19. Statistical Survival Analysis with Applications

This chapter discusses several important and interesting applications of statistical survival analysis which are relevant to both medical studies and reliability studies. Although it seems to be true that the proportional hazards models have been more extensively used in the application of biomedical research, the accelerated failure time models are much more popular in engineering and reliability research. Through several applications, this chapter not only offers some unified approaches to statistical survival analysis in biomedical research and reliability/engineering studies, but also sets up necessary connections between the statistical survival models used by biostatisticians and those used by statisticians working in engineering and reliability studies. The first application is the determination of sample size in a typical clinical trial when the mean or a certain percentile of the survival distribution is to be compared. The approach to the problem is based on an accelerated failure time model and therefore can have direct application in designing reliability studies to compare the reliability of two or more groups of differentially manufactured items. The other application we discuss in this chapter is the statistical analysis of reliability data collected from several variations of step-stress accelerated life test. The approach to the problem is based on the accelerated failure time model, but we will point out that these methodologies can be directly applied to medical and clinical studies when different doses of a therapeutic compound are administered in a sequential order to experimental subjects.

19.1 **Sample Size Determination to Compare Mean or Percentile of Two Lifetime Distributions** 349
19.1.1 The Model and Sample Size 350
19.1.2 Examples 351
19.1.3 Effect of Guarantee Time on Sample Size Determination 351
19.1.4 Application to NIA Aging Intervention Testing Program 354

19.2 **Analysis of Survival Data from Special Cases of Step-Stress Life Tests** 355
19.2.1 Analysis of Grouped and Censored Data from Step-Stress Life Tests..... 356
19.2.2 Analysis of a Very Simple Step-Stress Life Test with a Random Stress-Change Time 361

References 365

Failure time data or survival data are frequently encountered in biomedical studies, engineering, and reliability research. Applications of lifetime distribution methodologies range from investigations into the reliability of manufactured items to research involving human diseases. In medical studies, clinical endpoints for assessment of efficacy and safety of a promising therapy usually include occurrence of some predefined events such as deaths, the onset of a specific disease, the response to a new chemotherapy in treatment of some advanced cancer, the eradication of an infection caused by a certain microorganism, or serious adverse events. In engineering and reliability studies, manufactured items such as mechanical or electronic components are often subjected to life tests to obtain information on their endurance. This involves putting items into operation, often in a laboratory setting, and observing them until they fail. For example, *Nelson* [19.1] described a life test experiment in which specimens of a type of electronic insulating fluid were subjected to a constant voltage stress. The length of time until each specimen broke down was observed and investigated in its association with the voltage level. In all of the studies mentioned above, the primary variable of interest is usually the *survival time* to the occurrence of a specific predetermined event. One of the important features in survival data encountered

in both medical research and engineering studies is the existence of censored observations when only a lower (or upper) bound of the failure time on some experimental units are available. Censoring occurs frequently because of time limits and other restrictions during the process of data collection. In a life test experiment of manufactured items, for example, it may not be feasible to continue experimentation until all items under study have failed. If the study is terminated before all have failed, then for items that have not failed at the time of termination only a lower bound on lifetime is available.

The statistical analysis of survival data has been well developed in the literature. The estimation of the survival distribution can be done by the Kaplan–Meier product-limit estimator [19.2], which can also be viewed as a kind of nonparametric maximum likelihood estimator [19.3]. For studies in which the aim is to compare the survival distribution of two groups of subjects, the logrank test has been the most common method, although other rank tests such as the generalized Wilcoxon test are also used [19.2]. The logrank test can also be extended to allow an adjustment to be made for other covariates [19.4]. The major developments in the analysis of survival data have focused on several families of survival distributions. Two very important models of survival distribution are the model of proportional hazards and the accelerated failure time model. The proportional hazard model is a regression method introduced by *Cox* [19.5], which can be used to investigate the effects of several covariates on survival distribution at the same time. Cox's method is a semi-parametric approach – no particular type of distribution is assumed for the survival data, but a strong assumption is made on the effects of differences, which is referred to as the assumption of proportional hazards. Regression diagnostic procedures are also available to assess the assumption of proportional hazards [19.6,7], and some tests of the assumption of proportional hazards are also introduced through the incorporation of time-dependent covariates [19.8]. Extensions to Cox's proportional hazards model are the analysis of residuals, time-dependent coefficient, multiple/correlated observations, time-dependent strata, and estimation of underlying hazard function [19.8–10]. The accelerated failure time model, on the other hand, assumes that the covariates act by expanding or contracting time by a factor which is the exponential of a linear function of the covariates. In the logarithmic scale of the survival time, the accelerated failure time model is essentially a scale-location family of distributions.

It is quite interesting to observe that the proportional hazards models have been more extensively used in the application of biomedical research, while the accelerated failure time models are much more popular in engineering and reliability research. Part of the reason that Cox's proportional hazards models are popular in biomedical studies is the very fact that the assumption of proportional hazards summarizes the risk factor for a specific disease into a single quantity, the hazard ratio, which makes the interpretation easy to understand for clinicians. As an example, medical literature has demonstrated that a key protein, apolipoprotein E4 (ApoE4), contributes to the development of Alzheimer's disease [19.11]. Clinicians are interested in knowing how much the risk of Alzheimer's disease is increased for ApoE4-positive subjects compared to ApoE4-negative subjects. The point and confidence interval estimate to the hazard ratio associated with ApoE4 will adequately address the question if the assumption of proportional hazards can be adequately verified. On the other hand, the accelerated failure time models often make very good sense when the multiplicative time scale is assumed based on the level of covariate. As an example, assume that the lifetime of photocopiers has a hazard function that is a function of the total number of copies made, but the data on their failures were recorded in calendar time. Covariates that were related to the number of copies per day might be very successful in an accelerated failure time model. If the underlying hazard function had a particular form, say a sharp upturn after 25 000 cycles, a proportional hazard model would not fit as well. Similar examples can also be found for biological data related to cumulative toxicity or other damage.

Whether or not a statistical model is appropriate in a specific application depends on the distributional property of the observed variable and the specific research questions to be addressed. This chapter focuses on several applications of survival analysis in both medical/biological research and engineering/reliability research. We discuss several interesting applications which are relevant to both medical studies and reliability studies. The first application is the determination of sample size in a typical clinical trial when the mean or a certain percentile of the survival distribution is to be compared. The approach to the problem is based on an accelerated failure time model and therefore can have direct application in designing reliability studies to compare the reliability of two or more groups of differentially manufactured items. The other application we discuss in this chapter is the statistical

analysis of reliability data collected from several variations of step-stress accelerated life test. The approach to the problem is based on the accelerated failure time model, but we will point out that these methodologies can be directly applied to medical and clinical studies when different doses of a therapeutic compound are administered in a sequential order to experimental subjects.

19.1 Sample Size Determination to Compare Mean or Percentile of Two Lifetime Distributions

The determination of sample size is an important subject in planning long-term medical trials. In a randomized longitudinal clinical trial involving a treatment group and a control group, if the survival time to a particular event (e.g., death, relapse of symptoms) is the primary concern for the study, there are two important types of comparisons between the treatment group and the control group. One is the comparison of two survival curves and the other is the comparison of two common survival characteristics such as two means and two percentiles. Although the comparison of two survival curves is the major interest in many studies, the comparison of two means or two percentiles is important in many other applications. For example, in the announcement of the aging intervention testing program (RFA-AG-02-005), the National Institute on Aging (NIA) of USA states that one of the major research objectives of this program is to identify interventions that increase mean life expectancy by 10% in phase I studies, which may be terminated at 50% survivorship. This type of aging intervention study based on animal models has recently received much attention in the community of aging research. For example, caloric restriction has been identified as an intervention that extends the life span of both mammalian animal models and a variety of invertebrate animal models [19.12]. Mutations in the *dw* and *df* genes have been shown to attenuate the rate of aging in mice [19.13, 14]. *Warner* et al. [19.15] provided more details of biological interventions to promote healthy aging. The sample size computation for this type of study requires a statistical test that compares the mean lifetime between the control group and the treatment group based on type II censored observations.

Sample size determination methods are always based on certain parametric or semiparametric statistical models. This section concerns two important families of distributions used in the analysis of lifetime data: one is the family of proportional hazards and the other is the location-scale family of the log-transformed lifetime. The traditional approach to the sample size problem in planning long-term medical trials is based on the logrank test for the comparison of two lifetime distributions between the control group and the treatment group. Although the logrank test can be derived from both the proportional hazards family and the location-scale family of log-transformed lifetime, it is the proportional hazards family that most sample size computation methods with logrank test in the literature have been based on. In fact, the statistics literature on sample size calculation for failure time data is almost entirely devoted to tests based on exponential survival distributions [19.16–18] or binomial distributions [19.19, 20]. This is largely due to the fact that with the more general conditions hazard functions and ratios are no longer constant over time, so that the usual tests based on exponential models with constant hazard ratios no longer apply. *Schoenfeld* [19.21] and *Freedman* [19.22] presented methods for sample size calculation based on the asymptotic expectation and variance of the logrank statistic and the assumption of proportional hazards. *Lakatos* [19.23] proposed a Markov model to estimate the sample sizes for the comparison of lifetime distributions based on the logrank test. *Wu* et al. [19.24] provided a sample size computation method that allows time-dependent event (dropout) rate. *Lakatos* and *Lan* [19.25] compared several sample size calculation methods associated with the logrank test.

When the primary concern in a medical or reliability study is to compare the means or certain percentiles of two lifetime distributions, such as in the aging intervention testing program, the sample size determination based on proportional hazards assumption runs into the problem of expressing the difference or ratio of two means or two percentiles of lifetime distributions into the ratio of two hazard functions. Although this is no problem with exponential distributions or Weibull distributions with the same shape parameter, it might not always be possible for other families of proportional hazards. On the other hand, the location-scale family of log-transformed lifetime seems to be a very natural family of lifetime distributions to use for this type of sample size problem. This is based on the fact that,

when the scale parameters are assumed to be the same across different groups, the comparison of the means in lifetimes is equivalent to the comparison of location parameters. Although nonparametric tests such as the logrank test are appealing when the underlying lifetime distributions are unknown, they tend to be less efficient when the pilot information suggests a certain family of lifetime distributions and such information is ignored based on nonparametric tests. In fact, the logrank test bears 100% asymptotic relative efficiency when there is no censoring or when there is random but equal censoring in two groups within the family of Weibull distributions [19.26]. When data are from lognormal distributions differing only with respect to location, however, the asymptotic relative efficiency of the logrank test decreases to 82% [19.3].

Another common feature of most sample size determination methods in the literature is that they all deal with type I censored samples in which a prespecified time is used to terminate the experiment. This section studies the sample size determination to compare the means or certain percentiles of two lifetime distributions when both samples are subject to type II censoring. Our approach is based on a location-scale family of log-transformed lifetime distributions and the asymptotic normality of maximum likelihood estimates (MLEs). We also apply our methods to both the family of lognormal distributions and the family of Weibull distributions and compare our method with other well-known methods such as those based on *Rubinstein* et al. [19.18] and *Freedman* [19.22].

Although we will discuss the sample size determination in the context of designing a medical and biological study in this section, the basic ideas and results can be readily applied to the design of various engineering studies to compare the reliability of different groups of manufactured items. In fact, the application of our proposed methods to designing engineering studies will be even more intuitive to engineers as our approach is based on a location-scale family of log-transformed lifetime distributions, which is essentially equivalent to the accelerated failure time model popularly used in engineering and reliability studies.

19.1.1 The Model and Sample Size

Let T_c be the survival time for the control group, and T_t be the survival time for the treatment group. Assume that $Y_i = \ln T_i$ follows a probability distribution that belongs to a location-scale family with probability density function $\frac{1}{\sigma_i} g\left(\frac{y-\mu_i}{\sigma_i}\right)$, $-\infty < y < \infty$, $i = c, t$, where $g(s) > 0$ is a differentiable positive function whose derivative $g'(s) = \frac{dg(s)}{ds}$ makes all integrations used in this chapter exist. Let $G(s) = \int_{-\infty}^{s} g(t)\,dt$ be the cumulative distribution of $g(s)$. We also assume that both σ_c and σ_t are given and that $\sigma_c = \sigma_t = \sigma > 0$. The mean of $T_i = \exp(Y_i)$ is

$$ET_i = \int_{-\infty}^{\infty} e^y \frac{1}{\sigma} g\left(\frac{y-\mu_i}{\sigma}\right) dy = e^{\mu_i} \int_{-\infty}^{\infty} e^{\sigma s} g(s)\,ds\,, \tag{19.1}$$

for $i = c, t$. Hence,

$$\frac{ET_t}{ET_c} = e^{\mu_t - \mu_c}\,. \tag{19.2}$$

Let $0 < \delta < 1$. For $i = c, t$, a straightforward integration gives the $100\delta\%$ percentile of T_i as

$$\tau_i(\delta) = e^{\mu_i + \sigma G^{-1}(\delta)}\,, \tag{19.3}$$

where G^{-1} is the inverse function of G. It follows that

$$\frac{\tau_t(\delta)}{\tau_c(\delta)} = e^{\mu_t - \mu_c}\,. \tag{19.4}$$

Therefore, the problem of testing the ratio of two means or two percentiles between the control group and the treatment group can always be reduced to the problem of testing the difference between μ_t and μ_c.

