $\eta^{(j)}, \kappa^{(j)}, \beta^{(j)}, \gamma_0^{(j)}, \gamma_1^{(j)}$ given $\mathbf{t}$ and predictive draws $\omega^{(j)} \sim \Gamma(\eta^{(j)}, \kappa^{(j)})$. Let

$$
\begin{aligned}
\boldsymbol{\theta} &= (\omega, \gamma_0, \gamma_1, \beta) \quad \text{and} \\
q^{(j)} &= \kappa^{(j)} + n \exp(\gamma_0^{(j)}) s_0^{\gamma_1^{(j)}} t_0^{\beta^{(j)}} .
\end{aligned}
$$

Our test criterion is calculated as

$$
\begin{aligned}
&\mathbf{P}(t_R > t_{R^*} \mid Test\ Is\ Passed, \mathbf{t}) \qquad (10.27) \\
&= \frac{\int_0^\infty \int_0^\infty \int_0^\infty \int_0^{\frac{-\log(R) t_{R^*}^{-\beta}}{\exp(\gamma_0) s_0^{\gamma_1}}} \exp(-n \exp(\gamma_0) s_0^{\gamma_1} \omega t_0^{\beta}) p(\boldsymbol{\theta} \mid \mathbf{t}) \mathrm{d}\boldsymbol{\theta}}{\int_0^\infty \int_0^\infty \int_0^\infty \int_0^\infty \exp(-n \exp(\gamma_0) s_0^{\gamma_1} \omega t_0^{\beta}) p(\boldsymbol{\theta} \mid \mathbf{t}) \mathrm{d}\boldsymbol{\theta}} \\
&\approx \frac{\sum_{j=1}^N \exp(-n \exp(\gamma_0^{(j)}) s_0^{\gamma_1^{(j)}} \omega^{(j)} t_0^{\beta^{(j)}}) I \left[\omega^{(j)} \leq \frac{-\log(R) t_{R^*}^{-\beta^{(j)}}}{\exp(\gamma_0^{(j)}) s_0^{\gamma_1^{(j)}}} \right]}{\sum_{j=1}^N \exp(-n \exp(\gamma_0^{(j)}) s_0^{\gamma_1^{(j)}} \omega^{(j)} t_0^{\beta^{(j)}})} \\
&\approx \frac{\sum_{j=1}^N \frac{(\kappa^{(j)})^{\eta^{(j)}} \gamma \left[\eta^{(j)}, [-\log(R)] \exp(-\gamma_0^{(j)}) s_0^{-\gamma_1^{(j)}} q^{(j)} / t_{R^*}^{\beta^{(j)}} \right]}{\Gamma(\eta^{(j)}) (q^{(j)})^{\eta^{(j)}}}}{\sum_{j=1}^N \frac{(\kappa^{(j)})^{\eta^{(j)}}}{(q^{(j)})^{\eta^{(j)}}}} ,
\end{aligned}
$$

where $\gamma(q, z)$ denotes the lower incomplete gamma function.

More specifically, to analyze the data in Table 10.5, we use the following independent prior distributions:

$$
\begin{aligned}
\beta &\sim Exponential(1.0), \\
\eta &\sim InverseGamma(0.01, 0.01), \\
\kappa &\sim InverseGamma(0.01, 0.01), \\
\gamma_0 &\sim Normal(0, 10^6), \text{ and} \\
\gamma_1 &\sim Normal(0, 10^6).
\end{aligned}
$$

See Table 10.7, which summarizes the marginal posterior distributions for these five parameters. The hierarchical Weibull regression model fits these data well (see Exercise 7.23).

