

Simple step-stress models with a cure fraction

Nandini Kannan^a and Debasis Kundu^b

^aNational Science Foundation

^bIndian Institute of Technology Kanpur

Abstract. In this article, we consider models for time-to-event data obtained from experiments in which stress levels are altered at intermediate stages during the observation period. These experiments, known as step-stress tests, belong to the larger class of accelerated tests used extensively in the reliability literature. The analysis of data from step-stress tests largely relies on the popular cumulative exposure model. However, despite its simple form, the utility of the model is limited, as it is assumed that the hazard function of the underlying distribution is discontinuous at the points at which the stress levels are changed, which may not be very reasonable. Due to this deficiency, Kannan et al. (*Journal of Applied Statistics* **37** (2010b) 1625–1636) introduced the cumulative risk model, where the hazard function is continuous. In this paper, we propose a class of parametric models based on the cumulative risk model assuming the underlying population contains long-term survivors or ‘cured’ fraction. An EM algorithm to compute the maximum likelihood estimators of the unknown parameters is proposed. This research is motivated by a study on altitude decompression sickness. The performance of different parametric models will be evaluated using data from this study.

1 Introduction

In many applications in the physical, environmental, and biomedical sciences, the experimental conditions may change during the exposure period. This research is motivated by a study on altitude decompression sickness (DCS), a condition caused by a decrease in pressure and lower oxygen levels at high altitudes. Mountaineers and travelers to high mountainous regions frequently experience symptoms that include nausea, headaches, dizziness and fatigue. Decompression sickness is also observed in deep-sea divers and mine workers. The change in atmospheric pressure causes nitrogen to come out of solution and form bubbles in tissues and the bloodstream. Symptoms of DCS range from mild joint pain and swelling to dizziness, and in extreme cases death. To avoid/ decrease symptoms of decompression sickness, mountaineers are advised to acclimatize to the altitude conditions and perform gradual staged ascents. Deep-sea divers are provided with tables that specify the length of time at each stage in the ascent from depth (decompression stops) to avoid the “bends”. This staged ascent can delay or eliminate the onset of symptoms by allowing nitrogen levels to gradually dissipate. In the staged ascent, the experimental conditions (in this case, altitude) are changed throughout the exposure period.

Greven et al. (2004) described an experiment designed to examine the effects of water contamination on fish, in particular on their swimming performance. The underlying hypothesis was that fish exposed to toxins may exhibit a lower threshold for fatigue. Fatigue was induced by increasing the stress, in this case the water velocity, at fixed time points during the experiment.

The two examples described above are known as “step-stress tests” in the reliability literature. Step-stress tests belong to the larger class of accelerated tests that have been extensively

Key words and phrases. Cumulative exposure model, step-stress tests, cured fraction, maximum likelihood estimation, EM algorithm.

Received September 2017; accepted July 2018.

used in life-testing experiments. As products become more reliable, standard life-testing experiments often fail to provide adequate data on failure times, resulting in poor estimates of reliability. Accelerated testing allows the researcher to overcome this problem by subjecting the experimental units or items to higher than normal levels of stress. The increase in stress reduces the time to failure by accelerating the degradation or deterioration of the item. Using a model that relates the stress level to the failure distribution, properties of the underlying failure distribution under normal operating conditions may be obtained. In a standard step-stress experiment, all individuals or items are subject to an initial stress level. The stress is gradually increased at pre-specified times during the experiment. The stress factor may refer to the dose of a drug, incline/speed of a treadmill, temperature, voltage, or pressure.

The most commonly used model for step-stress experiments is the cumulative exposure model (CEM) introduced by Sedyakin (1966). The CEM and its variants have been studied by a number of authors including Bai, Kim and Lee (1989), Balakrishnan (2009), Balakrishnan et al. (2007), Gouno and Balakrishnan (2001), Nelson (1980, 1990), Xiong (1998), Xiong and Milliken (1999), Khamis and Higgins (1998), Ismail (2016), Kateri and Kamps (2015) and Sha and Pan (2014). The formulation of the CEM provides a relationship between the lifetime distribution of an experimental unit at one stress level and the lifetime distribution of the unit at the other stress level. As a consequence of this assumption, the underlying hazard function is discontinuous at the time point at which the stress level changes. This implies that the effect of a change in stress level is instantaneous, which may not be realistic in practice.

Kannan, Kundu and Balakrishnan (2010a) proposed a new simple step-stress model assuming that there is a latency or lag period δ before the effect of the change in stress level is observed. The cumulative risk model (CRM), proposed by the authors, assumes that the hazard function for the step stress experiment is continuous. The CEM and the time transformed Khamis and Higgins (1998) model may be obtained as limiting cases of the CRM. Kannan, Kundu and Balakrishnan (2010a) developed maximum likelihood and least squares estimators for the parameters of the CRM assuming exponential lifetime distributions at the two stress levels. Beltrami (2011) extended the CRM to account for competing causes of failure assuming exponential and Weibull lifetime distributions.

In recent years, the development of new drugs and treatment regimens has resulted in patients living longer with certain diseases. In some cases, patients show no recurrence of a disease, i.e. are viewed as being permanently cured. These patients are referred to as “cured” or long term survivors, while the remaining patients are referred to as “susceptible”. Traditional parametric survival models such as Weibull or Gamma do not account for the probability of cure. Although subtle, it is important to distinguish between the concepts of censoring and cure. Note that censoring refers to a subject who does not fail within the monitoring time window of a particular item, where as cure refers to one who will not fail within any reasonable time window. Clearly, the later is an abstraction, as we never observe a cure due to a finite monitoring time. Still estimating the probability of such an outcome is a important issue, see for example Tsodikov (1998). In the context of reliability or life-testing applications, cure refers to items that are resistant and not affected by the stress factor. It can happen even in a step-stress experiment also. For example, in the DCS study it has been observed, see for example Kannan, Raychaudhury and Pilmanis (1998), that even in the increased stress level some of the experimental units are immune to the stress factors and they do not fail at the higher stress level. They can be treated as the long term survivors.

The goal of this article is to develop a cumulative risk model assuming the underlying population includes both ‘cured’ and susceptible individuals. We consider different parametric distributions to model the lifetimes of susceptible individuals. In Section 2, we define the generalized cumulative risk model and discuss three special cases. In Section 3, using the standard mixture model formulation, the generalized CRM is adapted to include a cured

fraction. The probability of being cured is modeled as a function of covariates using a logistic link. Section 4 describes an EM algorithm for obtaining the MLEs of the parameters and associated likelihood ratio tests. In Section 5, we fit different parametric models to the data on altitude decompression sickness and evaluate the goodness of fit of these models. Conclusions and future research directions are discussed in Section 6.

2 Generalized cumulative risk model

Consider a life-testing experiment wherein the stress is changed only once during the exposure duration. We assume that all n individuals or items under study are exposed to an initial stress level x_1 . The items are continuously monitored, and at some pre-specified time τ_1 , the stress level is changed (increased) to x_2 . We assume that there is a latency period δ before the effects of the increased stress are observed.

