Statistical Analysis for Varying-Stress Accelerated Life Testing with Inverse Gaussian Distribution

Naijun Sha
(Corresponding Author)
Department of Mathematical Sciences
University of Texas at El Paso
TX 79968, USA
&
Richard Nii Okine
Department of Mathematical Sciences
University of Texas at El Paso
TX 79968, USA

Abstract

In this article, we consider a time-varying-stress accelerated life testing (ALT) under aWiener decay process. The failure time of products follows a time-transformed inverse Gaussian distribution. We outline some interesting properties of this highly flexible distribution, present the classical maximum likelihood estimation method, and propose a new Bayesian approach for inference. Simulation studies are carried out to assess the performance of the methods under various settings of parameter values and sample sizes. Real data are analyzed for illustrative purposes to demonstrate the efficiency and accuracy of the proposed Bayesian method over the likelihood-based procedure.

Keywords: varying-stress accelerated life testing, Wiener process, inverse Gaussian distribution, Bayesian inference, Fisher's information, MCMC sampling.

1 Introduction

In industrial experiments, it is often very costly and time consuming to obtain information about the lifetime of high reliable products under normal experimental conditions. To collect failure data quickly and improve the efficiency of experiment, people commonly apply accelerated life testing (ALT) where products or materials are subjected to severe stress conditions than those normally applied in practice. These stresses often include temperature, voltage, vibration, pressure, load, etc., either alone or in some combinations. Step-Stress ALT (SSALT) is one of the most widely used ALT in which the stress on each specimen is increased step-by-step over time.

Through the modeling the distributions of failure times under SSALT that relate to the distribution under usual conditions, the model parameters are estimated and inference is made from the accelerated-failure- time data. Many authors have contributed extensively to the development, parameter inference and the optimum design on SSALT for various lifetime distributions, see, for example, Bai and Chun (1991); Khamis (1997); Bagdonavicius et al. (2002); Srivastava and Shukla (2008); Sha and Pan (2014); Hu et al. (2015), and among others. In this article, we focus on the varying-stress ALT (an extension of SSALT described in Doksum and Hoyland (1992)) where the failure time is modeled in terms of accumulated decay governed by a continuous Wiener process whose distribution, at each time point, depends on the stress assigned to the experimental unit. Specifically, for products subject to constant stress, the accumulated decay is modeled as a Wiener process {W0(y) : y _ 0} with drift $\eta > 0$ and diffusion constant $\delta > 0$. The process is defined to be an independent Gaussian increment with $W_0(0) = 0$, mean $E(W0(y)) = \eta y$, and variance $\delta 2(y_2 - y_1)$ for each increment $W_0(y_2) - W_0(y_1)$, $0 < y_1 < y_2$. Product failure occurs (and the failure time is recorded as Y) when the decay process W0(y) crosses a critical boundary ω. Bhattacharyya and Fries (1982) showed that Y follows the inverse Gaussian distribution IG(μ, λ), whose density function is

$$f_0(y) = \sqrt{\frac{\lambda}{2\pi y^3}} \exp\left\{-\frac{\lambda(y-\mu)^2}{2\mu^2 y}\right\}, \; y > 0, \mu > 0, \lambda > 0, \tag{1}$$

where μ = ω/η is the mean and λ = ω2/δ2 is the shape parameter. The history and basic results of the distribution can be found in the book Chhikara and Folks (1989). In the following section, we focus on a continuous varying-stress ALT governed by aWiener process, and explore the distribution of failure time for products in the process.

We first briefly review the SSALT under a Wiener process described in Doksum and Hoyland (1992), and then extend to the continuous varying-stress ALT on experimental units. Suppose that, for a particular pattern of SSALT with k steps total, step i runs at stress level xi, starts at time ti−1, and runs to time ti(t0 = 0, tk = 1), i = 1, 2, ..., k. A general procedure of SSALT is illustrated as the step lines in Figure 1. The failure time of units is modeled by the accumulated decay Wiener process W(y) whose distribution depends on the stress x(y) assigned to the experimental unit at each time point. Failure time y is defined as the first time the accumulated decay W(y) reaches a fixed critical level ω. Suppose that W0(y) is the decay Wiener process at initial stress, and the decay rate of the process is changed by multiplicative factor αi as y crosses the stress change point ti. Then the accumulative decay from Wiener process is modeled as W(y) = W0(y), y 2 [0, t1), and W(y) = Wi(y), y 2 [ti, ti+1), where Wi(y) = Wi−1(ti + αi(y − ti)), y 2 [ti, ti+1), i = 1, 2, ..., k − 1.