Table 10.7. Posterior distribution summaries for $\eta, \kappa, \beta, \gamma_0$, and γ_1 given $\mathbf{t}$ and predictive distribution for ω for pressure vessels example

Parameter	Mean	Std Dev	Quantiles 0.025	0.50	0.975
β	1.199	0.085	1.038	1.199	1.3650
η	0.6522	0.3068	0.2198	0.5977	1.391
γ_0	−84.81	11.31	−105.50	−83.40	−66.32
γ_1	27.84	1.81	24.56	27.85	31.31
κ	3.704E+15	5.666E+16	2.403E−2	4.873E+6	7.071E+15
ω	2.512	19.42	1.34E-17	3.900E-08	19.73

We can see the effect that stress has on failure time by computing the posterior distribution of the Weibull mean time to failure (MTTF) as a function of s for a randomly selected spool of Kevlar 49. Recall that the MTTF of the Weibull distribution in Eq. 10.26 is $\lambda^{-1/\beta}\Gamma(1+1/\beta)$, which, substituting using Eq. 10.26, becomes $\exp(-\gamma_0/\beta)s^{-\gamma_1/\beta}\omega^{-1/\beta}\Gamma(1+1/\beta)$.

In Fig. 10.5, we plot the median log MTTF and a 90% central credible interval as a function of stress s. Note the decreasing trend in the Weibull log MTTF as s increases, as well as the extremely heavy right tail of the posterior log MTTF distribution for a given value of s.

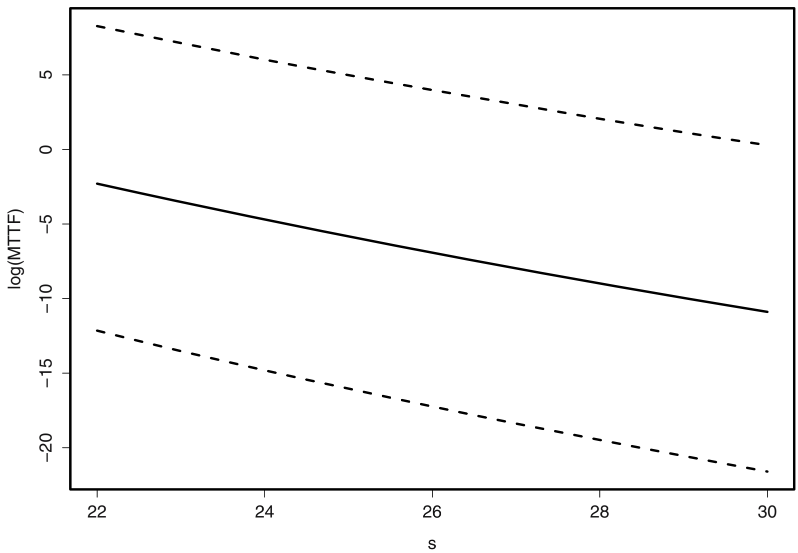

Fig. 10.5. The posterior median and 90% credible interval of the Weibull log MTTF as a function of stress s for pressure vessels example.

We can also see the effect of stress on pressure vessel reliability by computing the posterior distribution of reliability as a function of s for a randomly selected Kevlar-49 spool. Substituting using Eq. 10.26, the reliability at $t = 2,000$ hours is $\exp[-\exp(\gamma_0)s^{\gamma_1}\omega(2,000)^{\beta}]$. In Fig. 10.6, we plot the posterior median and 90% credible interval for pressure vessel reliability at $2,000$ hours as a function of stress s. Note the significant increase in the width of the 90% credible intervals as the stress s increases.