We assume that the hazard function is continuous for all $t > 0$ and is increasing in the interval $[\tau_1, \tau_2]$, where $\tau_2 = \tau_1 + \delta$. The piecewise hazard function for the Generalized Cumulative Risk Model has the following form;

$$h_0(t) = \begin{cases} h_{01}(t) & \text{if } 0 < t \leq \tau_1, \\ a + bt & \text{if } \tau_1 < t < \tau_2, \\ h_{03}(t) & \text{if } \tau_2 \leq t < \infty. \end{cases} \quad (2.1)$$

Here $h_{01}(t) > 0$ and $h_{03}(t) > 0$ are continuous functions for $t > 0$, and

$$\int_0^\infty h_0(t) dt = \infty. \quad (2.2)$$

The parameters a and b are such that the hazard function $h_0(t)$ is a continuous function for $t > 0$. Therefore, a and b satisfy the following conditions:

$$a + b\tau_1 = h_{01}(\tau_1), \quad (2.3)$$

$$a + b\tau_2 = h_{03}(\tau_2). \quad (2.4)$$

Note that, although for mathematical tractability, we assume the hazard function to be linear in the interval $[\tau_1, \tau_2]$, other smooth functions may be considered.

The generalized cumulative risk model, as defined through the hazard function (2.1), is an extension of the cumulative risk model proposed by Kannan, Kundu and Balakrishnan (2010a). The authors considered the case when $h_{01} = \lambda_1$ and $h_{03}(t) = \lambda_2$, that is, exponential distributions with different scale parameters.

When τ_2 decreases to τ_1 , then (2.1) takes the form

$$h_0(t) = \begin{cases} h_{01}(t) & \text{if } 0 < t \leq \tau_1, \\ h_{03}(t) & \text{if } \tau_1 < t < \infty. \end{cases} \quad (2.5)$$

Equation (2.5) is the hazard function for the cumulative exposure model in the case of exponential hazards (see Balakrishnan (2009)), and the hazard function of the model proposed by Khamis and Higgins (1998) in the case of Weibull hazards, when the two shape parameters are equal. It can also be observed as an generalization of the tampered failure rate (TFR) model of Bhattacharyya and Soejoeti (1989). In the TFR model, $h_{01}(t)/h_{03}(t)$ is a constant.

From (2.1), we obtain the cumulative hazard function as

$$H_0(t) = \begin{cases} H_{01}(t) & \text{if } 0 < t \leq \tau_1, \\ H_{02}(t) & \text{if } \tau_1 < t < \tau_2, \\ H_{03}(t) & \text{if } \tau_2 \leq t < \infty, \end{cases} \quad (2.6)$$

where

$$\begin{aligned} H_{01}(t) &= \int_0^t h_{01}(u) du, \\ H_{02}(t) &= H_{01}(\tau_1) + a(t - \tau_1) + \frac{b}{2}(t^2 - \tau_1^2), \\ H_{03}(t) &= H_{02}(\tau_2) + \int_{\tau_2}^t h_{03}(u) du. \end{aligned}$$

Therefore, the survival function is given by

$$S_0(t) = \begin{cases} e^{-H_{01}(t)} & \text{if } 0 < t \leq \tau_1, \\ e^{-H_{02}(t)} & \text{if } \tau_1 < t < \tau_2, \\ e^{-H_{03}(t)} & \text{if } \tau_2 \leq t < \infty. \end{cases} \quad (2.7)$$

Note that due to the condition (2.2), (2.7) is a proper survival function. The associated probability density function (PDF) is

$$f_0(t) = \begin{cases} f_{01}(t) & \text{if } 0 < t \leq \tau_1, \\ f_{02}(t) & \text{if } \tau_1 < t < \tau_2, \\ f_{03}(t) & \text{if } \tau_2 \leq t < \infty, \end{cases} \quad (2.8)$$

where

$$\begin{aligned} f_{01}(t) &= h_{01}(t)e^{-H_{01}(t)}, & f_{02}(t) &= (a + bt)e^{-H_{02}(t)}, \\ f_{03}(t) &= h_{03}(t)e^{-H_{03}(t)}. \end{aligned} \quad (2.9)$$

2.1 Special cases

In this subsection, we describe three special cases of the generalized cumulative risk model.

Weibull Model. The Weibull distribution is one of the most widely used parametric distribution in the survival analysis and reliability literature to analyze lifetime data. The PDF of a two-parameter Weibull distribution has the following form:

$$f(t) = \alpha \lambda t^{\alpha-1} e^{-\lambda t^\alpha}; \quad x > 0. \quad (2.10)$$

Here $\alpha > 0$ and $\lambda > 0$ are the shape and scale parameters, respectively.

Therefore, we have

$$h_{01}(t) = \alpha_1 \lambda_1 t^{\alpha_1-1}, \quad h_{03}(t) = \alpha_2 \lambda_2 t^{\alpha_2-1} \quad (2.11)$$

and

$$\begin{aligned} H_{01}(t) &= \lambda_1 t^{\alpha_1}, \\ H_{02}(t) &= \lambda_1 \tau_1^{\alpha_1} + a(t - \tau_1) + \frac{b}{2}(t^2 - \tau_1^2), \\ H_{03}(t) &= \lambda_1 \tau_1^{\alpha_1} + a(\tau_2 - \tau_1) + \frac{b}{2}(\tau_2^2 - \tau_1^2) + (\lambda_2 \tau_2^{\alpha_2} - \lambda_2 t^{\alpha_2}). \end{aligned}$$

Linear Failure Rate Distribution. The probability density function of a two-parameter linear failure rate distribution for $\alpha > 0$ and $\lambda > 0$, is given by

$$f(t) = (\alpha + \lambda t) e^{-\alpha t - \frac{\lambda t^2}{2}}. \quad (2.12)$$

In this case,

$$h_{01}(t) = \alpha_1 + \lambda_1 t, \quad h_{03}(t) = \alpha_2 + \lambda_2 t \quad (2.13)$$

and

$$\begin{aligned} H_{01}(t) &= \alpha_1 t + \frac{\lambda_1 t^2}{2}, \\ H_{02}(t) &= \alpha_1 \tau_1 + \frac{\lambda_1 \tau_1^2}{2} + a(t - \tau_1) + \frac{b}{2}(t^2 - \tau_1^2), \\ H_{03}(t) &= \alpha_1 \tau_1 + \frac{\lambda_1 \tau_1^2}{2} + a(\tau_2 - \tau_1) + \frac{b}{2}(\tau_2^2 - \tau_1^2) + \alpha_2(t - \tau_2) + \frac{\lambda_2}{2}(t^2 - \tau_2^2). \end{aligned}$$

Generalized Exponential Distribution. This lifetime model was introduced by [Gupta and Kundu \(1999\)](#) as an alternative to the gamma and Weibull distributions. The two-parameter generalized exponential distribution has the following PDF:

$$f(t) = \alpha \lambda e^{-\lambda t} (1 - e^{-\lambda t})^{\alpha-1}; \quad t > 0, \quad (2.14)$$

where $\alpha > 0$ and $\lambda > 0$ are the shape and scale parameters, respectively.