Suppose that two independent samples of size n_c and n_t are drawn from the distributions of T_c and T_t, respectively. For $i = c, t$, we assume that only the smallest $100\,q_i\%$ of the samples are observed for some given $0 < q_i < 1$. If we let $r_i = [q_i n_i]$, then only the order statistics up to r_i-th are observed for group $i = c, t$. Let γ be the ratio of two sample sizes: $\gamma = \frac{n_t}{n_c}$. We want to decide the sample sizes for testing the null hypothesis $H_0 : \mu_c = \mu_t$ against the alternative $H_1 : \mu_c \neq \mu_t$ at an asymptotic significance level α $(0 < \alpha < 1)$. If this test is to achieve $100(1-\beta)\%$ power to detect a difference of $d = \mu_t - \mu_c$, the required sample size n_c for the control group is the unique solution to the following equation:

$$\beta = \Phi\left(z_{\alpha/2} - \frac{d}{\sqrt{\frac{1}{n_c}\left(\frac{1}{K_c^2} + \frac{1}{\gamma K_t^2}\right)}}\right) - \Phi\left(-z_{\alpha/2} - \frac{d}{\sqrt{\frac{1}{n_c}\left(\frac{1}{K_c^2} + \frac{1}{\gamma K_t^2}\right)}}\right)\,, \tag{19.5}$$

where $z_{\alpha/2}$ is the upper $\alpha/2$ percentage point of the standard normal distribution, Φ is the cumulative distribution function of the standard normal distribution, K_i, $i = \mathrm{c}, \mathrm{t}$, is given by

$$K_i^2 = \frac{1}{\sigma^2} \int_{-\infty}^{(\lambda_i-\mu_i)/\sigma} \frac{[g'(s)]^2}{g(s)}\,\mathrm{d}s + \frac{1}{p_i\sigma^2}\left[g\left(\frac{\lambda_i-\mu_i}{\sigma}\right)\right]^2 , \tag{19.6}$$

and λ_i is such that

$$q_i = G\left(\frac{\lambda_i-\mu_i}{\sigma}\right) , \tag{19.7}$$

and $p_i = 1 - q_i$. The required sample size for the treatment group is then $n_\mathrm{t} = \gamma n_\mathrm{c}$. The proof of (19.5) is based on the asymptotic normality of the MLEs of μ_i and can be found in [19.27].

If we want to test the null hypothesis $H_0 : \mu_\mathrm{c} = \mu_\mathrm{t}$ against the one-sided alternative $H_1 : \mu_\mathrm{t} > \mu_\mathrm{c}$ at an asymptotic significance level α $(0 < \alpha < 1)$ and assume that this test is to achieve $100(1-\beta)\%$ power to detect a difference of $d = \mu_\mathrm{t} - \mu_\mathrm{c} > 0$, the required sample size for the control group n_c is given by

$$n_\mathrm{c} = \left(\frac{z_\alpha + z_\beta}{d}\right)^2 \left(\frac{1}{K_\mathrm{c}^2} + \frac{1}{\gamma K_\mathrm{t}^2}\right) , \tag{19.8}$$

and the required sample size for the treatment group is then $n_\mathrm{t} = \gamma n_\mathrm{c}$. The proof of (19.8) can also be found in [19.27].

19.1.2 Examples

Since the family of lognormal distributions and the family of Weibull distributions are two important location-scale families of log-transformed lifetime distributions, we apply our method to these two families.

Example 1: Lognormal Distribution

We first study the family of lognormal distribution in which

$$g(s) = \frac{1}{\sqrt{2\pi}}\,\mathrm{e}^{-s^2/2} . \tag{19.9}$$

For $i = \mathrm{c}, \mathrm{t}$, using integration by parts, we find

$$K_i^2 = -\frac{(\lambda_i-\mu_i)}{\sigma^3} g\left(\frac{\lambda_i-\mu_i}{\sigma}\right) + \frac{q_i}{\sigma^2} + \frac{1}{p_i\sigma^2}\left[g\left(\frac{\lambda_i-\mu_i}{\sigma}\right)\right]^2 , \tag{19.10}$$

where λ_i is such that

$$q_i = \Phi\left(\frac{\lambda_i-\mu_i}{\sigma}\right) . \tag{19.11}$$

If $q_\mathrm{c} = q_\mathrm{t} = 50\%$, then

$$K_i^2 = \frac{1}{2\sigma^2} + \frac{1}{\pi\sigma^2} . \tag{19.12}$$

Example 2: Weibull Distribution

In the family of Weibull distributions,

$$g(s) = \exp(s - \mathrm{e}^s) . \tag{19.13}$$

For $i = \mathrm{c}, \mathrm{t}$, by repeatedly using the technique of integration by parts, we have

$$K_i^2 = \frac{1-p_i}{\sigma^2} . \tag{19.14}$$

If $q_\mathrm{c} = q_\mathrm{t} = 50\%$, then

$$K_i^2 = \frac{1}{2\sigma^2} . \tag{19.15}$$

19.1.3 Effect of Guarantee Time on Sample Size Determination

A very simple feature of lifetime distributions is the existence of a threshold time, or *guarantee time*, during which no subjects will die. For example, the type of mice to be used in the aging intervention testing program of the National Institute on Aging exhibit a guarantee survival time of about 500 days in the survival distribution, as estimated from the survival curves reported by *Turturro* et al. [19.28]. When the comparison of two mean lifetimes is in terms of the difference and when the two distributions share the same guarantee time, this time contributes nothing to the comparison. When the comparison of two mean lifetimes is in terms of the ratio, however, the guarantee time plays an important role in the comparison, especially in the determination of sample sizes at the design stage of the clinical trials.

When the primary concern in a medical study is to compare the means or certain percentiles of two lifetime distributions with type II censored observations such as the aging intervention testing program from the National Institute on Aging, the sample size determination may be based on the method of *Rubinstein* et al. [19.18], the method of *Freedman* [19.22], and the method described by (19.5) and (19.8). The methods of *Rubinstein* et al. [19.18] can be used since the hazard ratio is simply

the reciprocal of the ratio of two means under the exponential distributions. The method of *Freedman* [19.22] refers to the logrank test of the hazard ratio and requires the assumptions of proportional hazards between two groups. It could be used to compare the means or certain percentiles of two lifetime distributions as long as the comparison of means or certain percentiles can be related to the hazard ratio between two distributions such as in the family of Weibull distributions. The method described by (19.5) and (19.8) directly applies to the comparison of means or certain percentiles of two lifetime distributions and requires the assumption of location-scale family of log-transformed lifetime distributions.

If two lifetime distributions under study exhibit survival thresholds, or *guarantee times* as in the survival distribution of mice used in the aging intervention testing program from the National Institute on Aging, these thresholds have an important role in the ratio of means from two distributions. Mathematically, let T_c be the survival time for the control group, and T_t be the survival time for the treatment group. Let ET_c and ET_t be the corresponding means of the two distributions. We are interested in testing the null hypothesis $H_0 : ET_t/ET_c = 1$ against the alternative $H_1 : ET_t/ET_c \neq 1$ at an asymptotic significance level α $(0 < \alpha < 1)$. For $i = c, t$, suppose that T_i follows a distribution with a threshold, or a *guarantee time* $\psi_i > 0$. We assume that both the control group and the treatment group share the same threshold parameter $\psi_1 = \psi_2 = \psi$ and that ψ is known. Let $T_i' = T_i - \psi$, then $ET_i' = ET_i - \psi$, $i = t, c$. The alternative hypothesis on which the sample size computation is based is

$$\rho = \frac{ET_t}{ET_c} = \frac{ET_t' + \psi}{ET_c' + \psi} . \tag{19.16}$$

Therefore

$$\frac{ET_t'}{ET_c'} = \frac{\rho ET_c - \psi}{ET_c - \psi} . \tag{19.17}$$

Since $ET_t/ET_c = 1$ if and only if $ET_t'/ET_c' = 1$. The original null and alternative hypotheses translate into the corresponding hypotheses in terms of distributions T_t' and T_c' : $H_0' : ET_t'/ET_c' = 1$ and $H_1' : ET_t'/ET_c' \neq 1$.

Since the distributions of T_t' and T_c' begin with time 0 and have no guarantee times, the sample size methods reviewed above can be directly applied to test the reduced hypotheses H_0' against H_1'. The alternative hypothesis on which the sample size computation should be based, however, now becomes $ET_t'/ET_c' = (\rho ET_c - \psi)/(ET_c - \psi)$. We call ET_t'/ET_c' the adjusted effect size. Note that for any $\rho \neq 1$, $ET_t'/ET_c' = \rho$ if and only if $\psi = 0$. Hence, when the ratio of the mean between two lifetime distributions is to be tested, it is crucial that the sample size determination based on the logrank test and the proportional hazards assumption or the location-scale family of log-transformed lifetime distributions takes into account the possible guarantee time in the lifetime distributions. Table 19.1 presents the sample size computation based on the method of *Rubinstein* et al. [19.18] for a selected set of the guarantee time ψ and the percentage p_t of censorship for the treatment group. Table 19.2 presents the sample size computation based on the method of *Freedman* [19.22] for the same selected set of ψ and p_t. The computation in Table 19.2 assumes Weibull distributions with the same shape parameter of 1.5 for both the treatment group and the control group so that the ratio of two means can be expressed as a function of the ratio of hazard functions between the two groups. Table 19.3 presents the sample size computation based on the method described by (19.8) for the same selected set of ψ and p_t under the assumption of lognormal distributions with a scale parameter of 0.8 in the log-transformed lifetime distribution. Table 19.4 presents the sample size computation based on the method described by (19.8) for the same selected set of ψ and p_t under the assumption of Weibull distributions with a scale parameter of 0.8 in the log-transformed lifetime distribution. All these computations in four tables are based on a one-sided test for

Table 19.1 Sample size per group based on the method of *Rubinstein*, et al. [19.18] $\alpha = 5\%$, $\beta = 20\%$

ψ	$p_t = 40\%$	$p_t = 50\%$	$p_t = 60\%$
0	589	702	871
10	518	617	765
20	452	537	666
30	390	463	574
40	332	394	489
50	279	331	410
60	231	274	339
70	187	222	274
80	148	175	216
90	114	134	165
100	84	98	121
110	58	68	83
120	37	43	53
130	21	24	29
140	9	10	12

the ratio of two means at a significance level of 5%, a mean lifetime of 150 units for the control group, and a statistical power of 80%. The 80% power is assumed at $\rho = 1.2$ in the original alternative hypotheses with the equal sample size between the treatment group and the control group. In addition, both groups are assumed the simultaneous entry to the study and the simultaneous stopping time, which is the time of achieving censorship p_t for the treatment group. The censorship for the control group is then decided by ρ and p_t under the appropriate distributional assumptions and is less than p_t by the assumption that $\rho > 1$. The sample size when there is no *guarantee time* or the *guarantee time* is ignored is given when $\psi = 0$.