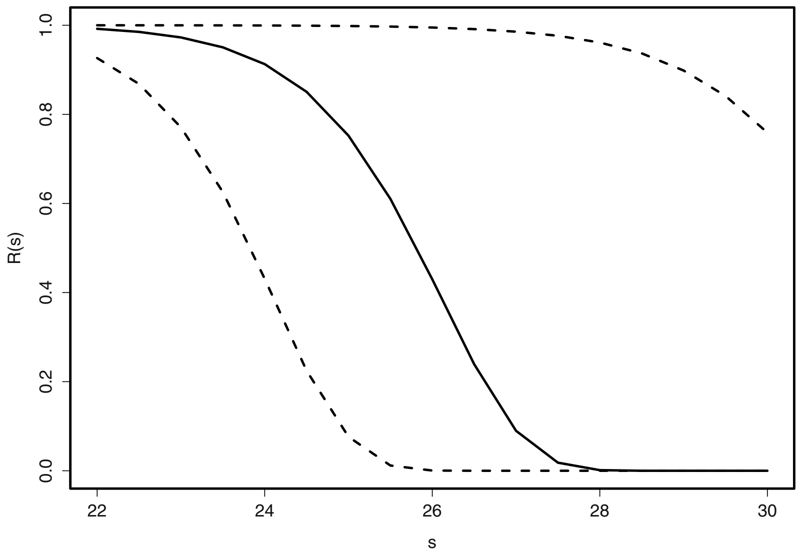

Fig. 10.6. The posterior median and 90% credible interval for $R(2000)$ as a function of stress s for pressure vessels example.

Now suppose that we want to find a Bayesian zero-failure test plan for $t_{R^*} = 2,000$ hours, $R = 0.9, \alpha = 0.05$, and $s_0 = 23$ MPa. By letting $n = 1, 2, \ldots, 15$ and solving Eq. 10.27 for the corresponding test length t_0, we obtain the solid curve in Fig. 10.7. For example, in testing $n = 10$ pressure vessels all wrapped from a spool of Kevlar-49 fiber, Fig. 10.7 indicates testing each of these vessels for approximately $t_0 = 1,900$ hours. Note that in Example 10.5, a shorter time of $t_0 = 1,132$ was required because the reliability at 2,000 hours was assessed to be higher based on the only 23 MPa data. If there are no failures, then we can claim with 0.95 probability that at least 90% of the pressure vessels wrapped from this spool will survive 2,000 hours at a stress of 23 MPa. Figure 10.7 also presents the test length t_0 as a function of n for stresses $s_0 = 24$ MPa and $s_0 = 27$ MPa.

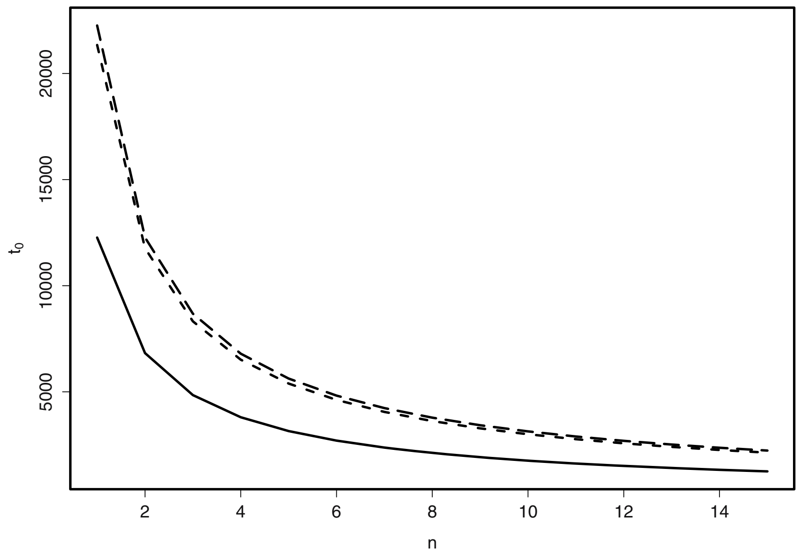

Fig. 10.7. The required test duration t_0 versus the number of test devices n for three different values of stress s_0 in MPa for pressure vessels example. The three stress values are 23 MPa (solid line), 24 MPa (short dashed line), and 27 MPa (long dashed line).

From Fig. 10.7, we can make an important observation; namely, for a given number of devices n, as the stress s_0 decreases, the required test time t_0 also decreases. This desirable situation arises because the failure time data model using the original accelerated life test data predicts that the probability of surviving the required 2,000 hours of operation increases dramatically as the stress decreases. In other words, because the fitted model predicts long failure times at low stress, achieving the required assurance needs very little additional data. For those cases, in which there is high reliability at low stress, the required assurance is already embedded in the fitted model. On the other hand, if the accelerated life test results indicate low reliability at low stress, then significant assurance testing would be required to overcome this situation.