In this case

$$\begin{aligned} h_{01}(t) &= \frac{\alpha_1 \lambda_1 e^{-\lambda_1 t} (1 - e^{-\lambda_1 t})^{\alpha_1-1}}{1 - (1 - e^{-\lambda_1 t})^{\alpha_1}}, \\ h_{03}(t) &= \frac{\alpha_2 \lambda_2 e^{-\lambda_2 t} (1 - e^{-\lambda_2 t})^{\alpha_2-1}}{1 - (1 - e^{-\lambda_2 t})^{\alpha_2}} \end{aligned} \quad (2.15)$$

and

$$\begin{aligned} H_{01}(t) &= -\ln(1 - (1 - e^{-\lambda_1 t})^{\alpha_1}), \\ H_{02}(t) &= -\ln(1 - (1 - e^{-\lambda_1 \tau_1})^{\alpha_1}) + a(t - \tau_1) + \frac{b}{2}(t^2 - \tau_1^2), \\ H_{03}(t) &= -\ln(1 - (1 - e^{-\lambda_1 \tau_1})^{\alpha_1}) + a(\tau_2 - \tau_1) + \frac{b}{2}(\tau_2^2 - \tau_1^2) \\ &\quad - \ln(1 - (1 - e^{-\lambda_2 t})^{\alpha_2}) - \ln(1 - (1 - e^{-\lambda_2 \tau_2})^{\alpha_2}). \end{aligned}$$

For all three cases, let $\Theta = (\alpha_1, \alpha_2, \lambda_1, \lambda_2)$ denote the parameter vector.

3 CRM with long-term survivors

Extensive work has been done in formulating parametric and non-parametric survival models incorporating a cured fraction, a non-zero tail probability of the survival function, starting with the seminal paper by [Boag \(1949\)](#), who proposed a two component mixture model to analyze breast cancer data. The population of interest may be regarded as a mixture of cured and susceptible individuals/ items. In this model, the survival function for the entire population can be written as

$$P(T > t) = S_p(t) = \pi + (1 - \pi)S(t),$$

where $\pi = S_p(+\infty)$ is the ‘cured fraction’, and $S(t)$ with $S(+\infty) = 0$ is the proper survival function for the ‘non-cured’ or the susceptible population. Since then there has been extensive research on the cure rate model, primarily in the survival analysis literature. Interested readers

may refer to the classical book by [Maller and Zhou \(1996\)](#) or [Yu et al. \(2004\)](#), [Cancho and Bolafarine \(2001\)](#), [Chen, Ibrahim and Sinha \(1999\)](#), [Gamel, Mclean and Rosenberg \(1999\)](#), [Kannan et al. \(2010b\)](#), and the references cited therein.

The cumulative risk model can be extended to incorporate the cured fraction that represents individuals/items resistant to the stress factors. We assume that the susceptible population has a survival function $S_0(\cdot)$ as defined in (2.7). The probability that an individual is cured is assumed to depend on a set of s covariates z_1, \dots, z_s . Let $\mathbf{z} = (1, z_1, \dots, z_s)$. For each individual under study with covariate vector \mathbf{z} , we define an indicator random variable $\Delta(\mathbf{z})$, with $\Delta(\mathbf{z}) = 1$, when the subject is susceptible, and $\Delta(\mathbf{z}) = 0$, when the subject is cured. It is assumed that

$$P(\Delta(\mathbf{z}) = 0) = p(\boldsymbol{\beta}, \mathbf{z}) = \frac{e^{\boldsymbol{\beta}'\mathbf{z}}}{1 + e^{\boldsymbol{\beta}'\mathbf{z}}} \quad \text{and}$$

$$P(\Delta(\mathbf{z}) = 1) = 1 - p(\boldsymbol{\beta}, \mathbf{z}) = \frac{1}{1 + e^{\boldsymbol{\beta}'\mathbf{z}}}.$$

Here $\boldsymbol{\beta}$ is an $s + 1$ dimensional vector of coefficients and β_0 represents the intercept term. Let $S(t; \boldsymbol{\Theta}, \boldsymbol{\beta}, \mathbf{z})$ denote the survival function of an individual with covariate vector \mathbf{z} . We have

$$S(t; \boldsymbol{\Theta}, \boldsymbol{\beta}, \mathbf{z}) = p(\boldsymbol{\beta}, \mathbf{z}) + (1 - p(\boldsymbol{\beta}, \mathbf{z}))S_0(t; \boldsymbol{\Theta}), \quad (3.1)$$

where $S_0(t; \boldsymbol{\Theta})$ has the form of the cumulative risk model defined in (2.7).

It may be noted that we have linked the covariates only with the cure fraction and not with the survival function $S_0(t; \boldsymbol{\Theta})$. It may be argued that linking covariates with the survival function as well as the cure fraction would provide a more general model. Although, it may be true but it raises some questions regarding choosing the proper covariates for two different purposes, and their identifiability issues, see also [Kannan et al. \(2010b\)](#). Due to which it has been avoided in this study.

4 Statistical inference

4.1 Maximum likelihood estimation

In this section, we propose an EM algorithm to compute the maximum likelihood estimators (MLEs) of the unknown parameters, namely, $\boldsymbol{\Theta}$ and $\boldsymbol{\beta}$. At the initial stage, we assume that δ , or equivalently, $\tau_2 = \tau_1 + \delta$ is known. We will then describe a procedure to estimate δ . It should be mentioned that for fixed δ , the consistency and asymptotic normality of the MLEs can be obtained along the same line as [Ghitany and Maller \(1992\)](#).

The observed failure times from the step-stress experiment are denoted by

$$0 < t_1 < \dots < t_{n_1} < \tau_1 < t_{n_1+1} < \dots < t_{n_1+n_2} < \tau_2 < t_{n_1+n_2+1} < \dots \\ \leq t_{n_1+n_2+n_3}. \quad (4.1)$$

Here $m = n_1 + n_2 + n_3$ represents the total number of failures observed. The remaining observations

$$t_{n_1+n_2+n_3+1} < \dots < t_{n_1+n_2+n_3+n_4}, \quad (4.2)$$

where $n_4 = n - m$, are censored. With each failure/censor time t_i , we have an associated covariate vector \mathbf{z}_i .