Clearly, it can be expressed as W(y) = W0(τ (y)) where the effective (non-accelerated) time is

$$\tau(y) = \sum_{l=0}^{i-1} \beta_l (t_{l+1} - t_l) + \beta_i (y - t_i), \ \ y \in [t_i, t_{i+1}) \qquad (2)$$

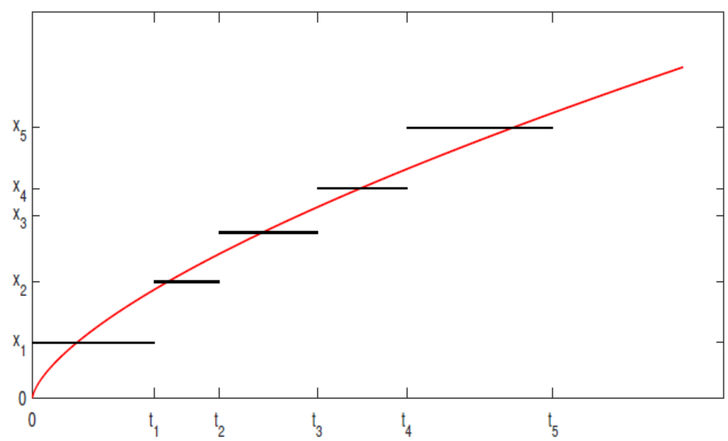


Figure 1: Step and varying stress accelerated life tests.

where the cumulative decay factor β_i = Q_i l=1 α_l with α_0 = 1. The decay rate at ith step αi is related to the stress levels xi−1 and xi. Doksum and Hoyland (1992) showed that the cumulative distribution function of random failure time Y is F(y) = F_0(τ (y)) where $F_0(\cdot)$ is the distribution function of IG(μ, λ). One of the association often used in many applications is α_i = 1+θ(x_i−x_i−1) with some positive value θ. Thus the cumulative decay rate becomes

$$\begin{aligned} \beta_i &= \prod_{l=1}^{i} \alpha_l = \prod_{l=1}^{i} [1 + \theta(x_l - x_{l-1})] = 1 + \sum_{l=1}^{i} \theta(x_l - x_{l-1}) \\ &= 1 + \theta(x_i - x_0) + \theta^2 \sum_{l,j=1}^{i} (x_l - x_{l-1})(x_j - x_{j-1}) + \cdots + \theta^i \prod_{l=1}^{i} (x_l - x_{l-1}) \qquad (3) \end{aligned}$$

As all time duration _i = ti − ti−1 ! 0, the step-stress levels xi become a varying-stress x(y) (see Figure 1) continuously over time in ALT, and the decay process W(y) at time y would have the cumulative decay rate (and denoted as) β_(y) = 1 + θ(x(y) − x(0)), where x(0) = x0.

As a result, the corresponding cumulative exposure (or damage) time τ_(y), which appears as a sum in (2) for a step-stress testing, becomes the integral τ_(y) = R y 0 β_(s)ds. Specifically, in the experiment reported by Nilsson and Uvell

(1985), x(y) is linear function of y, say x(y) = x0 + Ry where R is a known constant set by the experimenter, and thus the cumulative decay rate becomes

$\beta_\theta(y) = 1 + \theta Ry = 1 + \theta y$, where the constant R has been absorbed into θ. It follows that the cumulative exposure time of products at time y is $\tau_\theta(y) = \int_0^y \beta_\theta(s)ds = y + (\theta/2)y^2$, and the distribution and density functions of failure time y in the varying-stress test are

$$\begin{aligned} F(y) &= F_0(\tau_\theta(y)) \\ &= \Phi\left(\sqrt{\frac{\lambda}{\tau_\theta(y)}}\left(\frac{\tau_\theta(y)}{\mu} - 1\right)\right) + \exp\left\{\frac{2\lambda}{\mu}\right\}\Phi\left(-\sqrt{\frac{\lambda}{\tau_\theta(y)}}\left(\frac{\tau_\theta(y)}{\mu} + 1\right)\right), \qquad (4) \\ f(y) &= \beta_\theta(y) f_0(\tau_\theta(y)) \\ &= (1+\theta y)\sqrt{\frac{\lambda}{2\pi\tau_\theta^3(y)}}\exp\left\{-\frac{\lambda(\tau_\theta(y)-\mu)^2}{2\mu^2\tau_\theta(y)}\right\}, \quad y > 0, \mu > 0, \lambda > 0, \qquad (5) \end{aligned}$$