10.5 Related Reading

There is extensive literature on Bayesian assurance testing dating back to the late 1960s, and Martz and Waller (1982) describes this early research. Brush (1986) clarifies the distinction between a posterior Bayes' and a modified classical (or average) producer's risk and compares these two criteria for various test plans. Brush (1986) also highlights the importance of calculating a Bayes' producer's risk as a supplement to the modified classical producer's risk. Sharma and Bhutani (1992) analyzes the performance of Bayes' and classical assurance test plans for simultaneously specified consumer's and producer's risks. Both Brush (1986) and Sharma and Bhutani (1992) conclude that Bayes' and classical risks need consideration.

Since 1990 there have been several more articles concerned with either Bayesian acceptance sampling or assurance test plans. Hart (1990), with comments by Ganter et al. (1990), uses Bayesian methods to determine a test plan for qualifying the reliability of some industrial products, whereas Guess and Usher (1990) considers a Bayesian approach to the assurance testing of highly reliable devices. Fan (1991) proposes a Bayesian acceptance test plan for binomial testing, while Sheng and Fan (1992) presents a method for choosing a prior distribution for binomial testing. In a Master's thesis, Jin (1991) develops Bayesian acceptance test plans based on failure-free life tests and compares these with other test plans. Pham and Turkkan (1992) considers a four-parameter beta prior in binomial testing. In a somewhat different approach, Moskowitz and Tang (1992) uses both quadratic and step-function loss functions in determining Bayesian acceptance test plans. Berger and Sun (1993) develops Bayesian sequential reliability demonstration tests using two different approaches: posterior loss and predictive loss. In addition, Berger and Sun (1993) considers three test data models. Whitmore et al. (1994) proposes two different approaches for integrating life test data into a Bayesian analysis based on the exponential distribution. For the exponential distribution, Deely and Keats (1994) develops Bayesian stopping rules for use in terminating a sequential assurance test. Vintr (1999) considers the optimization of reliability requirements from a manufacturer's point of view. Tobias and Poore (2003) proposes Bayesian reliability testing for a new generation of semiconductor processing equipment, while Kleyner et al. (2004) develops reliability demonstration test plans based on minimizing life cycle costs.

10.6 Exercises for Chapter 10

10.1 In determining a Bayesian reliability assurance test plan, what happens if the prior distribution is especially strong and satisfies the desired criteria prior to testing?
   a) How does one know when this is the case?
   b) What calculations should one perform to check whether or not this is the case?

10.2 a) What is the (classical) producer's risk for a binomial test plan with $n = 15, c = 1$, and $\pi_0 = 0.9$?
   b) What is the (classical) consumer's risk for a binomial test plan with $n = 15, c = 1$, and $\pi_1 = 0.6$?

10.3 a) What is the average producer's risk for a binomial test plan with $n = 15, c = 1, \pi_0 = 0.9$, and $p(\pi) \sim Beta(10, 1)$?
b) What is the average consumer's risk for a binomial test plan with $n = 15, c = 1, \pi_1 = 0.6$, and $p(\pi) \sim Beta(10, 1)$?

10.4 a) What is the posterior producer's risk for a binomial test plan with $n = 15, c = 1, \pi_0 = 0.9$, and $p(\pi) \sim Beta(10, 1)$?
b) What is the posterior consumer's risk for a binomial test plan with $n = 15, c = 1, \pi_1 = 0.6$, and $p(\pi) \sim Beta(10, 1)$?

10.5 Discuss the similarities and differences between the producer's and consumer's risks calculated in Exercises 10.2, 10.3, and 10.4.

10.6 Calculate a binomial test plan with a $Uniform(0, 1)$ prior distribution for π, and $\pi_0 = 0.9, \pi_1 = 0.5, \alpha = \beta = 0.05$.