Let $I_1 = \{1, \dots, n_1\}$, $I_2 = \{n_1 + 1, \dots, n_1 + n_2\}$, $I_3 = \{n_1 + n_2 + 1, \dots, n_1 + n_2 + n_3\}$ and $I_4 = \{n_1 + n_2 + n_3 + 1, \dots, n_1 + n_2 + n_3 + n_4\}$. Here I_4 denotes the set of censored

observations. Based on the observations $\{(t_i, z_i); i = 1, \dots, n\}$, the likelihood function can be written as

$$L(\Theta, \beta) = \prod_{k=1}^3 \prod_{i \in I_k} \{(1 - p(\beta, z_i)) f_{0k}(t_i; \Theta)\} \\ \times \prod_{i \in I_4} \{p(\beta, z_i) + (1 - p(\beta, z_i)) S_0(t_i; \Theta)\}. \quad (4.3)$$

Hence, the log-likelihood function can be written as

$$l(\Theta, \beta) = \ln L(\Theta, \beta) = l_1(\Theta, \beta) + l_2(\Theta, \beta), \quad (4.4)$$

where

$$l_1(\Theta, \beta) = \sum_{i \in I_1 \cup I_2 \cup I_3} \ln(1 - p(\beta, z_i)) + \sum_{i \in I_1} \ln f_{01}(t_i; \Theta) + \sum_{i \in I_2} \ln f_{02}(t_i; \Theta) \\ + \sum_{i \in I_3} \ln f_{03}(t_i; \Theta),$$

and

$$l_2(\Theta, \beta) = \sum_{i \in I_4} \ln\{p(\beta, z_i) + (1 - p(\beta, z_i)) S_0(t_i; \Theta)\}.$$

The maximum likelihood estimators of Θ and β can be obtained by maximizing the log-likelihood function (4.4), and involves a $(4 + (s + 1))$ -dimensional optimization problem, where s denotes the number of covariates.

The MLEs may also be obtained by treating the problem as a missing data problem and using the EM algorithm. For an individual with covariate vector z in $I_1 \cup I_2 \cup I_3$, the associated random variable $\Delta(z)$ takes the value 1. For an individual in I_4 , $\Delta(z)$ is unobserved. Therefore, for these n_4 observations, the associated $\Delta(z)$ values are treated as missing.

At the k -th stage of the EM algorithm, let $\Theta^{(k)}$ and $\beta^{(k)}$ denote the estimates of the parameters Θ and β , respectively. We then compute the pseudo log-likelihood function based on the missing observations ('E'-step). For a censored time t , two partially complete 'pseudo observation' of the form $(t, w_1(t|\Theta^{(k)}, \beta^{(k)}, z))$ and $(t, w_2(t|\Theta^{(k)}, \beta^{(k)}, z))$ are constructed. Here, $w_1(t|\Theta^{(k)}, \beta^{(k)}, z)$ and $w_2(t|\Theta^{(k)}, \beta^{(k)}, z)$ denote the conditional probabilities that the individual with covariate z belongs to the cure and susceptible group, respectively, given that the individual has survived till time t . We have

$$w_1(t|\Theta^{(k)}, \beta^{(k)}, z) = P(\Delta(z) = 0 | T > t, \Theta^{(k)})$$

and

$$w_2(t|\Theta^{(k)}, \beta^{(k)}, z) = P(\Delta(z) = 1 | T > t, \Theta^{(k)}).$$

Note that

$$w_1(t|\Theta^{(k)}, \beta^{(k)}, z) = P(\Delta(z) = 0 | T > t, \Theta^{(k)}, \beta^{(k)}) \\ = \frac{P(T > t | \Delta(z) = 0, \Theta^{(k)}, \beta^{(k)}) \times P(\Delta(z) = 0 | \beta^{(k)})}{P(T > t | \Theta^{(k)}, \beta^{(k)})} \\ = \frac{p(\beta^{(k)}, z)}{p(\beta^{(k)}, z) + (1 - p(\beta^{(k)}, z)) S_0(t | \Theta^{(k)})},$$

and

$$\begin{aligned}
 w_2(t|\Theta^{(k)}, \beta^{(k)}, z) &= P(\Delta(z) = 1|T > t, \Theta^{(k)}, \beta^{(k)}) \\
 &= \frac{P(T > t|\Delta(z) = 1, \Theta^{(k)}, \beta^{(k)}) \times P(\Delta(z) = 1|\beta^{(k)})}{P(T > t|\Theta^{(k)}, \beta^{(k)})} \\
 &= \frac{(1 - p(\beta^{(k)}, z))S_0(t|\Theta^{(k)})}{p(\beta^{(k)}, z) + (1 - p(\beta^{(k)}, z))S_0(t|\Theta^{(k)})}.
 \end{aligned}$$

Therefore, the ‘pseudo log-likelihood’ function based on the missing observation is given by

$$\begin{aligned}
 l_{\text{pseudo}}(\Theta, \beta|\Theta^{(k)}, \beta^{(k)}) &= l_1(\Theta, \beta) + \sum_{i \in I_4} \{w_1(t_i|\Theta^{(k)}, \beta^{(k)}) \ln p(\beta, z_i) \\
 &\quad + w_2(t_i|\Theta^{(k)}, \beta^{(k)}) \ln[(1 - p(\beta, z_i))S_0(t_i|\Theta)]\} \\
 &= \sum_{i \in I_1 \cup I_2 \cup I_3} \ln(1 - p(\beta, z_i)) \\
 &\quad + \sum_{i \in I_4} w_1(t_i|\Theta^{(k)}, \beta^{(k)}) \ln p(\beta, z_i) \\
 &\quad + \sum_{i \in I_4} w_2(t_i|\Theta^{(k)}, \beta^{(k)}) \ln(1 - p(\beta, z_i)) \\
 &\quad + \sum_{i \in I_1} \ln f_{01}(t_i|\Theta) + \sum_{i \in I_2} \ln f_{02}(t_i|\Theta) \\
 &\quad + \sum_{i \in I_3} \ln f_{03}(t_i|\Theta) \\
 &\quad + \sum_{i \in I_4} w_2(t_i|\Theta^{(k)}, \beta^{(k)}) \ln S_0(t_i|\Theta) \\
 &= g_1(\beta|\Theta^{(k)}, \beta^{(k)}) + g_2(\Theta|\Theta^{(k)}) \quad (\text{say}), \tag{4.5}
 \end{aligned}$$

where

$$\begin{aligned}
 g_1(\beta|\Theta^{(k)}, \beta^{(k)}) &= \sum_{i \in I_1 \cup I_2 \cup I_3} \ln(1 - p(\beta, z_i)) \\
 &\quad + \sum_{i \in I_4} w_1(t_i|\Theta^{(k)}, \beta^{(k)}) \ln p(\beta, z_i) \\
 &\quad + \sum_{i \in I_4} w_2(t_i|\Theta^{(k)}, \beta^{(k)}) \ln(1 - p(\beta, z_i)), \tag{4.6}
 \end{aligned}$$

and

$$g_2(\Theta|\Theta^{(k)}) = g_3(\Theta) + \sum_{i \in I_4} w_2(t_i|\Theta^{(k)}, \beta^{(k)}) \ln S_0(t_i|\Theta). \tag{4.7}$$

Here

$$g_3(\Theta) = \sum_{i \in I_1} \ln f_{01}(t_i|\Theta) + \sum_{i \in I_2} \ln f_{02}(t_i|\Theta) + \sum_{i \in I_3} \ln f_{03}(t_i|\Theta). \tag{4.8}$$

The ‘M’-step of the EM algorithm involves maximization of the function $l_{\text{pseudo}}(\Theta, \beta|\Theta^{(k)}, \beta^{(k)})$ with respect to the unknown parameters.