Although Tables 19.1–19.4 are based on different sample size determination methods, they demonstrate two common important observations. First, if a guarantee time exists in the two lifetime distributions to be compared, the ignorance of the guarantee time leads to the overestimation of the sample size. Second, if the significance level of the test, the statistical power of the test, and the degree of censoring are fixed, the required sample size decreases as the guarantee time increases. All these can be explained by (19.17), which expresses the ratio of two means ET_t'/ET_c' after the guarantee time ψ is subtracted (i. e., the adjusted effect size) as a function of the guarantee time ψ, the original ratio ρ of mean lifetime when the guarantee time is included, and the mean lifetime ET_c for the control group. Let $\rho > 1$ and ET_c be fixed. We denote the adjusted effect size [the right-hand side of (19.17)] by $h(\psi)$. This function has two important features. First, $h(\psi) = \rho$ if and only if $\psi = 0$. Second, the derivative of $h(\psi)$ with respect to ψ is always positive, which implies that it is an increasing function of ψ. Since the sample size methods are applied to the lifetime distributions when the guarantee time is subtracted, the effect size used in these sample size computations is based on $h(\psi)$ instead of ρ. The fact that $h(\psi) = \rho$ if and only if $\psi = 0$ implies that the ig-

Table 19.2 Sample size per group based on the method of *Freedman* [19.22] (Weibull distribution with a shape parameter 1.5 assumed) $\alpha = 5\%$, $\beta = 20\%$

ψ	$p_t = 40\%$	$p_t = 50\%$	$p_t = 60\%$
0	258	305	377
10	227	268	331
20	198	233	288
30	171	201	248
40	146	171	210
50	122	144	176
60	101	119	145
70	83	96	118
80	66	76	93
90	51	59	71
100	38	44	52
110	27	31	36
120	18	20	24
130	12	13	14
140	7	7	8

Table 19.3 Sample size per group based on (19.8); The lognormal case $\alpha = 5\%$, $\beta = 20\%$, $\sigma = 0.8$

ψ	$p_t = 40\%$	$p_t = 50\%$	$p_t = 60\%$
0	267	283	307
10	235	249	270
20	205	217	235
30	178	188	203
40	152	161	174
50	128	135	146
60	106	112	121
70	87	91	99
80	69	73	78
90	53	56	60
100	40	42	45
110	28	29	31
120	18	19	20
130	11	11	12
140	5	5	5

Table 19.4 Sample size per group based on (19.8); the Weibull case $\alpha = 5\%$, $\beta = 20\%$, $\sigma = 0.8$

ψ	$p_t = 40\%$	$p_t = 50\%$	$p_t = 60\%$
0	377	449	558
10	332	395	490
20	289	344	427
30	249	297	368
40	213	253	313
50	179	212	263
60	148	175	217
70	120	142	175
80	95	112	138
90	73	86	106
100	54	63	77
110	37	44	53
120	24	28	34
130	14	16	19
140	6	7	8

norance of the guarantee time (i. e., by assuming $\psi = 0$) will always lead to an inadequate sample size when in fact $\psi > 0$. Since a common feature of the sample size methods is that the sample size decreases when the effect size $h(\psi)$ increases, this explains why the sample size decreases as the guarantee time ψ increases from Tables 19.1 to 19.4.

19.1.4 Application to NIA Aging Intervention Testing Program

We now demonstrate the sample size determination by applying it to the aging intervention testing program (RFA-AG-02-005) of the National Institute on Aging (NIA). One of the major research objectives of this program is to identify interventions that increase mean life expectancy by 10% in phase 1 studies which may be terminated at 50% survivorship. The experimental units in this study are the 4WCNIA mice obtained from the National Institute of Health (NIH) aging rodent colony. Pilot data such as the survival curves reported by *Turturro* et al. [19.28] on similar mice have suggested a guarantee survival time of about 500 days in the survival distribution. In addition, *Pugh* et al. [19.29] reported a mean life expectancy of 876 days and a standard deviation of 18 days for similar mice.

Assume that we ignore the guarantee time in the sample size computation and use $\rho = 1.1$ as the ratio of mean lifetime between the intervention group and the control group. The method of *Rubinstein* et al. [19.18] gives a sample size of 2637 per group. The method of *Freedman* [19.22] gives a sample size of 323 per group based on two Weibull distributions with the same shape parameter which is estimated as 2.793 by the survival curves reported by *Turturro* et al. [19.28]. Our proposed method gives a sample size of 387 per group based on two lognormal distributions with the same scale parameter σ (in the log-transformed lifetime). This computation uses $\sigma = 0.482$ as estimated by the survival curves reported in [19.28]. When applied under the family of Weibull distributions, the projected sample size per group based on our method is 338.

When the 500-days guarantee survival time is taken into account in the sample size computation, these methods are applied to the survival distributions after the 500-days guarantee survival time is subtracted. The pilot information of $ET_c = 876$ and $\rho = 1.1$ along with (19.17) implies that $ET_t'/ET_c' = 1.233$. The sample size methods are then applied to the distributions of T_t' and T_c' when testing H_0' against H_1' at a 5% significance level and an 80% statistical power. We assume that both groups use the same number of mice, that the treatment group is terminated at the 50% censorship, and that the control group is allowed to continue until the treatment group terminates. The method of *Rubinstein* et al. [19.18] gives a sample size of 528 per group. The method of *Freedman* [19.22] under the assumption of Weibull distributions gives a sample size of 64 per group. Assuming the lognormal distribution for the lifetimes with the same scale parameter σ in the log-transformed lifetime distributions between the control and treatment groups, our proposed method gives the sample size required per group as 81. Assuming a Weibull distribution for the lifetimes with the same scale parameter σ in the log-transformed lifetime distributions between the control and treatment groups, our proposed method gives the sample size required per group as 68.

Similar to observations from Tables 19.1–19.4, the real-life example again demonstrates the importance of taking into account the guarantee time in sample size computation when it exists. A considerable waste of resources would occur if the guarantee time is ignored in the sample size projection. Notice also that, while the methods of *Freedman* [19.22] and ours give fairly consistent results about the sample size per group, the method of *Rubinstein* et al. [19.18], however, results in a very different sample size compared to the others. The reason behind this difference is the assumption of an exponential distribution for the method in [19.18]. The mathematically attractive but practically unrealistic property of the exponential distribution is its constant hazard function over time, which then implies the memoryless feature for the survival distribution [19.3]. Although the exponential distribution is a distribution extensively discussed in the fields of biometrics, reliability and industrial life testing literature [19.30, 31], it has long been pointed out by many authors such as *Zelen* and *Dannemiller* [19.32] that the estimations and inferences associated with an exponential distribution are not robust and that exponential distribution is a very unrealistic distribution in many applications, especially in studies associated with the aging process.

19.2 Analysis of Survival Data from Special Cases of Step-Stress Life Tests

We now discuss some applications of survival analysis in engineering and reliability studies. Accelerated life tests (ALT) consist of a variety of test methods for shortening the life of products or hastening the degradation of their performance. The aim of such testing is to obtain data quickly which, properly modeled and analyzed, yield desired information on product life or performance under normal use. ALT can be carried out using constant stress, step-stress, or linearly increasing stress. The step-stress scheme applies stress to test units in the way that the stress setting of test units will be changed at specified times. Generally, a test unit starts at a specified low stress. If the unit does not fail at a specified time, the stress on it is raised and held for a specified time. The stress is repeatedly increased and held, until the test unit fails or a censoring time is reached. A simple step-stress ALT (SSALT) uses only two stress levels. The problem of modeling data from ALT and making inferences from such data has been studied by many authors. *Chernoff* [19.33] considered optimal life tests for estimation of model parameters based on data from ALT. *Meeker* and *Nelson* [19.34] obtained optimum ALT plans for Weibull and extreme-value distributions with censored data. *Nelson* and *Kielpinski* [19.35] further studied optimum ALT plans for normal and lognormal life distributions based on censored data. *Nelson* [19.36] considered data from SSALT and obtained maximum likelihood estimates (MLE) for the parameters of a Weibull distribution under the inverse power law using the breakdown time data of an electrical insulation. *Miller* and *Nelson* [19.37] studied optimum test plans which minimized the asymptotic variance of the MLE of the mean life at a design stress for simple SSALT where all units were run to failure. *Bai* et al. [19.38] further studied the similar optimum simple SSALT plan for the case where a specified censoring time was involved. *Tyoskin* and *Krivolapov* [19.39] presented a nonparametric approach for making inferences for SSALT data. *Dorp* et al. [19.40] developed a Bayes model and studied the inferences of data from SSALT. *Xiong* [19.41] obtained inferences based on pivotal quantities for type II censored exponential data from a simple SSALT. *Alhadeed* and *Yang* [19.42] discussed the optimal simple step-stress plan for the Khamis–Higgins model. *Teng* and *Yeo* [19.43] used the method of least squares to estimate the life–stress relationship in SSALT. *Hobbs* [19.44] gave detailed discussion on highly accelerated life test (HALT) and highly accelerated stress screens (HASS). *Mann* et al. [19.45] and *Lawless* [19.3] provided the general theory and applications of lifetime data analysis. *Meeker* and *Escobar* [19.46] briefly surveyed optimum test plans for different types of ALT. *Nelson* [19.1, 47] provided an extensive and comprehensive source for theory and examples for ALT and SSALT.

During the step-stress life test, test units can be continuously or intermittently inspected for failure. The latter type of test is frequently used since it generally requires less testing effort and can be administratively more convenient. In some other cases, intermittent inspection is the only feasible way of checking the status of test units (see, for example, [19.48]). The data obtained from intermittent inspections are called grouped data and consist of only the number of failures in the inspection intervals. The first problem we consider in this section is the statistical inference of model parameters and optimum test plans based on only grouped and type I censored data obtained from a step-stress life test. We will also study another important and interesting variation associated with grouped and censored data from a simple SSALT, when both the stress change time and the censoring time are random variables, such as the order statistics at the current stress levels, and when only these order statistics (stress-change time and type II censoring time) are observed during the test.

Throughout the section, we denote the design stress by x_0, the i-th test stress by x_i, $i = 1, 2, \cdots, m$, $x_1 < x_2 < \cdots < x_m$, where m is the total number of test stress levels. We assume that the i-th stress change time is constant τ_i, $i = 1, 2, \cdots, m-1$, and the fixed censoring time is $\tau_m > \tau_{m-1}$. Let $\tau_{m+1} = \infty$, $\tau_0 = 0$, $\Delta\tau_i = \tau_i - \tau_{i-1}$. We also make following assumptions:

(A1). At any constant stress x_i, $i = 0, 1, 2, \cdots, m$, the cumulative distribution function (CDF)of a test unit lifetime is

$$F_i(t) = F(t/\theta_i) \quad \text{for} \quad t > 0\,, \tag{19.18}$$

where the stress–response relationship (SRR) θ_i is a function of stress x_i and F is a strictly increasing distribution function.

(A2). The stresses are applied in the order $x_1 < x_2 < \cdots < x_m$.

(A3). The lifetimes of test units under SSALT are statistically independent.

For the step-stress life test, there is a probability distribution $G(t)$ of time T to failure on test. Data from this distribution are observed during the test. The cumulative exposure model of time T assumes that the remaining

life of a test unit depends only on the current cumulative fraction failed and the current stress, regardless of how the fraction is accumulated. Moreover, if held at the current stress, survivors will fail according to the cumulative distribution for that stress but starting at the previously accumulated fraction failed. Also, the change in stress has no effect on life, only the level of the stress does. As pointed out by *Miller* and *Nelson* [19.37] and *Yin* and *Sheng* [19.49], this model has many applications in industrial life testing.

Mathematically, the cumulative distribution $G(t)$ of time T to failure from a step-stress test described above is

$$G(t) = \begin{cases} F_i(t - \tau_{i-1} + s_{i-1}) \,, \\ \quad \text{for} \quad \tau_{i-1} \le t < \tau_i \,, \\ \qquad\qquad i = 1, 2, \cdots, m-1, \\ F_m(t - \tau_{m-1} + s_{m-1}) \,, \\ \quad \text{for} \quad \tau_{m-1} \le t < \infty \,, \end{cases} \tag{19.19}$$

where s_{i-1} is the equivalent start time at step i satisfying

$$F_i(s_{i-1}) = F_{i-1}(\tau_{i-1} - \tau_{i-2} + s_{i-2}) \,. \tag{19.20}$$

We further assume:

(A4). The stress-response relationship (SRR) $\theta_i = \theta(x_i)$ is a log–linear function of stress x_i. That is,

$$\log \theta(x_i) = \alpha + \beta x_i \,, \tag{19.21}$$

where α and β are unknown model parameters which typically depend on the nature of the product and the test method. Although the specification of $\theta(x)$ looks rather restrictive, it covers some of the most important models used in industry, such as the power-law model, the Eyring model and the Arrhenius model [19.47]. With the above specifications, it is straightforward to find that $s_{i-1} = (\tau_{i-1} - \tau_{i-2} + s_{i-2})\theta_i/\theta_{i-1} = \theta_i \sum_{j=1}^{i-1} \Delta\tau_j/\theta_j$, $i = 2, 3, 4, \cdots, m$. Thus, the distribution function of the step-stress failure time T is

$$G(t) = \begin{cases} F\left(\dfrac{t - \tau_{i-1}}{\theta_i} + \sum_{j=1}^{i-1} \dfrac{\Delta\tau_j}{\theta_j}\right) \,, \\ \quad \text{for } \tau_{i-1} \le t < \tau_i, \, i = 1, 2, ..., m-1 \,. \\ F\left(\dfrac{t - \tau_{m-1}}{\theta_m} + \sum_{j=1}^{m-1} \dfrac{\Delta\tau_j}{\theta_j}\right) \,, \\ \quad \text{for } \tau_{m-1} \le t < \infty \end{cases} \tag{19.22}$$

Although we will discuss the statistical analysis of data collected from SSALT in the context of engineering studies, we point out that the statistical models and methods discussed in this section can be readily used in medical and biological research. Many clinical trials on therapeutical compounds of a disease contain a novel treatment group and a control group. At the completion of well-designed clinical studies, especially when the novel treatment has been found efficacious for the protection against the disease development based on the available data, many clinical trials are extended for another specified time period so that subjects from the original control group can receive the treatment. When a survival endpoint such as the time from the study baseline to the onset of a specific disease is measured on subjects from the original control group throughout the entire trial period, the resulting survival data are very analogous to these collected from the standard SSALT in engineering studies. In fact, the drug dose used in such clinical trials can be thought of as the stress level in which 0 is the stress at the initial phase of the original control group and a positive dose is the stress at the extension phase of the trial.