where F0(·), f0(·) are the distribution and density functions of IG(μ, λ), respectively, _(·) is the distribution function of standard normal. The expression of normal mixture in (4) comes from the property of inverse Gaussian distribution function Johnson et al. (1995). Notice that the transformed variate $\tau_\theta(Y) = Y + (\theta/2)Y^2$ follows $\mathrm{IG}(\mu, \lambda)$, and so we regard it in (4) as a transformed inverse Gaussian distribution, denoted by TIG(μ, λ, θ). Since the inverse Gaussian is a unimodal and the transformation τ_(y) is a quadratic increasing function in y > 0, the unimodality will be kept for the three-parameter TIG family in (5). Figure 2 shows various graphs of the density function for different values of μ, λ and θ, where μ is a location-like parameter, λ plays a shape role to make the distribution more like a Gaussian as it tends to be large, and another shape parameter θ moves the density curve distinctly away from the origin to become less right-skewed as it decreases. Additionally, Figure 3 presents the failure (or hazard) function given by h(y) = f(y)/(1 − F(y)) with various values of parameters, showing that generally it is increasing except that it is increasing followed by a bathtub-shaped when both μ and λ are small. Hence, it seems that the distribution is flexible enough to model various situations of product life.

In this article, we investigate the maximum likelihood (ML) estimation of parameters and pro- vide interval estimation based on the Fisher’s information. Further, to contribute a novel research of analysis, we propose a Bayesian approach for parameter inference. The rest of the paper is organized as follows. Section 3 presents the methodology of estimation procedure. Subsequently, we carry out simulation studies to investigate the performance of proposed methods in Section 4. For illustrative purpose, two real data sets are analyzed in Section 5, followed by some concluding remarks in Section 6.

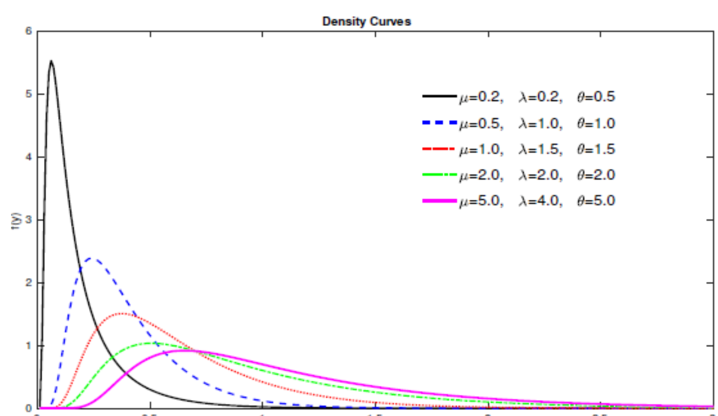


Figure 2: TIG (μ, λ, θ) density curves for various parameter values.

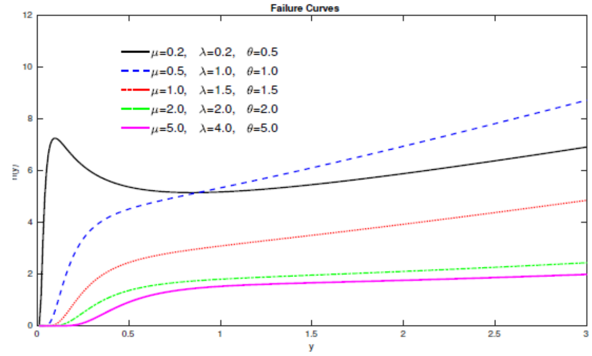


Figure 3: TIG(μ, λ, θ) failure curves for various parameter values.

3 Estimations

Let y = (y1, y2, ..., yn)′ be n observational failure times from the distribution in (4). We first consider the maximum likelihood (ML) estimation approach in the following.

3.1 Likelihood-Based Method

The expression of density function in (5) results in the likelihood and log-likelihood functions, up to a constant, given by

$$L(\mu,\lambda,\theta|\boldsymbol{y}) = \lambda^{n/2}\prod_{i=1}^{n}[(1+\theta y_i)(\tau_\theta(y_i))^{-3/2}]\exp\left\{-\frac{\lambda}{2\mu^2}\sum_{i=1}^{n}\frac{[\tau_\theta(y_i)-\mu]^2}{\tau_\theta(y_i)}\right\}, \quad (6)$$

$$\ell = \ell(\mu,\lambda,\theta|\boldsymbol{y}) = \frac{n}{2}\log(\lambda)+\sum_{i=1}^{n}\left\{\log(1+\theta y_i)-\frac{3}{2}\log(\tau_\theta(y_i))-\frac{\lambda[\tau_\theta(y_i)-\mu]^2}{2\mu^2\tau_\theta(y_i)}\right\}. \quad (7)$$