Assuming a model with no covariates, and denoting $p(\boldsymbol{\beta}, \mathbf{z}) = p$, the function (4.6) reduces to

$$g_1(p|\boldsymbol{\Theta}^{(k)}, p^{(k)}) = m \ln(1 - p) + \ln p \sum_{i \in I_4} w_1(t_i|\boldsymbol{\Theta}^{(k)}, p^{(k)}) + \ln(1 - p) \sum_{i \in I_4} w_2(t_i|\boldsymbol{\Theta}^{(k)}, p^{(k)}), \quad (4.9)$$

where

$$w_1(t_i|\boldsymbol{\Theta}^{(k)}, p^{(k)}) = \frac{p^{(k)}}{p^{(k)} + (1 - p^{(k)})S_0(t_i|\boldsymbol{\Theta}^{(k)})} \quad (4.10)$$

and

$$w_2(t_i|\boldsymbol{\Theta}^{(k)}, p^{(k)}) = \frac{(1 - p^{(k)})S_0(t_i|\boldsymbol{\Theta}^{(k)})}{p^{(k)} + (1 - p^{(k)})S_0(t_i|\boldsymbol{\Theta}^{(k)})}. \quad (4.11)$$

Therefore,

$$p^{(k+1)} = \frac{\sum_{i \in I_4} w_1(t_i|\boldsymbol{\Theta}^{(k)}, p^{(k)})}{n}. \quad (4.12)$$

If the model includes covariates, a Newton–Raphson type method or other optimization techniques may be used to maximize the functions $g_1(\boldsymbol{\beta}|\boldsymbol{\Theta}^{(k)}, \boldsymbol{\beta}^{(k)})$ and $g_2(\boldsymbol{\Theta}|\boldsymbol{\Theta}^{(k)})$.

So far we have assumed that δ is known. Since in practice δ is unknown we estimate δ by maximizing the profile likelihood function with respect to δ . Since it cannot be done analytically we have suggested to use grid search method for this purpose. For each δ , we compute the MLEs of the unknown parameters as suggested above and obtain the corresponding likelihood value, and then choose that δ for which the likelihood value is maximum. The detailed implementation procedure will be explained in the Data Analysis section.

4.2 Testing of hypothesis

In this section, we describe three hypotheses of interest and derive the corresponding likelihood ratio tests. In all the cases the exact tests are not available, hence, we rely on the asymptotic tests. Since the sample size is not very small, it is well known that the asymptotic tests work reasonably well.

Problem 1.

$$\mathcal{H}_{01} : \alpha_1 = \alpha_2 \quad \text{vs.} \quad \mathcal{H}_{11} : \alpha_1 \neq \alpha_2. \quad (4.13)$$

The hypothesis tests the equality of the shape parameters of the corresponding lifetime distributions at the two different stress levels. Under the null hypothesis, the maximization of $g_2(\boldsymbol{\Theta}|\boldsymbol{\Theta}^{(k)})$ involves a three-dimensional optimization. Letting l_{01} and l_{11} denote the maximized log-likelihood values under the null and alternative hypothesis, respectively, we have

$$-2(l_{01} - l_{11}) \sim \chi_1^2. \quad (4.14)$$

Problem 2.

$$\mathcal{H}_{02} : \alpha_1 = \alpha_2 = 1 \quad \text{vs.} \quad \mathcal{H}_{12} : \alpha_1 = \alpha_2 \neq 1. \quad (4.15)$$

The null hypothesis states that the lifetime distributions at the two different stress levels are exponential. Under the null hypothesis, the maximization of $g_2(\boldsymbol{\Theta}|\boldsymbol{\Theta}^{(k)})$ involves a two-dimensional optimization. Letting l_{02} and l_{12} ($= l_{11}$) denote the maximized log-likelihood values under the null and alternative hypothesis, respectively, we have

$$-2(l_{02} - l_{12}) \sim \chi_1^2. \quad (4.16)$$

Problem 3.

$$\mathcal{H}_{03} : \beta_0 = \beta_1 = \cdots = \beta_s = 0 \quad \text{vs.} \quad \mathcal{H}_{13} : \text{At least one of them is not 0.} \quad (4.17)$$

This is a test of the significance of the covariates. In the EM algorithm, the ‘M’-step that involves the maximization of $g_1(p|\Theta^{(k)}, p^{(k)})$ can be performed explicitly. From the likelihood ratio test, it follows that

$$-2(l_{03} - l_{13}) \sim \chi_s^2, \quad (4.18)$$

where l_{03} and l_{13} denote the maximized log-likelihood values under the null and alternative hypothesis, respectively.

Problem 4.

$$\mathcal{H}_{04} : p = 0 \quad \text{vs.} \quad \mathcal{H}_{14} : p > 0. \quad (4.19)$$

This test assesses whether or not the underlying population has a cure fraction. Since 0 is a boundary point of p , using Theorem 3 of [Self and Liang \(1987\)](#), it follows that

$$-2(l_{04} - l_{14}) \sim \frac{1}{2} + \frac{1}{2}\chi_1^2. \quad (4.20)$$

Here l_{04} and l_{14} denote the maximized log-likelihood values under the null and alternative hypothesis, respectively. It may be mentioned that in this case $l_{14} = l_{03}$.

5 Data analysis and simulation experiments**5.1 Data analysis**

The motivation for this research comes from a study on altitude decompression sickness conducted by the Air Force Research Laboratory (AFRL). In the study, human subjects were exposed to simulated altitude in a hypobaric chamber. Subjects were exposed to different altitudes, varying denitrogenation times, exposure duration, different breathing gas mixtures during the exposure, and levels of exercise as determined by the flight protocol. During the exposure, subjects were constantly monitored for symptoms of DCS. If a subject exhibited any of the symptoms associated with DCS, the experiment was terminated and the subject repressurized to ground level. Subjects who did not exhibit any symptoms during the exposure period are assumed to be censored.

During exposure to altitude, the gas exchange processes in the tissues are unable to expel the excess nitrogen causing supersaturation. These gases which come out of solution when tissues are sufficiently supersaturated collect as bubbles in the tissue. To delay the onset of bubbles, researchers suggest either a staged decompression to altitude or prebreathing pure oxygen prior to exposure to allow nitrogen diffusion through desaturation.

The data that we consider here comes from one of the flight profiles that involved a staged ascent. 40 subjects were placed in the hypobaric chamber and exposed to an initial altitude of 18,000 feet. After four hours (240 minutes), the altitude was increased to 35,000 feet and the experiment continued for another three hours. In the first four hours of the exposure, prior to the increase in altitude, 7 subjects reported DCS symptoms. At the end of the study, a total of 31 subjects reported DCS. The remaining 9 subjects were asymptomatic at the end of the study, that is, they do not show any symptoms of DCS. In the AFRL study, researchers noted that there were several individuals who remained asymptomatic regardless of the exposure duration and the final altitude. These individuals represent the cure population, and researchers were interested in determining the characteristics that caused immunity.

The staged ascent profile is an example of a simple step-stress experiment and the presence of a cure population makes the dataset ideal for the proposed model.