19.2.1 Analysis of Grouped and Censored Data from Step-Stress Life Tests

MLE

We first consider the case when all test units are subject to the same censoring time and the same stress-change patterns with the same set of stresses and the same stress-change times. We assume that data obtained from such step-stress tests are grouped and type I censored. More specifically, we assume that the intermittent inspection times during the step-stress test coincide with the stress-change times and the censoring time. Suppose that n test units begin at low stress x_1. If the unit does not fail at a specified time τ_1, the stress on it is raised to a higher stress x_2. The stress is repeatedly increased and held in this fashion, until the test unit fails or the fixed censoring time τ_m ($> \tau_{m-1}$) is reached. Assume that n_i units fail during the inspection time interval $[\tau_{i-1}, \tau_i)$, $i = 1, 2, \cdots, m, m+1$, $(\tau_{m+1} = \infty)$. To simplify the notations, we denote, for $i = 1, 2, \cdots, m+1$,

$$\begin{aligned} u_i(\alpha, \beta) &= \sum_{j=1}^{i} \Delta\tau_j \, \exp(-\alpha - \beta x_j) \\ v_i(\alpha, \beta) &= \sum_{j=1}^{i} x_j \Delta\tau_j \, \exp(-\alpha - \beta x_j) \,. \end{aligned} \tag{19.23}$$

Let $p_i = \Pr(\tau_{i-1} \le T < \tau_i)$ for $1 \le i \le m+1$. The cumulative exposure model (19.22) implies that, for

$1 \le i \le m+1$,

$$p_i = F[u_i(\alpha,\beta)] - F[u_{i-1}(\alpha,\beta)] \,. \tag{19.24}$$

The likelihood function based on data vector $(n_1, n_2, \cdots, n_{m+1})$ is (up to a constant):

$$L(\alpha,\beta) \propto \prod_{i=1}^{m+1} \{F[u_i(\alpha,\beta)] - F[u_{i-1}(\alpha,\beta)]\}^{n_i} \,, \tag{19.25}$$

where θ_i is specified by the SRR (19.21) and $F(\infty) = 1$. Thus, the log likelihood function is a function of the unknown parameters α and β :

$$\log L(\alpha,\beta) \propto \sum_{i=1}^{m+1} n_i \log\{F[u_i(\alpha,\beta)] - F[u_{i-1}(\alpha,\beta)]\} \,. \tag{19.26}$$

To find the maximum likelihood estimators (MLE) for α and β, we maximize $\log L(\alpha,\beta)$ over α and β. The maximization of $\log L(\alpha,\beta)$ requires the solution to the system:

$$\begin{cases} \dfrac{\partial L(\alpha,\beta)}{\partial \alpha} = -\sum_{i=1}^{m+1} n_i \times \dfrac{u_i(\alpha,\beta) f[u_i(\alpha,\beta)]}{F[u_i(\alpha,\beta)] - F[u_{i-1}(\alpha,\beta)]} - \dfrac{u_{i-1}(\alpha,\beta) f[u_{i-1}(\alpha,\beta)]}{F[u_i(\alpha,\beta)] - F[u_{i-1}(\alpha,\beta)]} = 0 \,, \\ \dfrac{\partial L(\alpha,\beta)}{\partial \beta} = -\sum_{i=1}^{m+1} n_i \times \dfrac{v_i(\alpha,\beta) f[u_i(\alpha,\beta)]}{F[u_i(\alpha,\beta)] - F(u_{i-1}(\alpha,\beta))} - \dfrac{v_{i-1}(\alpha,\beta) f[u_{i-1}(\alpha,\beta)]}{F[u_i(\alpha,\beta)] - F(u_{i-1}(\alpha,\beta))} = 0 \,, \end{cases} \tag{19.27}$$

where $f(t) = \mathrm{d}F(t)/\mathrm{d}t$ is the probability density function of $F(t)$. Generally, the solution of (19.27) requires a numerical method such as Newton–Raphson. *Seo* and *Yum* [19.50] proposed several approximate ML estimators and compared with the MLE by a Monte Carlo simulation when the lifetime distribution is assumed exponential. The expected second partial derivatives of the log likelihood function at $(\alpha,\beta)'$ are

$$\begin{aligned} \sigma_{11} &= -E\frac{\partial^2 L(\alpha,\beta)}{\partial \alpha^2} \\ &= -n\sum_{i=1}^{m+1} \frac{\partial^2 p_i}{\partial \alpha^2} + \sum_{i=1}^{m+1} \frac{1}{p_i}\left(\frac{\partial p_i}{\partial \alpha}\right)^2 \,, \\ \sigma_{12} &= -E\frac{\partial^2 L(\alpha,\beta)}{\partial \alpha \partial \beta} \\ &= -n\sum_{i=1}^{m+1} \frac{\partial^2 p_i}{\partial \alpha \partial \beta} + \sum_{i=1}^{m+1} \frac{1}{p_i}\frac{\partial p_i}{\partial \alpha}\frac{\partial p_i}{\partial \beta} \,, \\ \sigma_{22} &= -E\frac{\partial^2 L(\alpha,\beta)}{\partial \beta^2} \\ &= -n\sum_{i=1}^{m+1} \frac{\partial^2 p_i}{\partial \beta^2} + \sum_{i=1}^{m+1} \frac{1}{p_i}\left(\frac{\partial p_i}{\partial \beta}\right)^2 \,. \end{aligned} \tag{19.28}$$

Let D be the $(m+1)$ by $(m+1)$ diagonal matrix with $1/p_i$, $i = 1, 2, \cdots, m+1$, as its diagonal elements. Let $J = (j_{st})$, $1 \le s \le m+1$, $1 \le t \le 2$, be the $(m+1)\times 2$ Jacobian matrix of $(p_1, p_2, \cdots, p_{m+1})'$ with respect to $(\alpha,\beta)'$, i. e.,

$$\begin{aligned} j_{s1} = \frac{\partial p_i}{\partial \alpha} &= -u_i(\alpha,\beta) f[u_i(\alpha,\beta)] \\ &\quad + u_{i-1}(\alpha,\beta) f[u_{i-1}(\alpha,\beta)] \,, \\ j_{s2} = \frac{\partial p_i}{\partial \beta} &= -v_i(\alpha,\beta) f[u_i(\alpha,\beta)] \\ &\quad + v_{i-1}(\alpha,\beta) f[u_{i-1}(\alpha,\beta)] \,, \end{aligned} \tag{19.29}$$

for $s = 1, 2, \cdots, m+1$. Because

$$E\left(\sum_{i=1}^{m+1} \frac{\partial p_i}{\partial \alpha}\right) = E\left(\sum_{i=1}^{m+1} \frac{\partial p_i}{\partial \beta}\right) = 0 \,,$$

the expected Fisher information matrix $\boldsymbol{\Sigma} = (\sigma_{ij})$, $i, j = 1, 2$, is given by $\boldsymbol{\Sigma} = n \cdot J' D J$. Let $(\widehat{\alpha}, \widehat{\beta})'$ be the MLE of $(\alpha,\beta)'$ obtained from solving (19.27). $n^{-1}\boldsymbol{\Sigma}$ can be consistently estimated by $n^{-1}\widehat{\boldsymbol{\Sigma}}$, where $\widehat{\boldsymbol{\Sigma}} = (\widehat{\sigma_{ij}})$, $i, j = 1, 2$, is obtained by replacing $(\alpha,\beta)'$ in $\boldsymbol{\Sigma}$ by its MLE $(\widehat{\alpha}, \widehat{\beta})'$.

Based on the asymptotic normality of $(\widehat{\alpha}, \widehat{\beta})'$ with estimated covariance matrix $\widehat{\Sigma}^{-1}$, we can set up the asymptotic confidence interval (CI) for α, β, the SRR of lifetime at design stress $\theta_0 = \theta(x_0) = \exp(\alpha + \beta x_0)$, and the reliability function at design stress $R_0(t) = 1 - F(t/\theta_0)$. Let $\widehat{\Sigma}^{-1} = (\widehat{m_{ij}})$, $i, j = 1, 2$, be the estimated asymptotic covariance matrix of $(\widehat{\alpha}, \widehat{\beta})'$. It is straightforward to show that an asymptotic $100(1-\gamma)\%$ CI for α is $\widehat{\alpha} \pm z_{\gamma/2}\widehat{m}_{11}$, and an asymptotic $100(1-$

γ)% CI for β is $\widehat{\beta} \pm z_{\gamma/2}\widehat{m}_{22}$, where $z_{\gamma/2}$ is the $\gamma/2$ point of the standard normal distribution. The asymptotic variance for $\log\widehat{\theta(x_0)} = \widehat{\alpha} + \widehat{\beta}x_0$ is given by

$$\widehat{\sigma} = (1, x_0)\widehat{\boldsymbol{\Sigma}}^{-1}(1, x_0)' \,. \tag{19.30}$$

An asymptotic $100(1-\gamma)$% CI for θ_0 is

$$\exp\left(\widehat{\alpha} + \widehat{\beta}x_0 \pm z_{\gamma/2}\widehat{\sigma}\right) \,.$$

Finally, because $F(t)$ is a strictly increasing function of t, an asymptotic $100(1-\gamma)$% CI for $R_0(t) = 1 - F(t/\theta_0)$ at a given time t is $1 - F\left[t/\exp\left(\widehat{\alpha} + \widehat{\beta}x_0 \pm z_{\gamma/2}\widehat{\sigma}\right)\right]$.

When the step-stress test is a simple SSALT, i. e., when $m = 2$, there exist closed form MLE for α and β. The MLE of α and β solves

$$\begin{cases} F\left(\dfrac{\tau_1}{\theta_1}\right) & = \dfrac{n_1}{n} \\ F\left(\dfrac{\tau_2 - \tau_1}{\theta_2} + \frac{\tau_1}{\theta_1}\right) - F\left(\dfrac{\tau_1}{\theta_1}\right) & = \dfrac{n_2}{n} \,. \end{cases} \tag{19.31}$$

The solutions are

$$\begin{aligned} \widehat{\alpha} &= \frac{x_2}{x_2 - x_1}\log\frac{\tau_1}{F^{-1}\left(\frac{n_1}{n}\right)} \\ &\quad - \frac{x_1}{x_2 - x_1}\log\frac{(\tau_2 - \tau_1)}{F^{-1}\left(\frac{n_1+n_2}{n}\right) - F^{-1}\left(\frac{n_1}{n}\right)} \,, \\ \widehat{\beta} &= \frac{1}{x_2 - x_1}\log\frac{(\tau_2 - \tau_1)F^{-1}\left(\frac{n_1}{n}\right)}{\tau_1\left[F^{-1}\left(\frac{n_1+n_2}{n}\right) - F^{-1}\left(\frac{n_1}{n}\right)\right]} \,, \end{aligned} \tag{19.32}$$

where F^{-1} is the inverse function of F.