Since $\partial\tau_\theta(y)/\partial\theta = y^2/2$, we have the first partial derivatives for all the parameters

$$\frac{\partial\ell}{\partial\mu} = \frac{\lambda}{\mu^3}\sum_{i=1}^{n}[\tau_\theta(y_i)-\mu],\ \frac{\partial\ell}{\partial\lambda} = \frac{n}{2\lambda}-\frac{1}{2\mu^2}\sum_{i=1}^{n}\frac{[\tau_\theta(y_i)-\mu]^2}{\tau_\theta(y_i)},$$
$$\frac{\partial\ell}{\partial\theta} = \sum_{i=1}^{n}\frac{y_i}{1+\theta y_i}-\frac{3}{4}\sum_{i=1}^{n}\frac{y_i^2}{\tau_\theta(y_i)}-\frac{\lambda}{4\mu^2}\sum_{i=1}^{n}\frac{[\tau_\theta^2(y_i)-\mu^2]y_i^2}{\tau_\theta^2(y_i)}. \quad (8)$$

There are no analytical forms of maximum likelihood estimates (MLEs), a numerical method has to be applied to obtain MLEs ˆμ, ˆλ, ˆθ by solving the equations ∂ℓ/∂μ = 0, ∂ℓ/∂λ = 0 and ∂ℓ/∂θ = 0 simultaneously. To make inference for the parameters, let the Fisher's information matrix be

$$I(\mu,\lambda,\theta) = \begin{pmatrix} v_{\mu\mu} & v_{\mu\lambda} & v_{\mu\theta} \\ v_{\lambda\mu} & v_{\lambda\lambda} & v_{\lambda\theta} \\ v_{\theta\mu} & v_{\theta\lambda} & v_{\theta\theta} \end{pmatrix}, \quad (9)$$

where the elements are the negative expectation of the second partial derivatives of parameters for the log-likelihood function in (7), such as vμμ = −E(∂²ℓ/∂μ²), vμ_ = −E(∂²ℓ/∂μ∂λ), etc. These elements are given by (the derivations are provided in the appendix)

$$v_{\mu\mu} = -\mathrm{E}\left(\frac{\partial^2\ell}{\partial\mu^2}\right) = \frac{n\lambda}{\mu^3},\ v_{\mu\lambda} = v_{\lambda\mu} = -\mathrm{E}\left(\frac{\partial^2\ell}{\partial\mu\partial\lambda}\right) = 0, \quad (10)$$

$$v_{\mu\theta} = v_{\theta\mu} = -\mathrm{E}\left(\frac{\partial^2\ell}{\partial\mu\partial\theta}\right) = -\frac{n\lambda}{2\mu^3}E(Y^2),\ v_{\lambda\lambda} = -\mathrm{E}\left(\frac{\partial^2\ell}{\partial\lambda^2}\right) = \frac{n}{2\lambda^2}, \quad (11)$$

$$v_{\lambda\theta} = v_{\theta\lambda} = -\mathrm{E}\left(\frac{\partial^2\ell}{\partial\lambda\partial\theta}\right) = \frac{n}{4\mu^2}\left[\mathrm{E}(Y^2)-\mu^2\mathrm{E}\left(\frac{Y^2}{\tau_\theta^2(Y)}\right)\right], \quad (12)$$

$$v_{\theta\theta} = -\mathrm{E}\left(\frac{\partial^2\ell}{\partial\theta^2}\right) = n\left[\mathrm{E}\left(\frac{Y^2}{(1+\theta Y)^2}\right)-\frac{3}{8}\mathrm{E}\left(\frac{Y^4}{\tau_\theta^2(Y)}\right)+\frac{\lambda}{4}\mathrm{E}\left(\frac{Y^4}{\tau_\theta^3(Y)}\right)\right]. \quad (13)$$

So the variance-covariance matrix of the asymptotic normal distribution for the MLEs is given by

$$I^{-1}(\mu,\lambda,\theta)=\frac{1}{d^2}\begin{pmatrix} v_1^2 & v_{\mu\theta}v_{\lambda\theta} & -v_{\mu\theta}v_{\lambda\lambda} \\ v_{\mu\theta}v_{\lambda\theta} & v_2^2 & -v_{\mu\mu}v_{\lambda\theta} \\ -v_{\mu\theta}v_{\lambda\lambda} & -v_{\mu\mu}v_{\lambda\theta} & v_3^2 \end{pmatrix}, \qquad (14)$$

with $v_1^2 = v_{\lambda\lambda}v_{\theta\theta} - v_{\lambda\theta}^2, v_2^2 = v_{\mu\mu}v_{\theta\theta} - v_{\mu\theta}^2, v_3^2 = v_{\mu\mu}v_{\lambda\lambda}$, the determinant $d^2 = v_{\mu\mu}v_1^2 - v_{\lambda\lambda}v_{\mu\theta}^2$.