In the experiment, we have $\tau_1 = 240$ but the value of τ_2 is unknown. We fit a cumulative risk model to the data and estimate the unknown parameters for fixed τ_2 , assuming it to be known. We vary the value of τ_2 and employ a discrete optimization with a grid size of 5 minutes to find the value at which the likelihood is maximum.

The MLEs of the parameters of the cumulative risk model assuming Weibull, linear failure rate (LFR) and generalized exponential (GE) hazards, their associated standard errors (within parentheses), the corresponding maximized log-likelihood (MLL) values, the Kolmogorov–Smirnov (K–S) distances between the empirical and fitted distributions, and the p values for the K–S test are presented in Table 1.

For comparison purposes, we also fit the [Khamis and Higgins \(1998\)](#) model with a cure fraction. The [Khamis and Higgins \(1998\)](#) model with Weibull hazards is a special case of the cumulative risk model with $\delta = 0$, that is, $\tau_1 = \tau_2$

If we examine the maximized log-likelihood values and the K–S distances, it is clear that the Weibull model provides the best fit. The Kaplan–Meier estimator of the survival function and the survival function based on the Weibull model are provided in Figure 1. The two curves are very close, indicating the Weibull model is indeed a good fit for the data. Using a grid search, the estimate of τ_2 was found to be 350 minutes. This implies a lag of 110 minutes (almost 2 hours) before the effects of the increased altitude are observed. The cumulative hazard function with the estimated value of τ_2 is provided in Figure 2. We also obtain an estimate of the cure fraction p . For the Weibull model, we estimate 18% of the population to be asymptomatic.

Table 1 *Parameter estimates for the models without covariates*

Model	α_1	α_2	λ_1	λ_2	p	LL	K–S	p -value
Weibull	2.4972 (0.4213)	1.5390 (0.2587)	0.2341 (0.0378)	0.9762 (0.1175)	0.1820 (0.0298)	−38.7093 —	0.0502 —	0.7519 —
LFR	0.0985 (0.0118)	6.8254 (1.2319)	0.0102 (0.0008)	1.0915 (0.2478)	0.2214 (0.0215)	−41.1311 —	0.0949 —	0.2114 —
GE	5.0066 (0.9567)	2.6427 (0.4565)	1.3461 (0.3534)	3.6877 (0.7567)	0.1589 (0.0198)	−39.6485 —	0.0589 —	0.6998 —
KH	1.2454 (0.1775)	1.9290 (0.3318)	0.2003 (0.0397)	0.7446 (0.1114)	0.1975 (0.0298)	−39.0482 —	0.1007 —	0.1723 —

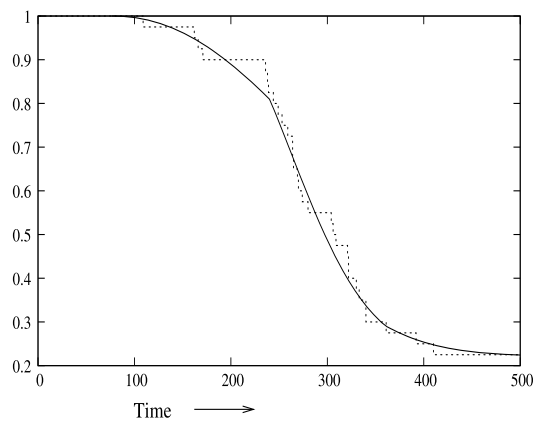


Figure 1 *Kaplan–Meier estimator and the estimated survival function based on the Weibull model. Here time is measured in mins.*

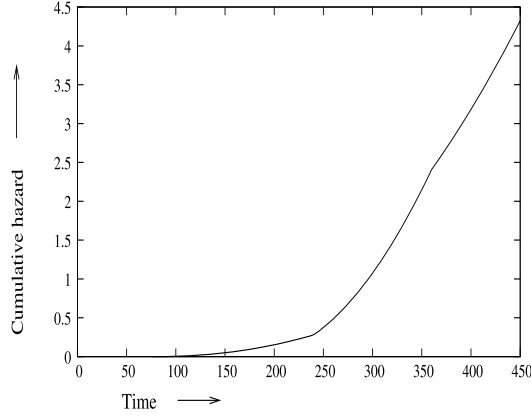


Figure 2 Cumulative hazard function of the Weibull model. Here time is measured in mins.

Table 2 Summary statistics of the covariates

Cov.	min	max	Q_1	Q_2	Q_3
Body fat	5.9	29.9	14.6	17.1	21.1
VO ₂ max	1.9	6	2.7	3.2	3.95
Age	21	43	27	34	37

We can test whether or not the population has a cure fraction by using the procedures outlined in Section 4.2. Under the null hypothesis \mathcal{H}_{04} , the MLEs of the unknown parameters and MLL value are

$$\hat{\alpha}_1 = 1.9685, \quad \hat{\alpha}_2 = 1.8170, \quad \hat{\lambda}_1 = 0.1539, \quad \hat{\lambda}_2 = 1.0582, \\ \text{MLL} = -41.2654.$$

The value of the test statistic is $-2(-41.2654 + 38.7093) = 5.3122$. Using the result (4.20), the p -value is 0.0019, leading to rejection of the null hypothesis. This indicates the population has a cure fraction.

In addition to the event times, researchers also collected covariate information on study participants. For each subject, we have their age (in years), height (in cms), weight (in lbs), body fat percentage (BF) and maximal oxygen uptake (VO₂ max in liter/min). Since the variability in the height and weight measurements was small, we only considered age, body fat and VO₂ max in the model. In Table 2, we provide information about the minimum (min), maximum (max), median (Q_2), first quartile (Q_1) and third quartile (Q_3) of the covariates.

Given that the Weibull model provided the best fit for the data without covariates, we assume a Weibull model with covariate vector $Z = (1, \text{BF}, \text{VO}_2, \text{Age})$. We assume that there is a lag of 110 minutes (i.e., $\tau_2 = 350$). The function $p(\beta, z)$ is as follows:

$$p(\beta, z) = \frac{\exp(\beta_0 + \beta_1 * \text{BF} + \beta_2 * \text{VO}_2 + \beta_3 * \text{Age})}{1 + \exp(\beta_0 + \beta_1 * \text{BF} + \beta_2 * \text{VO}_2 + \beta_3 * \text{Age})}.$$

We fit Weibull models with all possible subsets of covariates. The maximum likelihood estimates of the unknown parameters and the associated standard errors (within parentheses) for all the sub-models are provided in Table 3.