10.7 The auxiliary feedwater (AFW) system is an important standby safety system in a nuclear power plant (Poloski et al., 1998). The AFW system probability of starting on demand is an important indicator of its reliability. The data in Table 10.8 are the number of AFW system failures to start on demand x_i in n_i demands at 68 U.S. commercial nuclear power plants.
a) Find the Bayesian hierarchical test plan having the posterior consumer's and producer's risk values $\pi_1 = 0.985, \beta = 0.05, \pi_0 = 0.995$, and $\alpha = 0.10$.
b) What are the actual posterior risks when using this test plan?
c) What is the unconditional probability of passing the test when using this test plan?
d) Is there anything unusual about this problem?

10.8 Derive Eq. 10.15.

10.9 Borg (1962) provides data that apparently originated at the U.S. Bureau of Naval Weapons regarding the number of minor, major, and critical defectives in successive MIL-STD-105B samples of some material. The data in Table 10.9 consist of the observed frequencies of the number of minor defectives x in samples of size $n = 150$ from $m = 205$ lots each of size 2016 items of this material.
a) Using the "hybrid" posterior consumer's and average producer's risk criteria, find the Bayesian hierarchical test plan having the risk values $\beta = 0.10$ and $\alpha = 0.05$ for $\pi^* = 0.975$.
b) What are the actual risks when using this test plan?
c) What is the unconditional probability of passing this test?

10.10 Using the data in Exercise 10.9, find the Bayesian hierarchical test plan having the posterior consumer's and producer's risk values $\pi_1 = 0.96, \beta = 0.10, \pi_0 = 0.975$, and $\alpha = 0.05$.
a) What are the actual risks when using this test plan?
b) What is the unconditional probability of passing this test?

10.11 a) What is the posterior producer's risk for a Poisson test plan with $T = 10, c = 3, \lambda_0 = 3.0$, and $p(\lambda) \sim Gamma(5, 1)$?

Table 10.8. Number of AFW system failures to start on demand x in n demands at 68 U.S. commercial nuclear power plants (Poloski et al., 1998)

Plant	x	n	Plant	x	n
Arkansas 1	0	14	North Anna 2	0	18
Arkansas 2	0	9	Oconee 1	0	18
Beaver Valley 1	0	24	Oconee 2	0	18
Beaver Valley 2	0	43	Oconee 3	0	12
Braidwood 1	0	13	Palisades	0	13
Braidwood 2	0	24	Palo Verde 1	0	7
Byron 1	0	11	Palo Verde 2	0	12
Byron 2	0	26	Palo Verde 3	0	9
Callaway	0	57	Point Beach 1	0	8
Calvert Cliffs 1	0	12	Point Beach 2	0	16
Calvert Cliffs 2	0	15	Prairie Island 1	0	3
Catawba 1	0	41	Prairie Island 2	0	7
Catawba 2	0	89	Robinson 2	1	28
Comanche Pk 1	0	66	Salem 1	0	24
Comanche Pk 2	0	14	Salem 2	0	32
Cook 1	0	18	San Onofre 2	0	13
Cook 2	0	36	San Onofre 3	0	17
Crystal River 3	1	16	Seabrook	0	17
Diablo Canyon 1	0	46	Sequoyah 1	0	30
Diablo Canyon 2	0	30	Sequoyah 2	0	41
Farley 1	0	34	South Texas 1	0	69
Farley 2	0	54	South Texas 2	0	87
Fort Calhoun	0	5	St. Lucie 1	0	35
Ginna	0	28	St. Lucie 2	0	21
Harris	0	98	Summer	0	24
Indian Point 2	1	24	Surry 1	0	26
Indian Point 3	2	32	Surry 2	0	32
Kewaunee	0	26	Three Mile Isl 1	0	6
Maine Yankee	0	23	Vogtle 1	0	103
McGuire 1	0	45	Vogtle 2	0	45
McGuire 2	0	44	Waterford 3	0	38
Millstone 2	1	11	Wolf Creek	0	51
Millstone 3	0	54	Zion 1	0	13
North Anna 1	0	20	Zion 2	0	8

b) What is the posterior consumer's risk for a Poisson test plan with $T = 10, c = 3.0, \lambda_1 = 7.0$, and $p(\lambda) \sim Gamma(5, 1)$?