In the more general situation when different test units are subject to different censoring times and different stress-change patterns with different sets of stresses and even different stress-change times, the likelihood function for each test unit can be given by (19.25) for each test unit with $\sum_{i=1}^{m+1} n_i = 1$. The likelihood function for a sample of size n test units is the product of all n individual likelihood functions by assumption (A3). The MLE of α and β can be obtained by maximizing this likelihood function using a numerical method such as Newton–Raphson. Although the lifetime distributions of n test units are independent, they are not identical. The asymptotic confidence interval estimates for various model parameters given above, however, are still valid when the Fisher information matrix $\boldsymbol{\Sigma}/n$ is replaced by the average information matrix to take into account of the difference in the lifetime distributions. The detailed theoretical justification can be found in Chapt. 9 of *Cox* and *Hinkley* [19.51].

A Statistical Test for the Cumulative Exposure Model when m>2

We only consider the case when all test units are subject to the same censoring time and the same stress-change patterns with the same set of stresses and the same stress-change times in this section. We again let $p_i = \Pr(\tau_{i-1} \le T < \tau_i)$ for $1 \le i \le m+1$. The cumulative exposure model (19.22) implies that

$$p_i = F[u_i(\alpha, \beta)] - F[u_{i-1}(\alpha, \beta)] \,. \tag{19.33}$$

A statistical test for the cumulative exposure model can be obtained by testing the null hypothesis $H_0: \; p_i = F[u_i(\alpha, \beta)] - F[u_{i-1}(\alpha, \beta)], \; 1 \le i \le m+1$, against the alternative H_a : there is no constraint on p_i, $1 \le i \le m+1$. When grouped and type I censored data are available from n test units, the likelihood function of p_i, $1 \le i \le m+1$, is

$$L \propto \prod_{i=1}^{m+1} {p_i}^{n_i} \,. \tag{19.34}$$

The MLE of p_i, $1 \le i \le m+1$, under H_0 are given by

$$\widehat{p_i^0} = F\left[u_i\left(\widehat{\alpha}, \widehat{\beta}\right)\right] - F\left[u_{i-1}\left(\widehat{\alpha}, \widehat{\beta}\right)\right] \,, \tag{19.35}$$

where $\widehat{\alpha}, \widehat{\beta}$ are the MLE of α and β. Under H_a, a straightforward maximization of the likelihood function gives the MLE of p_i, $1 \le i \le m+1$, as

$$\widehat{p_i^a} = \frac{n_i}{n} \,. \tag{19.36}$$

Therefore, an asymptotic likelihood ratio test of significance level γ $(0 < \gamma < 1)$ rejects H_0 if

$$\begin{aligned} &-2\sum_{i=1}^{m+1} n_i \left(\log\left\{F\left[u_i\left(\widehat{\alpha}, \widehat{\beta}\right)\right]\right. \right. \\ &\left. \left. - \; F\left[u_{i-1}\left(\widehat{\alpha}, \widehat{\beta}\right)\right]\right\} - \log\frac{n_i}{n}\right) \\ &> \chi^2_\gamma(m-2) \,, \end{aligned} \tag{19.37}$$

where $\chi^2_\gamma(m-2)$ is the upper 100γ% percentile of the χ^2 distribution with $m-2$ degrees of freedom. Because

$$\begin{aligned} &-2\sum_{i=1}^{m+1} n_i \left(\log\left\{F\left[u_i\left(\widehat{\alpha}, \widehat{\beta}\right)\right]\right. \right. \\ &\left. \left. - \; F\left[u_{i-1}\left(\widehat{\alpha}, \widehat{\beta}\right)\right]\right\} - \log\frac{n_i}{n}\right) \end{aligned} \tag{19.38}$$

is stochastically equivalent to

$$\sum_{i=1}^{m+1} \frac{\left(n_i - n\left\{F\left[u_i\left(\widehat{\alpha},\widehat{\beta}\right)\right] - F\left[u_{i-1}\left(\widehat{\alpha},\widehat{\beta}\right)\right]\right\}\right)^2}{n\left\{F\left[u_i\left(\widehat{\alpha},\widehat{\beta}\right)\right] - F\left[u_{i-1}\left(\widehat{\alpha},\widehat{\beta}\right)\right]\right\}}\,, \tag{19.39}$$

another asymptotically equivalent test of significance level γ $(0<\gamma<1)$ is the well-known χ^2 goodness-of-fit test, which rejects H_0 if

$$\sum_{i=1}^{m+1} \frac{\left(n_i - n\left\{F\left[u_i\left(\widehat{\alpha},\widehat{\beta}\right)\right] - F\left[u_{i-1}\left(\widehat{\alpha},\widehat{\beta}\right)\right]\right\}\right)^2}{n\left\{F\left[u_i\left(\widehat{\alpha},\widehat{\beta}\right)\right] - F\left[u_{i-1}\left(\widehat{\alpha},\widehat{\beta}\right)\right]\right\}} > \chi^2_\gamma(m-2)\,. \tag{19.40}$$

The mathematical verification of these tests can be found in *Agresti* [19.52] and *Pearson* [19.53].

Optimum Test Plans

We next discuss the optimum test plan for choosing τ_1 in a particular case. Suppose that n test units are tested under a simple SSALT which uses the same censoring time τ_2 and the same stress-change patterns with the same set of stresses $(x_1 < x_2)$ and the same stress-change times τ_1. Assume that the censoring time τ_2 is given. Suppose that the lifetimes at constant stresses x_1 and x_2 are exponential with means θ_1 and θ_2, respectively, where $\theta_i = \exp(\alpha+\beta x_i)$, $i=1,2$. Thus, $F(t) = 1-\exp(-t)$ for $t>0$. The expected Fisher information matrix $\boldsymbol{\Sigma}$ is now simplified as

$$\boldsymbol{\Sigma} = n\begin{pmatrix} A\left(\frac{\tau_1}{\theta_1}\right)^2 + B\left(\frac{\Delta\tau_2}{\theta_2}\right)^2 & x_1 A\left(\frac{\tau_1}{\theta_1}\right)^2 + x_2 B\left(\frac{\Delta\tau_2}{\theta_2}\right)^2 \\ x_1 A\left(\frac{\tau_1}{\theta_1}\right)^2 + x_2 B\left(\frac{\Delta\tau_2}{\theta_2}\right)^2 & A\left(\frac{x_1\tau_1}{\theta_1}\right)^2 + B\left(\frac{x_2\Delta\tau_2}{\theta_2}\right)^2 \end{pmatrix}, \tag{19.41}$$

where

$$A = (1-p_1)p_1$$
$$B = (1-p_1)/[\exp(\Delta\tau_2/\theta_2)-1]\,.$$

Because

$$\boldsymbol{\Sigma}^{-1} = \frac{(\theta_1\theta_2)^2}{n\cdot AB\tau_1^2\Delta\tau_2^2(x_2-x_1)^2} \times \begin{pmatrix} A\left(\frac{x_1\tau_1}{\theta_1}\right)^2 + B\left(\frac{x_2\Delta\tau_2}{\theta_2}\right)^2 & -x_1 A\left(\frac{\tau_1}{\theta_1}\right)^2 - x_2 B\left(\frac{\Delta\tau_2}{\theta_2}\right)^2 \\ -x_1 A\left(\frac{\tau_1}{\theta_1}\right)^2 - x_2 B\left(\frac{\Delta\tau_2}{\theta_2}\right)^2 & A\left(\frac{\tau_1}{\theta_1}\right)^2 + B\left(\frac{\Delta\tau_2}{\theta_2}\right)^2 \end{pmatrix}, \tag{19.42}$$

we find that the asymptotic variance of $\log\widehat{\theta}_0 = \widehat{\alpha}+\widehat{\beta}x_0$, denoted by $\mathrm{Asvar}\left(\log\widehat{\theta}_0\right)$, is given by

$$n\cdot\mathrm{Asvar}\left(\log\widehat{\theta}_0\right) = \xi^2\frac{\theta_2^2\left[\exp(\Delta\tau_2/\theta_2)-1\right]}{\exp(-\tau_1/\theta_1)\,(\Delta\tau_2)^2} + (1+\xi)^2 \times \frac{\theta_1^2\left[1-\exp(-\tau_1/\theta_1)\right]}{\exp(-\tau_1/\theta_1)\tau_1^2}\,, \tag{19.43}$$

where $\xi = \frac{x_1-x_0}{x_2-x_1}$ is the amount of stress extrapolation. Our optimum criterion is to find the optimum stress change time τ_1 $(0<\tau_1<\tau_2)$ such that the $\mathrm{Asvar}(\log\widehat{\theta}_0)$ is minimized. Because

$$\lim_{\tau_1\to 0^+}\mathrm{Asvar}\left(\log\widehat{\theta}_0\right) = \lim_{\tau_1\to\tau_2^-}\mathrm{Asvar}\left(\log\widehat{\theta}_0\right) = +\infty\,, \tag{19.44}$$

the minimum of $\mathrm{Asvar}\left(\log\widehat{\theta}_0\right)$ is attained at some τ_1 between 0 and τ_2 based on the fact that $\mathrm{Asvar}\left(\log\widehat{\theta}_0\right)$ is a continuous function of τ_1 when τ_1 is between 0 and τ_2. The minimization of $\mathrm{Asvar}\left(\log\widehat{\theta}_0\right)$ over τ_1 solves the equation

$$\frac{\partial\left[n\cdot\mathrm{Asvar}\left(\log\widehat{\theta}_0\right)\right]}{\partial\tau_1} = 0\,, \tag{19.45}$$

where

$$\begin{aligned}&\frac{\partial\left[n\cdot\mathrm{Asvar}\left(\log\widehat{\theta}_0\right)\right]}{\partial\tau_1}\\ &= \frac{\xi^2\theta_2^2}{(\Delta\tau_2)^3}\exp\left(\frac{\tau_1}{\theta_1}\right)\left\{\left[2+\left(\frac{1}{\theta_1}-\frac{1}{\theta_2}\right)\Delta\tau_2\right]\exp\left(\frac{\Delta\tau_2}{\theta_2}\right) - \left(2+\frac{\Delta\tau_2}{\theta_1}\right)\right\}\\ &\quad + \frac{(1+\xi)^2\theta_1^2}{\tau_1^3}\left[2+\left(\frac{\tau_1}{\theta_1}-2\right)\exp\left(\frac{\tau_1}{\theta_1}\right)\right]\,.\end{aligned} \tag{19.46}$$

The uniqueness of the solution to (19.45) is shown in [19.54]. In general, the solution to (19.45) is not in a closed form and therefore requires a numerical method such as the Newton–Raphson method.

An Example

We use a real data set reported in Table 17.2.1 of Chapt. 10 in *Nelson* [19.47] to demonstrate our estimation and testing procedure. The data set was obtained from a step-stress test of cryogenic cable insulation. Each specimen was first stressed for 10 min each at voltages of 5 kV, 10 kV, 15 kV, and 20 kV before it went into

step 5. Thereafter one group of specimens was stressed for 15 min at each step given in Table 19.5.

Three other groups were held for 60, 240, and 960 min at each step. Thus there were four step-stress patterns. The stress on a specimen (x) is the natural logarithm of the ratio between the voltage and the insulation thickness.

The original data were observed as exact failure times. To demonstrate our estimation process, we grouped the failure time data according to the intervals formed by consecutive stress-change times. There were five censored failure times in the data set. The grouped and censored data are summarized in Table 19.6.

Because of the different thickness for different specimens and different voltages at different steps in the testing, each specimen has its own stress pattern and censoring time. A likelihood function can be written for each specimen according to (19.25). The likelihood function for the whole sample is the product of all these individual likelihood functions in the sample. By using exact failure times instead of grouped count data, *Nelson* [19.47] fitted the Weibull model to the step-stress data and presented the MLE of model parameters on Table 17.2.2 of Chapt. 10 in *Nelson* [19.47]. The MLE estimate for the Weibull shape parameter is 0.755 97 with an asymptotic 95% confidence interval from 0.18 to 1.33. Because the confidence interval contains the value 1, there is no significant evidence against the hypothesis that the failure times of these specimens follow an exponential distribution when tested against the larger family of Weibull distributions based on the standard normal test at a significance level of 5%. We choose to base our analysis on exponential failure time in the step-stress test.

The analysis provided by Chapt. 10 in *Nelson* [19.47] assumed that the SRR is an inverse power-law model and used the stress as the ratio between the voltage and the insulation thickness. In our set up of log-linear SRR, the inverse power-law model translates into $\log\theta(x) = \alpha + \beta x$, where $\theta(x)$ is the mean of the exponential distribution at stress x, and stress x now becomes the natural logarithm of the ratio between the voltage and the insulation thickness. The design stress is at 400 V/mm, therefore, $x_0 = 5.99$. Table 19.7 presents the MLE and CI of various parameters.