To test whether or not the covariates are significant, we use the procedure outlined for Problem 3 in Section 4.2. The value of the test statistic is $-2(31.3099 - 38.7093) = 14.7988$, with a p -value of 0.002. We reject the null hypothesis and conclude that the covariates considered

Table 3 *Parameter estimates for the Weibull model with covariates*

Covariates	α_1	α_2	λ_1	λ_2	β_0	β_1	β_2	β_3	MLL
$z = (1, \text{BF}, \text{VO}_2, \text{Age})$	2.6299 (0.4789)	1.8450 (0.2918)	0.2215 (0.0445)	1.1348 (0.2191)	-1.8124 (0.5412)	-3.9196 (1.1212)	3.6604 (1.1318)	-0.2154 (0.0519)	-31.3009 -
$z = (1, \text{BF}, \text{VO}_2)$	2.6037 (0.4339)	1.9657 (0.2711)	0.2159 (0.0401)	0.9892 (0.1999)	-1.8503 (0.4798)	-3.9707 (1.1198)	3.4585 (1.1010)	- -	-33.3294 -
$z = (1, \text{BF}, \text{Age})$	2.5249 (0.4412)	1.8836 (0.2798)	0.2339 (0.0466)	1.0436 (0.1814)	-0.2281 (0.0445)	-3.9895 (1.1253)	- -	-1.0607 (0.0498)	-33.6776 -
$z = (1, \text{VO}_2, \text{Age})$	2.5808 (0.4398)	1.8008 (0.2810)	0.2307 (0.0511)	1.1334 (0.1889)	-2.1468 (0.5798)	- -	3.9776 (1.2156)	-1.7856 (0.0610)	-34.2477 -
$z = (1, \text{BF})$	2.4368 (0.4310)	1.9457 (0.2716)	0.2269 (0.0515)	1.0348 (0.1817)	-0.6206 (0.2656)	-3.9817 (1.1191)	- -	- -	-34.0121 -
$z = (1, \text{VO}_2)$	2.5491 (0.4397)	1.7902 (0.2926)	0.2228 (0.0511)	1.1726 (0.1898)	-2.6528 (0.6545)	- -	3.9823 (1.1871)	- -	-34.5618 -
$z = (1, \text{Age})$	2.6066 (0.4365)	1.9207 (0.2710)	0.2268 (0.0516)	0.9736 (0.1912)	-0.3374 (0.1151)	- -	- -	-3.0710 (1.1097)	-34.9897 -

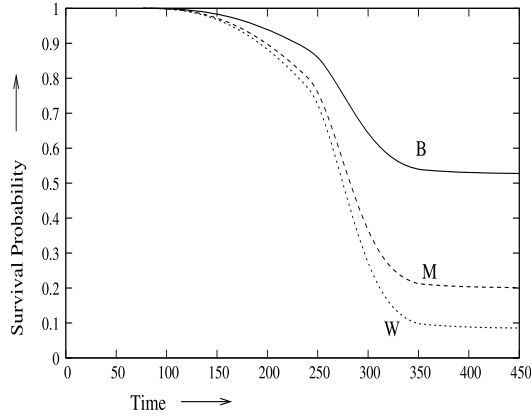


Figure 3 Predicted survival function at three different sets of covariates. Time is measured in mins. Here ‘B’ denotes low risk, ‘M’ denotes medium risk and ‘W’ denotes high risk.

here are significant. As body fat or age increases, the probability of immunity decreases, and as VO_2 increases, the probability of immunity increases. These results are consistent with the physiology of decompression sickness. Using the Weibull model, we provided the predicted survival function at three different sets of covariates. (i) Low Risk (B): $BF = 5.9$, $VO_2 = 6.0$, Age = 21, (ii) Medium Risk (M): $BF = 17.1$, $VO_2 = 3.2$, Age = 34, (iii) High Risk (W): $BF = 29.9$, $VO_2 = 1.9$, Age = 43, in Figure 3. We see that the survival functions are quite close up to 240 minutes (the initial exposure level) but the effects of the change in altitude is different for the three groups. The High Risk group has the lowest proportion of cure individuals.

5.2 Simulation experiments

In this section, we perform some small simulation experiments mainly to see how the proposed methods work for different sample sizes and for the above parametric and covariate set up. We have generated samples from the Weibull model with covariates with the following parameters: $\alpha_1 = 2.6$, $\alpha_2 = 1.8$, $\lambda_1 = 0.22$, $\lambda_2 = 1.14$, $\beta_0 = -1.80$, $\beta_1 = -4.0$, $\beta_2 = 3.7$, $\beta_3 = -0.20$. We have considered the lag-period $\delta = 100$. We have chosen these parameters values since they closely mimic the original data. We have taken different sample sizes namely 40, 80, 120, 160 and 200. Since in the original data set the BF varies between 5.9 and 29.9, we have chosen at random between these two values for a generated sample. Similarly, we have generated at random VO_2 between 1.9 and 6.0, and Age between 20 and 43. We have calculated the MLEs and the associated mean squared errors (MSEs) based on 1000 replications. All the results are based on 20% of censored data. We have reported in Table 4 the average and the square root of the MSEs of all the estimates. It is observed that as sample size increases the MSEs and biases decrease. In all the cases the EM algorithm stops within 35 iterations. It seems the proposed method works well in this particular set up.

6 Conclusions

In this article, we propose a new model for step stress experiments that includes a cure fraction. The susceptible population has been modeled using three different parametric forms of the generalized cumulative risk model, and the probability of being cured is treated as a function of covariates. Different models are fit to data from a study on altitude decompression sickness, that was the motivation for this research. The results show that the Weibull cumulative risk model with a cure fraction provides an excellent fit to the data. Using the model,

Table 4 MLEs for the Weibull model with covariates for different sample sizes

Sample size	α_1	α_2	λ_1	λ_2	β_0	β_1	β_2	β_3
$n = 40$	2.8411 (0.6723)	2.1105 (0.3871)	0.2541 (0.0671)	1.1768 (0.3928)	-2.6816 (0.7129)	-6.1141 (1.3154)	2.1786 (1.4116)	-0.2576 (0.0723)
$n = 80$	2.7921 (0.3817)	2.0156 (0.2011)	0.2487 (0.0401)	1.1523 (0.1987)	-2.4225 (0.3610)	-5.8791 (0.6578)	2.6517 (0.7750)	-0.2443 (0.0411)
$n = 120$	2.7015 (0.2218)	1.9767 (0.1254)	0.2365 (0.0251)	1.1511 (0.1310)	-2.1065 (0.2454)	-4.7910 (0.4516)	2.8917 (0.4519)	-0.2312 (0.0218)
$n = 160$	2.6811 (0.2013)	1.8917 (0.1045)	0.2310 (0.0222)	1.1487 (0.0967)	-1.9167 (0.1810)	-4.4578 (0.3331)	3.2166 (0.3843)	-0.2189 (0.0202)
$n = 200$	2.6115 (0.1319)	1.8214 (0.0798)	0.2242 (0.0156)	1.1411 (0.0861)	-1.8382 (0.1452)	-4.2331 (0.2512)	3.5714 (0.2815)	-0.2067 (0.0114)

we were able to determine the significant covariates, estimate the cure proportion, and find an estimate of the lag period. The model is extremely flexible and can be easily adapted to different distributions and censoring schemes.

Acknowledgments

The research of Nandini Kannan was supported by the NSF IR/D program. However, any opinion, finding, and conclusions and recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation. The authors would like to thank the reviewers for their constructive suggestions which have helped to improve the manuscript significantly.