10.12 For Poisson testing presented in Sect. 10.3, show that

$$\mathbf{P}(Test\ Is\ Failed|\lambda \leq \lambda_0, \eta, \kappa) = 1 - \frac{\kappa^\eta \sum_{x=0}^{c} \frac{T_0^x \gamma[x+\eta, (T_0+\kappa)\lambda_0]}{x!(T_0+\kappa)^{x+\eta}}}{\gamma(\eta, \kappa\lambda_0)} .$$

10.13 Using the expression given in Exercise 10.12 and the pump failure data in Table 10.3, find the Bayesian hierarchical test plan having the "hybrid"

Table 10.9. Minor defectives from MIL-STD-105B sampling of material (Borg, 1962)

x	Frequency	x	Frequency
0	68	9	2
1	45	10	1
2	24	12	1
3	20	13	1
4	8	15	4
5	7	18	1
6	8	20	1
7	10	22	1
8	3		

posterior consumer's and average producer's risk values $\lambda_1 = 0.3, \beta = 0.10, \lambda_0 = 0.2$, and $\alpha = 0.05$.
a) What are the actual risks for this test plan?
b) What is the unconditional probability of passing this test?

10.14 Using the pump failure data in Table 10.3, find the Bayesian hierarchical test plan having the posterior consumer's and producer's risk values $\lambda_1 = 0.1, \beta = 0.10, \lambda_0 = 0.05$, and $\alpha = 0.05$.
a) What are the actual risks for this test plan?
b) What is the unconditional probability of passing this test?
c) Is this a good test plan to use?

10.15 For the model in Sect. 10.4.1, show that we may approximate the unconditional probability of passing the Bayesian minimum sample size test plan by

$$\mathbf{P}[\mathit{Test\ Is\ Passed} \,|\, \mathbf{t}] \approx \frac{1}{N} \sum_{j=1}^{N} \left(\frac{\kappa^{(j)}}{\kappa^{(j)} + n t_0^{\beta^{(j)}}} \right)^{\eta^{(j)}} .$$

10.16 Using the expression in Exercise 10.15, what is the approximate unconditional probability of passing the Bayesian test plan ($n = 10, t_0 = 700$) given in Example 10.5?

10.17 Gerstle and Kunz (1983) gives the failure times for Kevlar-49-wrapped pressure vessels at a stress of 25.5 MPa. Table 10.10 displays these data. For $t_{R^*} = 300$ hours, $R = 0.9, \alpha = 0.05$, and these Weibull distributed data, find the Bayesian minimum sample size test plan time t_0 that we must test each of $n = 5$ items. What is the unconditional probability of passing this test? Is this a satisfactory test plan?

10.18 For the Weibull testing described in Sect. 10.4, suppose that we want to find a Bayesian minimum sample size test plan to assure that, at some specified time t_R, the Weibull reliability R is at least as large as a requirement R^* at the $100 \times (1 - \alpha)\%$ credible level. How does this test plan compare to the one based on the reliable life criterion?

Table 10.10. Failure times of Kevlar-49-wrapped pressure vessels at a stress of 25.5 MPa (Gerstle and Kunz, 1983)

Spool	Failure Time (hours)
1	11487.3, 14032.0, 31008.0
2	1134.3, 1824.3, 1920.1, 2383.0, 3708.9, 5556.0
3	1087.7, 2442.5
4	13501.3, 29808.0
5	11727.1
6	225.2, 6271.1, 7996.0
7	503.6
8	2974.6, 4908.9, 7332.0, 7918.7, 9240.3, 9973.0