To demonstrate how to find the optimum design under a simple step-stress life test, we assume that the voltage levels from step 5 (26 kV) and step 6

Table 19.5 Step-stress pattern after step 4

Step	5	6	7	8	9	10	11
kV	26.0	28.5	31.0	33.4	36.0	38.5	41.0

Table 19.6 Count data

Holding time (min)	Final step	Count (uncensored)	Censoring time	Count (censored)	Thickness (mm)
15	9	3		0	27
60	10	1	370	1	29.5
60	10	0	345	1	28
240	9	2		0	29
240	10	2	1333	1	29
240	10	1		0	30
960	5	1		0	30
960	5	0	363.9	1	30
960	6	1		0	30
960	7	3	2460.9	1	30
960	8	1		0	30
960	9	1		0	30

Table 19.7 Parameter estimates

Parameter	MLE	95% CI
α	97.5	[60.3, 134.7]
β	−12.9	[−13.6, −12.2]
$\theta(x_0) = \exp(\alpha + \beta x_0)$	6.1×10^{8}	$[2.05\times10^{-8}, 1.76\times10^{9}]$
$R_0(t) = \exp[-t/\theta(x_0)]$	$\exp(-10^{-8}t/6.1)$	$[\exp(-10^{-8}t/2.05), \exp(-10^{-9}t/1.76)]$

(28.5 kV) in the step-stress pattern are used to conduct a future simple step-stress life test. We also assume that the test uses the cable insulation with thickness equal to 30 mm (one of the four types used in the study). Therefore, the two stress levels for this simple step-stress test are $x_1 = \log(26\,000/30) = 6.765$ and $x_2 = \log(28\,500/30) = 6.856$. We still use a design stress of $x_0 = 5.99$. The amount of stress extrapolation is $\xi = 8.516$. We assume that the simple step-stress test has to stop after 1800 min (censoring time τ_2). Using the MLE of α and β obtained from the grouped and censored data in Table 19.7, we numerically solved (19.45) and found that the optimum stress-change time is after 1191.6 min (τ_1) of testing under stress x_1.

19.2.2 Analysis of a Very Simple Step-Stress Life Test with a Random Stress-Change Time

In this section we deal with a very special case of a simple SSALT that is subject to type II censoring. The traditional cumulative exposure model assumes that the stress-change time is a prespecified constant. The stress-change time in many applications, however, can be a random variable which follows a distribution. Here we consider a specific case of a simple step-stress life testing in which the stress-change time T_1 is an order statistic at the low-stress level. This type of simple SSALT occurs when experimenters want to change the stress level after a certain number of failures are observed at the low-stress level.

Throughout the section, we also denote the design stress by x_0, the i-th test stress by x_i, $i = 1, 2$, $x_1 < x_2$. We further assume that a sample of n test units begin at the low stress x_1 until the first n_1 units fail. The stress is then raised to the high stress x_2 and held until another n_2 units fail. Let $C = n(n-n_1)\binom{n-n_1-1}{n_2-1}\binom{n-1}{n_1-1}$ and $\binom{n}{k} = n!/[k!(n-k)!]$ for $0 \le k \le n$. For $0 \le i \le n_1 - 1$ and $0 \le j \le n_2 - 1$, let $\xi(n, n_1, i) = n - n_1 + i + 1$ and $\eta(n, n_1, n_2, j) = n - n_1 - n_2 + j + 1$. In addition to the assumption (A1) and (A2) made above, we further assume that only two order statistics are observed during the entire simple SSALT: one is the stress-change time, which is the n_1-th order statistic under the low stress x_1, the other is the final failure time of SSALT, which is the n_2-th order statistic under the high stress x_2. We will present the joint and marginal distributions of the two observed order statistics from the simple SSALT. We will also discuss the maximum likelihood estimates (MLE) and the method of moment estimates (MME) for the model parameters based on the joint distribution and present the exact confidence interval estimates for the model parameters based on various pivotal quantities.

Joint Distribution of Order Statistics under SSALT

Let T_1 be the stress-change time and T be the lifetime under such a simple SSALT. We further assume that the lifetime under the simple SSALT, given $T_1 = t_1$, follows the cumulative exposure model. Therefore, the conditional cumulative distribution function $G_{T|T_1}$ of T, given the stress-change time $T_1 = t_1$, is given by the classic cumulative exposure model [19.47]:

$$G_{T|T_1}(t) \begin{cases} F_1(t), & \text{for} \quad 0 \le t < t_1 \\ F_2(t - \tau_1 + s), & \text{for} \quad t_1 \le t < \infty \end{cases}, \tag{19.47}$$

where $s = t_1\theta_2/\theta_1$. The conditional probability distribution function (PDF) $g(t|t_1)$ of T, given $T_1 = t_1$, is then

$$g(t|t_1) = \begin{cases} \dfrac{1}{\theta_1} f\left(\dfrac{t_1}{\theta_1}\right), & \text{for} \quad 0 \le t < t_1 \\ \dfrac{1}{\theta_2} f\left(\dfrac{t_2 - t_1}{\theta_2} + \dfrac{t_1}{\theta_1}\right), & \text{for} \quad t_1 \le t < \infty \end{cases}. \tag{19.48}$$

The marginal probability density function (PDF) of T is given by $g(t|t_1)l(t_1)$, where $l(t_1)$ is the PDF of T_1 [19.55].

Suppose that T_2 ($T_1 < T_2$) is the final censoring observation under the simple SSALT. The observed data in such a test are the vector (T_1, T_2). Since T_1 is the n_1-th smallest observation from the distribution $F(\frac{t_1}{\theta_1})$. The probability density function of T_1 is [19.3]:

$$l(t_1) = \binom{n-1}{n_1-1} \frac{n}{\theta_1} f\left(\frac{t_1}{\theta_1}\right) F^{n_1-1}\left(\frac{t_1}{\theta_1}\right) \times \left[1 - F\left(\frac{t_1}{\theta_1}\right)\right]^{n-n_1}. \tag{19.49}$$

Given $T_1 = t_1$, the conditional cumulative exposure model implies that T_2 is the n_2-th order statistic from a sample of size $n - n_1$ with probability density function $(1/\theta_2) f[(t_2 - t_1)/\theta_2 + t_1/\theta_1]/[1 - F(t_1/\theta_1)]$, $t \ge t_1$. Thus, the conditional probability density function for T_2,

given $T_1 = t_1$, is [19.3]:

$$
\begin{aligned}
f_{T_2|T_1}(t_2) = & \binom{n-n_1-1}{n_2-1} \frac{(n-n_1)}{\theta_2} \\
& \times f\left(\frac{t_2-t_1}{\theta_2}+\frac{t_1}{\theta_1}\right) \\
& \times \left[F\left(\frac{t_2-t_1}{\theta_2}+\frac{t_1}{\theta_1}\right) - F\left(\frac{t_1}{\theta_1}\right)\right]^{n_2-1} \\
& \times R^{n-n_1-n_2}\left(\frac{t_2-t_1}{\theta_2}+\frac{t_1}{\theta_1}\right) \\
& \times R^{n_1-n}\left(\frac{t_1}{\theta_1}\right) , \qquad (19.50)
\end{aligned}
$$

where $R(t) = 1 - F(t)$. Therefore, the joint probability density for (T_1, T_2) is

$$
\begin{aligned}
f(t_1, t_2) &= f_{T_2|T_1}(t_2) l_1(t_1) \\
&= \frac{C}{\theta_1\theta_2} f\left(\frac{t_1}{\theta_1}\right) f\left(\frac{t_2-t_1}{\theta_2}+\frac{t_1}{\theta_1}\right) \\
&\quad \times F^{n_1-1}\left(\frac{t_1}{\theta_1}\right) \\
&\quad \times \left[F\left(\frac{t_2-t_1}{\theta_2}+\frac{t_1}{\theta_1}\right) - F\left(\frac{t_1}{\theta_1}\right)\right]^{n_2-1} \\
&\quad \times R^{n-n_1-n_2}\left(\frac{t_2-t_1}{\theta_2}+\frac{t_1}{\theta_1}\right) . \qquad (19.51)
\end{aligned}
$$

When the lifetime is exponential under constant stress, i. e., $f(t) = \exp(-t)$, $t > 0$,

$$
\begin{aligned}
f(t_1, t_2) = & \frac{C}{\theta_1\theta_2}\left\{\exp\left[-(n-n_1+1)\frac{t_1}{\theta_1}\right]\right\} \\
& \times \left[1-\exp\left(-\frac{t_1}{\theta_1}\right)\right]^{n_1-1} \\
& \times \left\{\exp\left[-(n-n_1-n_2+1)\frac{t_2-t_1}{\theta_2}\right]\right\} \\
& \times \left[1-\exp\left(-\frac{t_2-t_1}{\theta_2}\right)\right]^{n_2-1} . \qquad (19.52)
\end{aligned}
$$

MLE and MME

From now on we concentrate on the SRR, which assumes that $\theta(x)$ is a log-linear function of the stress x, i. e., $\ln[\theta(x)] = \alpha + \beta x$. The parameters α and β are characteristics of the products and test methods and we assume that $x > 0$ and $\beta < 0$. Notice that $\theta(x)$ is a multiple of the mean lifetime under the stress x based on the assumption (A1). In fact, if the lifetime distribution is exponential, then $\theta(x)$ is the mean lifetime under stress x. We discuss the point estimates for α, β, and $\theta_0 = \exp(\alpha + \beta x_0)$ in this section based on the method of maximum likelihood and the method of moment.

As a function of α and β, the joint density function $f(t_1, t_2)$ in (19.51) becomes the likelihood function $L(\alpha, \beta)$ based on the data vector (T_1, T_2). The maximum likelihood estimate for α and β can be obtained by solving the system of equations:

$$
\begin{cases}
\begin{aligned}
\frac{\partial \log L}{\partial \alpha} &= -2 - \frac{t_1 f'\left(\frac{t_1}{\theta_1}\right)}{\theta_1 f\left(\frac{t_1}{\theta_1}\right)} - \left(\frac{t_2-t_1}{\theta_2}+\frac{t_1}{\theta_1}\right) \\
&\quad \times \frac{f'\left(\frac{t_2-t_1}{\theta_2}+\frac{t_1}{\theta_1}\right)}{f\left(\frac{t_2-t_1}{\theta_2}+\frac{t_1}{\theta_1}\right)} - \frac{(n_1-1)t_1 f\left(\frac{t_1}{\theta_1}\right)}{\theta_1 F\left(\frac{t_1}{\theta_1}\right)} \\
&\quad - (n_2-1) \\
&\quad \times \frac{\left[\left(\frac{t_2-t_1}{\theta_2}+\frac{t_1}{\theta_1}\right) f\left(\frac{t_2-t_1}{\theta_2}+\frac{t_1}{\theta_1}\right) - \frac{t_1}{\theta_1} f\left(\frac{t_1}{\theta_1}\right)\right]}{F\left(\frac{t_2-t_1}{\theta_2}+\frac{t_1}{\theta_1}\right) - F\left(\frac{t_1}{\theta_1}\right)} \\
&\quad + (n-n_1-n_2)\left(\frac{t_2-t_1}{\theta_2}+\frac{t_1}{\theta_1}\right) \\
&\quad \times \frac{f\left(\frac{t_2-t_1}{\theta_2}+\frac{t_1}{\theta_1}\right)}{R\left(\frac{t_2-t_1}{\theta_2}+\frac{t_1}{\theta_1}\right)} = 0 \\
\frac{\partial \log L}{\partial \beta} &= -\frac{x_1 t_1 f'\left(\frac{t_1}{\theta_1}\right)}{\theta_1 f\left(\frac{t_1}{\theta_1}\right)} - \left[\frac{x_2(t_2-t_1)}{\theta_2}+\frac{x_1 t_1}{\theta_1}\right] \\
&\quad \times \frac{f'\left(\frac{t_2-t_1}{\theta_2}+\frac{t_1}{\theta_1}\right)}{f\left(\frac{t_2-t_1}{\theta_2}+\frac{t_1}{\theta_1}\right)} - \frac{(n_1-1)x_1 t_1 f\left(\frac{t_1}{\theta_1}\right)}{\theta_1 F\left(\frac{t_1}{\theta_1}\right)} \\
&\quad - (x_1+x_2) \\
&\quad - (n_2-1) \\
&\quad \times \left[\frac{\left(\frac{x_2(t_2-t_1)}{\theta_2}+\frac{x_1 t_1}{\theta_1}\right) f\left(\frac{t_2-t_1}{\theta_2}+\frac{t_1}{\theta_1}\right)}{F\left(\frac{t_2-t_1}{\theta_2}+\frac{t_1}{\theta_1}\right) - F\left(\frac{t_1}{\theta_1}\right)} - \frac{\frac{x_1 t_1}{\theta_1} f\left(\frac{t_1}{\theta_1}\right)}{F\left(\frac{t_2-t_1}{\theta_2}+\frac{t_1}{\theta_1}\right) - F\left(\frac{t_1}{\theta_1}\right)}\right] \\
&\quad + (n-n_1-n_2)\left[\frac{x_2(t_2-t_1)}{\theta_2}+\frac{x_1 t_1}{\theta_1}\right] \\
&\quad \times \frac{f\left(\frac{t_2-t_1}{\theta_2}+\frac{t_1}{\theta_1}\right)}{R\left(\frac{t_2-t_1}{\theta_2}+\frac{t_1}{\theta_1}\right)} = 0
\end{aligned}
\end{cases} , \qquad (19.53)
$$