References

Bai, D. S., Kim, M. S. and Lee, S. H. (1989). Optimum simple step-stress accelerated life test with censoring. *IEEE Transactions on Reliability* **38**, 528–532.

Balakrishnan, N. (2009). A synthesis of exact inferential results for exponential step-stress models and associated optimal accelerated life-tests. *Metrika* **69**, 351–396. [MR2481928](#) <https://doi.org/10.1007/s00184-008-0221-4>

Balakrishnan, N., Kundu, D., Ng, H. K. T. and Kannan, N. (2007). Point and interval estimation for a simple step-stress model with type-II censoring. *Journal of Quality Technology* **9**, 35–47.

Beltrami, J. (2011). Competing risks in the step-stress model with lagged effects. Ph.D. thesis, The University of Texas at San Antonio. [MR2942153](#)

Bhattacharyya, G. K. and Soejoeti, Z. (1989). A tampered failure rate model for step-stress accelerated life test. *Communications in Statistics Theory and Methods* **18**, 1627–1643. [MR1010126](#) <https://doi.org/10.1080/03610928908829990>

Boag, J. W. (1949). Maximum likelihood estimates of the proportion of patients cured by cancer therapy. *Journal of the Royal Statistical Society, Series B* **11**, 15–53.

Cancho, V. G. and Bolafarine, H. (2001). Modeling the presence of immunes by using the exponentiated-Weibull model. *Journal of Applied Statistics* **28**, 659–671. [MR1858829](#) <https://doi.org/10.1080/02664760120059200>

Chen, M.-H., Ibrahim, J. and Sinha, D. (1999). A new Bayesian model for survival data with surviving fraction. *Journal of the American Statistical Association* **94**, 909–918. [MR1723307](#) <https://doi.org/10.2307/2670006>

Gamel, J. W., Mclean, I. W. and Rosenberg, S. H. (1999). Proportion cured and mean long-survival time as function of tumor size. *Statistics in Medicine* **9**, 999–1006.

Ghitany, M. E. and Maller, R. A. (1992). Asymptotic results for exponential mixture models with long-term survivors. *Statistics* **23**, 321–336. [MR1238254](#) <https://doi.org/10.1080/02331889208802379>

Gouno, E. and Balakrishnan, N. (2001). *Step-Stress Accelerated Life Test. Hand Book of Statistics*. Amsterdam, The Netherlands: North-Holland. [MR1861939](#) [https://doi.org/10.1016/S0169-7161\(01\)20025-X](https://doi.org/10.1016/S0169-7161(01)20025-X)

Greven, S., Bailer, J., Kupper, L. L., Muller, K. E. and Craft, J. L. (2004). A parametric model for studying organism fitness using step-stress experiments. *Biometrics* **60**, 793–799. [MR2089456](#) <https://doi.org/10.1111/j.0006-341X.2004.00230.x>

- Gupta, R. D. and Kundu, D. (1999). Generalized exponential distribution. *Australian & New Zealand Journal of Statistics* **41**, 173–188. [MR1705342 https://doi.org/10.1111/1467-842X.00072](https://doi.org/10.1111/1467-842X.00072)
- Ismail, A. A. (2016). Statistical inference for a step-stress partially-accelerated life test model with an adaptive type-I progressively hybrid censored data from Weibull distribution. *Statistical Papers* **57**, 271–301. [MR3483189 https://doi.org/10.1007/s00362-014-0639-x](https://doi.org/10.1007/s00362-014-0639-x)
- Kannan, N., Kundu, D. and Balakrishnan, N. (2010a). Survival models for step-stress experiments with lagged effects. In *Advances in Degradation Modelling* (M. Nikulin, N. Limnios and N. Balakrishnan, eds.) 355–369. New York: Birkhauser. [MR2642594 https://doi.org/10.1007/978-0-8176-4924-1_23](https://doi.org/10.1007/978-0-8176-4924-1_23)
- Kannan, N., Kundu, D., Nair, R. C. and Tripathi, R. C. (2010b). The generalized exponential cure rate model with covariates. *Journal of Applied Statistics* **37**, 1625–1636. [MR2758672 https://doi.org/10.1080/02664760903117739](https://doi.org/10.1080/02664760903117739)
- Kannan, N., Raychaudhury, A. and Pilmanis, A. A. (1998). A logistic model for altitude decomposition sickness. *Aviation, Space, and Environmental Medicine* **69**, 965–970.
- Kateri, M. and Balakrishnan, N. (2008). Inference for a simple step-stress model with type-II censoring and Weibull distributed lifetimes. *IEEE Transactions on Reliability* **57**, 616–626.
- Kateri, M. and Kamps, U. (2015). Inference in step-stress models based on failure rates. *Statistical Papers* **56**, 639–660. [MR3369423 https://doi.org/10.1007/s00362-014-0601-y](https://doi.org/10.1007/s00362-014-0601-y)
- Khamis, I. H. and Higgins, J. J. (1998). A new model for step-stress testing. *IEEE Transactions on Reliability* **47**, 131–134.
- Maller, R. A. and Zhou, S. (1996). *Survival Analysis with Long-Term Survivors*. New York: Wiley. [MR1453117](https://doi.org/10.1002/9781118133231)
- Nelson, W. B. (1980). Accelerated life testing: Step-stress models and data analysis. *IEEE Transactions on Reliability* **29**, 103–108.
- Nelson, W. B. (1990). *Accelerated Life Testing, Statistical Models, Test Plans and Data Analysis*. New York: John Wiley & Sons.
- Sedyakin, N. M. (1966). On one physical principle in reliability theory. *Technical Cybernetics* **3**, 80–87 (in Russian).
- Self, S. G. and Liang, K.-L. (1987). Asymptotic properties of the maximum likelihood estimators and likelihood ratio test under non-standard conditions. *Journal of the American Statistical Association* **82**, 605–610. [MR0898365](https://doi.org/10.1080/01621459.1987.10477455)
- Sha, N. and Pan, R. (2014). Bayesian analysis for step-stress accelerated life testing using Weibull proportional hazard model. *Statistical Papers* **55**, 715–726. [MR3227548 https://doi.org/10.1007/s00362-013-0521-2](https://doi.org/10.1007/s00362-013-0521-2)
- Tsodikov, A. (1998). A proportional hazards model taking account of long-term survivors. *Biometrics* **54**, 1508–1516.
- Xiong, C. (1998). Inferences on a simple step-stress model with type-II censored exponential data. *IEEE Transactions on Reliability* **47**, 142–146.
- Xiong, C. and Milliken, G. A. (1999). Step-stress life-testing with random change times for exponential data. *IEEE Transactions on Reliability* **48**, 141–148.
- Yu, B., Tiwari, R. C., Cronin, K. A. and Feuer, E. C. (2004). Cure fraction estimation from the mixture cure models for grouped survival data. *Statistics in Medicine* **23**, 1733–1747.

Division of Mathematical Sciences
National Science Foundation
Arlington, Virginia 22230
USA

Department of Mathematics and Statistics
Indian Institute of Technology Kanpur
Kanpur, Pin 208016
India
E-mail: kundu@iitk.ac.in