where $f'(t) = \mathrm{d}f(t)/\mathrm{d}t$. In general, the solution of (19.53) requires a numerical method such as the Newton–Raphson method. The methods in *Seo* and *Yum* [19.50] can also be used. To find the MME of α and β, we notice that (by a change of variable)

$$
\begin{aligned}
ET_1 = & \int_0^\infty t_1 l(t_1)\,\mathrm{d}t_1 = n\binom{n-1}{n_1-1}\theta_1 \\
& \times \int_0^1 u^{n_1-1}(1-u)^{n-n_1} F^{-1}(u)\,\mathrm{d}u , \qquad (19.54)
\end{aligned}
$$

and

$$E(T_2|T_1=t_1)=\int_{t_1}^{\infty} t_2 f_{T_2|T_1}(t_2)\,\mathrm{d}t_2$$
$$=\binom{n-n_1-1}{n_2-1}\frac{(n-n_1)}{R^{n-n_1}\left(\frac{t_1}{\theta_1}\right)}$$
$$\times\int_{F(t_1/\theta_1)}^{1}\left[\theta_2 F^{-1}(v)+\left(1-\frac{\theta_2}{\theta_1}\right)t_1\right]$$
$$\times\left[v-F\left(\frac{t_1}{\theta_1}\right)\right]^{n_2-1}(1-v)^{n-n_1-n_2}\,\mathrm{d}v\,, \tag{19.55}$$

where F^{-1} is the inverse function of F. Thus, by letting $w=F(t_1/\theta_1)$,

$$E(T_2)=E_{T_1}\left[E(T_2|T_1)\right]$$
$$=C\int_0^1 w^{n_1-1}\,\mathrm{d}w$$
$$\times\int_w^1\left[\theta_2 F^{-1}(v)+(\theta_1-\theta_2)\,F^{-1}(w)\right]$$
$$\times[v-w]^{n_2-1}(1-v)^{n-n_1-n_2}\,\mathrm{d}v\,. \tag{19.56}$$

The MME of α and β can be found by solving the system of equations:

$$\begin{cases} T_1 &= n\binom{n-1}{n_1-1}\theta_1\int_0^1 u^{n_1-1}(1-u)^{n-n_1}F^{-1}(u)\,\mathrm{d}u \\ T_2 &= C\int_0^1 w^{n_1-1}\,\mathrm{d}w \\ &\quad\times\int_w^1\left[F^{-1}(v)+(\theta_1-\theta_2)\,F^{-1}(w)\right] \\ &\quad\times(v-w)^{n_2-1}(1-v)^{n-n_1-n_2}\,\mathrm{d}v\,. \end{cases} \tag{19.57}$$

When the lifetime is exponential under constant stress, i.e., $f(t)=\exp(-t)$, $t>0$, $ET_1=\theta_1\sum_{i=0}^{n_1-1}(n-i)^{-1}$ [19.3]. A direct binomial expansion in (19.52) along with the repeated use of integration by parts yields

$$E(T_2)=\theta_1\sum_{i=0}^{n_1-1}(n-i)^{-1}$$
$$+C\sum_{i=0}^{n_1-1}\sum_{j=0}^{n_2-1}(-1)^{i+j}\binom{n_1-1}{i}$$
$$\times\binom{n_2-1}{j}\frac{\theta_2}{\eta^2(n,n_1,n_2,j)\xi(n,n_1,i)}\,. \tag{19.58}$$

This also gives a closed form solution to the MME of α and β as

$$\widetilde{\beta}=\frac{1}{x_2-x_1}\left\{\ln\left[(T_2-T_1)\sum_{i=0}^{n_1-1}(n-i)^{-1}\right]\right.$$
$$-\ln\left[T_1 C\sum_{i=0}^{n_1-1}\sum_{j=0}^{n_2-1}(-1)^{i+j}\binom{n_1-1}{i}\binom{n_2-1}{j}\right.$$
$$\left.\left.\times\,\eta^{-2}(n,n_1,n_2,j)\zeta^{-1}(n,n_1,i)\right]\right\}$$
$$\widetilde{\alpha}=\ln\left(\frac{T_1}{\sum_{i=0}^{n_1-1}(n-i)^{-1}}\right)-\widetilde{\beta}x_1\,. \tag{19.59}$$

Confidence Interval Estimates of Model Parameters

Now we set up the exact confidence intervals for α, β, and $\theta_0=\exp(\alpha+\beta x_0)$ under the assumption that F is given. We first observe an important fact which will be used for the estimation of model parameters in this section. Let $S_1=\frac{T_1}{\theta_1}$, and $S_2=\frac{T_2-T_1}{\theta_2}$. The joint probability density function of (S_1, S_2) is

$$g(s_1,s_2)=Cf(s_1)f(s_2+s_1)F^{n_1-1}(s_1)$$
$$\times[F(s_2+s_1)-F(s_1)]^{n_2-1}$$
$$\times[1-F(s_2+s_1)]^{n-n_1-n_2}\,. \tag{19.60}$$

Therefore, (S_1, S_2) is a pivotal vector whose distribution does not depend on the unknown parameters θ_1 and θ_2.

We now set up a confidence interval for β. Let $S_3=\frac{S_2}{S_1}$. The marginal distribution of S_3 is given by

$$g_3(s_3)=\int_0^{\infty} g(s_1,s_1s_3)s_1\,\mathrm{d}s_1\,. \tag{19.61}$$

Since $S_3=\frac{T_2-T_1}{T_1}\exp\left[\beta(x_1-x_2)\right]$, a $100(1-\gamma)\%$ $(0<\gamma<1)$ confidence interval for β is $[\beta_1,\beta_2]$, where

$$\beta_1=\frac{1}{x_2-x_1}\ln\left(\frac{T_2-T_1}{T_1 S_{3,\gamma/2}}\right),$$
$$\beta_2=\frac{1}{x_2-x_1}\ln\left(\frac{T_2-T_1}{T_1 S_{3,1-\gamma/2}}\right), \tag{19.62}$$

and for $0<c<1$, $S_{3,c}$ is such that

$$\int_0^{S_{3,c}} g_3(s_3)\,\mathrm{d}s_3=1-c\,. \tag{19.63}$$

To set up a confidence interval for α, we let $S_4 = S_1^{x_2/x_1}$ and $S_5 = \frac{S_2}{S_4}$. The marginal distribution of S_5 is given by

$$g_5(s_5) = \frac{x_1}{x_2} \int_0^\infty g\left(s_4^{\frac{x_1}{x_2}}, s_4 s_5\right) s_4^{\frac{x_1}{x_2}} \, \mathrm{d}s_4 \,. \tag{19.64}$$

Since $S_5 = (T_2 - T_1)\, T_1^{-\frac{x_2}{x_1}} \exp\left(\frac{x_2 - x_1}{x_1}\alpha\right)$, a $100(1-\gamma)\%$ $(0 < \gamma < 1)$ confidence interval for α is $[\alpha_1, \alpha_2]$, where

$$\begin{aligned} \alpha_1 &= \frac{x_1}{x_1 - x_2}\left[\ln(T_2 - T_1) - \ln\left(T_1^{\frac{x_2}{x_1}} S_{5,1-\gamma/2}\right)\right], \\ \alpha_2 &= \frac{x_1}{x_1 - x_2}\left[\ln(T_2 - T_1) - \ln\left(T_1^{\frac{x_2}{x_1}} S_{5,\gamma/2}\right)\right], \end{aligned} \tag{19.65}$$

and, for $0 < c < 1$, $S_{5,c}$ is such that

$$\int_0^{S_{5,c}} g_5(s_5)\mathrm{d}s_5 = 1 - c \,. \tag{19.66}$$

The confidence interval for $\theta_0 = \exp(\alpha + \beta x_0)$ can also be obtained based on the distribution of a similar pivotal quantity to S_5. In fact, for $i = 1, 2$, we can always write $\theta_i = \exp(\alpha + \beta x_i) = \exp[(\alpha + \beta x_0) + \beta(x_i - x_0)]$. Therefore, by replacing the stress x_i by the transformed stress $x_i - x_0$ in the derivation of pivotal quantity S_5, we obtain another pivotal quantity:

$$\widetilde{S}_5 = (T_2 - T_1)\, T_1^{-\frac{x_2 - x_0}{x_1 - x_0}} \exp\left[\frac{x_2 - x_1}{x_1 - x_0}(\alpha + \beta x_0)\right] . \tag{19.67}$$

The distribution of $\widetilde{S}_5$ is given by the marginal density function

$$\widetilde{g}_5(s_5) = \frac{x_1 - x_0}{x_2 - x_0} \int_0^\infty g\left(s_4^{\frac{x_1 - x_0}{x_2 - x_0}}, s_4 s_5\right) s_4^{\frac{x_1 - x_0}{x_2 - x_0}} \, \mathrm{d}s_4 \,. \tag{19.68}$$

By using the distribution of pivotal quantity $\widetilde{S}_5$, we can set up a confidence interval for $\alpha + \beta x_0$ similar to the way that the confidence interval for α was set up based on the original stress x_i and the pivotal quantity S_5. Then a confidence interval for θ_0 can be obtained by the exponentiation of the confidence interval of $\alpha + \beta x_0$. More specifically, let $\xi = (x_1 - x_0)/(x_2 - x_1)$ be the amount of stress extrapolation [19.37]. A $100(1-\gamma)\%$ $(0 < \gamma < 1)$ confidence interval for $\theta_0 = \exp(\alpha + \beta x_0)$ is $[\theta_{01}, \theta_{02}]$, where

$$\begin{aligned} \theta_{01} &= \frac{T_1^{1+\xi}\, \widetilde{S}_{5,1-\gamma/2}^{\xi}}{(T_2 - T_1)^{\xi}} \,, \\ \theta_{02} &= \frac{T_1^{1+\xi}\, \widetilde{S}_{5,\gamma/2}^{\xi}}{(T_2 - T_1)^{\xi}} \,, \end{aligned} \tag{19.69}$$

Table 19.8 Percentiles of S_3 and S_5

Variable	Percentile	$n_2 = 6$	$n_2 = 8$	$n_2 = 10$
S_3	1	0.40	0.69	1.07
	2.5	0.52	0.86	1.31
	5	0.63	1.03	1.55
	10	0.80	1.27	1.88
	90	4.01	5.82	8.20
	95	5.12	7.38	10.36
	97.5	6.37	9.14	12.79
	99	8.28	11.83	16.52
S_5	1	0.84	1.32	1.91
	2.5	1.08	1.73	2.57
	5	1.42	2.25	3.32
	10	1.97	3.07	4.51
	90	27.14	40.05	56.98
	95	42.16	61.96	87.96
	97.5	62.95	92.25	130.75
	99	102.81	150.25	212.60

and, for $0 < c < 1$, $\widetilde{S}_{5,c}$ is such that

$$\int_0^{\widetilde{S}_{5,c}} \widetilde{g}_5(s_5)\,\mathrm{d}s_5 = 1 - c\,. \tag{19.70}$$

When the lifetime is exponential under constant stress, i. e., $f(t) = \exp(-t)$, $t > 0$, the marginal density functions for S_3 and S_5 are simplified. The marginal probability density function of S_3 is

$$g_3(s_3) = C \sum_{i=0}^{n_1-1} \sum_{j=0}^{n_2-1} (-1)^{i+j} \binom{n_1-1}{i} \binom{n_2-1}{j} \times \frac{1}{[\xi(n, n_1, i) + \eta(n, n_1, n_2, j)s_3]^2}\,. \tag{19.71}$$

A direct integration gives the marginal CDF of S_3 as

$$G_3(s_3) = C \sum_{i=0}^{n_1-1} \sum_{j=0}^{n_2-1} (-1)^{i+j} \binom{n_1-1}{i} \binom{n_2-1}{j} \times \frac{1}{\eta(n, n_1, n_2, j)} \left[\frac{1}{\xi(n, n_1, i)} - \frac{1}{[\xi(n, n_1, i) + \eta(n, n_1, n_2, j)s_3]} \right]. \tag{19.72}$$

The marginal density function of S_5 is

$$g_5(s_5) = C \frac{x_1}{x_2} \sum_{i=0}^{n_1-1} \sum_{j=0}^{n_2-1} (-1)^{i+j} \binom{n_1-1}{i} \binom{n_2-1}{j} \times \int_0^{\infty} \exp\left[-\xi(n, n_1, i) s_4^{\frac{x_1}{x_2}} - \eta(n, n_1, n_2, j) s_4 s_5 \right] s_4^{\frac{x_1}{x_2}}\,\mathrm{d}s_4\,. \tag{19.73}$$

A change in the order of integration and the use of integration by parts gives the marginal CDF of S_5 as

$$G_5(s_5) = C \frac{x_1}{x_2} \sum_{i=0}^{n_1-1} \sum_{j=0}^{n_2-1} (-1)^{i+j} \binom{n_1-1}{i} \binom{n_2-1}{j} \times \int_0^{\infty} \frac{1 - \exp[-\eta(n, n_1, n_2, j)s_4, s_5]}{\eta(n, n_1, n_2, j)s_4} \times \exp\left[-\zeta(n, n_1, i) s_4^{\frac{x_1}{x_2}} \right] s_4^{\frac{x_1}{x_2}}\,\mathrm{d}s_4 \tag{19.74}$$

The marginal density function and the marginal distribution function of $\widetilde{S}_5$ can be obtained by replacing the stress x_i by the transformed stress $x_i - x_0$, $i = 1, 2$, in the corresponding function of S_5. Except for trivial situations, numerical integration subroutines are typically required for the evaluation of the distribution functions associated with the pivotal quantities even when the exponential distributions are assumed. In addition, the approximation of these distributions can also be obtained through large simulations of the pivotal quantities.

Assume that a sample of 20 experimental units are placed under a simple step-stress life test. The test stress is changed from the lower stress x_1 to the higher stress x_2 after the fifth failure from the lower stress level x_1 is observed ($n_1 = 5$). The test is finished after another n_2 failures are observed at the higher stress x_2. Assume that the lifetime distribution under constant stress $x_i (i = 0, 1, 2)$ is exponential with mean parameter $\theta_i = \exp(\alpha + \beta x_i)$. For $x_0 = 0$, $x_2 = 2x_1 > 0$ and $n_2 = 6, 8, 10$, Table 19.8 presents the 1, 2.5, 5, 10, 90, 95, 97.5 and 99 percentiles for the distributions of S_3 and S_5. These percentiles are computed by numerical integration of the distribution function for S_3 and S_5. They can be used to set up appropriate confidence intervals for β, α, and θ_0.

References

19.1 W. B. Nelson: *Applied Life Data Analysis* (Wiley, New York 1982)

19.2 J. D. Kalbfleisch, R. L. Prentice: *The Statistical Analysis of Failure Time Data* (Wiley, New York 1980)

19.3 J. F. Lawless: *Statistical Models and Methods for Lifetime Data* (Wiley, New York 1982)

19.4 R. Peto, M. C. Pike, P. Armitage, N. E. Breslow, D. R. Cox, S. V. Howard, N. Mantel, K. McPherson, J. Peto, P. G. Smith: Design and analysis of randomized clinical trials requiring prolonged observation of each patient. Part II: Analysis and examples, Br. J. Cancer **35**, 1–39 (1977)

19.5 D. R. Cox: Regression models and life tables (with Discussion), J. R. Stat. Soc. B **74**, 187–200 (1972)

19.6 D. Schoenfeld: Partial residuals for the proportional hazards regression model, Biometrika **69**, 239–241 (1982)

19.7 T. M. Therneau, P. M. Grambsch, T. R. Fleming: Martingale-based residuals and survival models, Biometrika **77**, 147–160 (1990)

19.8 T. M. Therneau, P. M. Grambsch: *Modeling Survival Data: Extending the Cox Model* (Springer, Berlin Heidelberg New York 2000)

19.9 T. R. Fleming, D. P. Harrington: *Counting Processes and Survival Analysis* (Wiley, New York 1991)

19.10 P. K. Anderson, R. D. Gill: Cox's regression model for counting processes: a large sample study, Ann. Stat. **10**, 1100–1120 (1982)

19.11 P. Tiraboschi, L. A. Hansen, E. Masliah, M. Alford, L. J. Thal, J. Corey-Bloom: Impact of APOE genotype on neuropathologic and neurochemical markers of Alzheimer disease, Neurology **62(11)**, 1977–1983 (2004)

19.12 R. Weindruch, R. L. Walford: *The Retardation of Aging and Disease by Dietary Restriction* (Thomas, Springfield 1988)

19.13 H. M. Brown-Borg, K. E. Borg, C. J. Meliska, A. Bartke: Dwarf mice and the aging process, Nature **33**, 384 (1996)

19.14 R. A. Miller: Kleemeier award lecture: are there genes for aging?, J Gerontol. **54A**, B297–B307 (1999)

19.15 H. R. Warner, D. Ingram, R. A. Miller, N. L. Nadon, A. G. Richardson: Program for testing biological interventions to promote healthy aging., Mech. Aging Dev. **155**, 199–208 (2000)

19.16 S. L. George, M. M. Desu: Planning the size and duration of a clinical trial studying the time to some critical event, J. Chron. Dis. **27**, 15–24 (1974)

19.17 D. A. Schoenfeld, J. R. Richter: Nomograms for calculating the number of patients needed for aclinical trial with survival as an endpoint, Biometrics **38**, 163–170 (1982)

19.18 L. V. Rubinstein, M. H. Gail, T. J. Santner: Planning the duration of acomparative clinical trial with loss to follow-up and a period of continued observation, J. Chron. Dis. **34**, 469–479 (1981)

19.19 J. Halperin, B. W. Brown: Designing clinical trials with arbitrary specification of survival functions and for the log rank or generalized Wilcoxon test, Control. Clin. Trials **8**, 177–189 (1987)

19.20 E. Lakatos: Sample sizes for clinical trials with time-dependent rates of losses and noncompliance, Control. Clin. Trials **7**, 189–199 (1986)

19.21 D. Schoenfeld: The asymptotic properties of nonparametric tests for comparing survival distributions, Biometrika **68**, 316–318 (1981)

19.22 L. S. Freedman: Tables of the number of patients required in clinical trials using the log-rank test, Stat. Med. **1**, 121–129 (1982)

19.23 E. Lakatos: Sample sizes based on the log-rank statistic in complex clinical trials, Biometrics **44**, 229–241 (1988)

19.24 M. Wu, M. Fisher, D. DeMets: Sample sizes for long-term medical trial with time-dependent noncompliance and event rates, Control. Clin. Trials **1**, 109–121 (1980)

19.25 E. Lakatos, K. K. G. Lan: A comparison of sample size methods for the logrank statistic, Stat. Med. **11**, 179–191 (1992)

19.26 J. Crowley, D. R. Thomas: *Large Sample Theory for the Log Rank Test*, Technical Report, Vol. 415 (University of Wisconsin, Department of Statistics, 1975)

19.27 C. Xiong, Y. Yan, M. Ji: Sample sizes for comparing means of two lifetime distributions with type II censored data: application in an aging intervention study, Control. Clin. Trials **24**, 283–293 (2003)

19.28 A. Turturro, W. W. Witt, S. Lewis et al.: Growth curves and survival characteristics of the animals used in the biomarkers of aging program, J. Gerontol. Biol. Sci. Med. Sci. **A54**, B492–B501 (1999)

19.29 T. D. Pugh, T. D. Oberley, R. I. Weindruch: Dietary intervention at middle age: caloric restriction but not dehydroepiandrosterone sulfate increases lifespan and lifetime cancer incidence in mice, Cancer Res. **59**, 1642–1648 (1999)

19.30 A. S. Little: Estimation of the T-year survival rate from follow-up studies over alimited period of time, Human Biol. **24**, 87–116 (1952)

19.31 B. Epstein: Truncated life tests in the exponential case, Ann. Math. Stat. **23**, 555–564 (1954)

19.32 M. Zelen, M. C. Dannemiller: The robustness of life testing procedures derived from the exponential distribution, Technometrics **3**, 29–49 (1961)

19.33 H. Chernoff: Optimal accelerated life designs for estimation, Technometrics **4**, 381–408 (1962)

19.34 W. Q. Meeker, W. B. Nelson: Optimum accelerated life tests for Weibull and extreme value distributions and censored data, IEEE Trans. Reliab. **24**, 321–332 (1975)

19.35 W. B. Nelson, T. J. Kielpinski: Theory for optimum censored accelerated life tests for normal and lognormal life distributions, Technometrics **18**, 105–114 (1976)

19.36 W. B. Nelson: Accelerated life testing–step-stress models and data analysis, IEEE Trans. Reliab. **29**, 103–108 (1980)

19.37 R. W. Miller, W. B. Nelson: Optimum simple step-stress plans for accelerated life testing, IEEE Trans. Reliab. **32**, 59–65 (1983)

19.38 D. S. Bai, M. S. Kim, S. H. Lee: Optimum simple step-stress accelerated life tests with censoring, IEEE Trans. Reliab. **38**, 528–532 (1989)

19.39 O. I. Tyoskin, S. Y. Krivolapov: Nonparametric model for step-stress accelerated life test, IEEE Trans. Reliab. **45**, 346–350 (1996)

19.40 J. R. Dorp, T. A. Mazzuchi, G. E. Fornell, L. R. Pollock: A Bayes approach to step-stress accelerated life test, IEEE Trans. Reliab. **45**, 491–498 (1996)

19.41 C. Xiong: Inferences on a simple step-stress model with type II censored exponential data, IEEE Trans. Reliab. **47**, 142–146 (1998)

19.42 A. A. Alhadeed, S. S. Yang: Optimal simple step-stress plan for Khamis–Higgins model, IEEE Trans. Reliab. **51**, 212–215 (2002)
19.43 S. L. Teng, K. P. Yeo: A least-square approach to analyzing life–stress relationship in step-stress accelerated life tests, IEEE Trans. Reliab. **51**, 177–182 (2002)
19.44 G. K. Hobbs: *Accelerated Reliability Engineering* (Wiley, New York 2000)
19.45 N. R. Mann, R. E. Schafer, N. D. Singpurwalla: *Methods for Statistical Analysis of Reliability and Life Data* (Wiley, New York 1974)
19.46 W. Q. Meeker, L. A. Escobar: A review of recent research, current issues in accelerated testing, **61**, 147–168 (1993)
19.47 W. B. Nelson: *Accelerated Life Testing, Statistical Models, Test Plans, and Data Analysis* (Wiley, New York 1990)
19.48 S. Ehrenfeld: Some experimental design problems in attribute life testing, J. Am. Stat. Assoc. **57**, 668–679 (1962)
19.49 X. K. Yin, B. Z. Sheng: Some aspects of accelerated life testing by progressive stress, IEEE Trans. Reliab. **36**, 150–155 (1987)
19.50 S. K. Seo, B. J. Yum: Estimation methods for the mean of the exponential distribution based on grouped censored data, IEEE Trans. Reliab. **42**, 87–96 (1993)
19.51 D. R. Cox, D. V. Hinkley: *Theoretical Statistics* (Chapman Hall, London 1974)
19.52 A. Agresti: *Categorical Data Analysis* (Wiley, New York 1990)
19.53 K. Pearson: On a criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling, Philos. Mag. **50**, 157–175 (1900)
19.54 C. Xiong, M. Ji: Analysis of grouped and censored data from step-stress life testing, IEEE Trans. Reliab. **53**(1), 22–28 (2004)
19.55 C. Xiong: Step-stress life-testing with random stress-change times for exponential data, IEEE Trans. Reliab. **48**, 141–148 (1999)