Due to the positiveness of the parameters, we may use the natural log transformation to obtain approximate confidence intervals (CIs) for the parameters. In particular, for parameter μ and its MLE ˆμ, we have the approximate normal distribution log $(\hat{\mu}) \sim \mathrm{N}(\log(\mu),\ \mathrm{Var}(\log(\hat{\mu})))$, where the variance can be approximated by delta method as $\widehat{\mathrm{Var}(\log(\hat{\mu}))} = \widehat{\mathrm{Var}(\hat{\mu})}/\hat{\mu}^2 = \hat{v}_1^2/(\hat{d}\hat{\mu})^2$ with $\hat{v}_1, \hat{d}$, being the values of v1 and d, respectively, evaluated at MLEs ˆμ,ˆλ, ˆθ and observed data y. Then a (1 − α)100% CI for μ is then given by

$$\left[\hat{\mu}\times\exp\left\{-\frac{z_{\alpha/2}\sqrt{\mathrm{Var}(\hat{\mu})}}{\hat{\mu}}\right\},\ \hat{\mu}\times\exp\left\{\frac{z_{\alpha/2}\sqrt{\mathrm{Var}(\hat{\mu})}}{\hat{\mu}}\right\}\right] \qquad (15)$$

where zα/2is the upper 100 × α/2-th percentile of the standard normal distribution. The normal- approximated CIs for other two parameters λ and θ can be constructed in the same way to have the form in (15) with ˆμ being replaced by ˆλ and ˆθ, respectively, and $\widehat{\mathrm{Var}(\hat{\lambda})} = \hat{v}_2^2/\hat{d}^2, \widehat{\mathrm{Var}(\hat{\theta})} = \hat{v}_3^2/\hat{d}^2$.

It is worth noting that these intervals can exhibit wide width and less accurate coverage for small samples. It is worth noting that these intervals can exhibit wide width and less accurate coverage for small samples.

3.2 Bayesian Approach

Since there are no tractable forms of MLEs and CIs, it is not much accurate and efficient by use of the ML estimation method. Bayesian analysis of the IG distribution has received considerable attention in the past decades, see, for example Padgett (1981); Sinha (1986); Pandey and Bandyopadhyay (2012), where independent priors for the parameters μ and λ were imposed. However, from the model development of the inverse Gaussian distribution in Bhattacharyya and Fries (1982), we notice that the parameters $^{\mu2}$ and λ has a proportional relation λ / $^{\mu2}$. Hence we consider a dependent prior as an alternative in our Bayesian inference approach. Combining the fact that the parameter θ is independent of μ and λ, we propose a joint prior π(μ, λ, θ) = π(μ)π(λ|μ)π(θ) with the conditional prior mean E(λ|μ) / $^{\mu2}$.From the likelihood function in (6), first it is easily discernible that for λ|μ, a gamma is a conjugate prior for the conditional likelihood L(λ|μ, θ, y).

Secondly, the conditional likelihood $L(\mu|\lambda,\theta,y) \propto \exp\{-\lambda/2\sum_{i=1}^{n}[\tau_\theta(y_i)\mu^{-2} - 2\mu^{-1}]\}$, and the gamma prior π(λ|μ) updates the coefficient $\sum_{i=1}^{n}\tau_\theta(y_i)$ of μ^{-2}, hence we choose an inverse gamma for μ to update the coefficient of μ^{-1}. Therefore, we specify the following priors

$$\mu \sim \text{Inv-Gamma}(a_0, a_1),\ \ \lambda|\mu \sim \text{Gamma}\left(\frac{b_0}{2}, \frac{b_1}{2\mu^2}\right) \qquad (16)$$

with the hyperparameters a0, a_1, b_0, b_1 > 0. It is also clear that there is no conjugate prior for θ.However, we may consider a prior distribution which has a similar functional form as its conditional likelihood function. In this case, we pick the prior of θ to be a gamma distribution,

$$\theta \sim \text{Gamma}\,(c_0, c_1)\,, \qquad (17)$$

with c_0,c_1 > 0. The hyperparameter values can be specified by considering the following factors: (i) Since μ and μ^3/λ are the mean and variance of the inverse Gaussian distribution for the transformed variable τ_(Y), we can refer to the sample mean and variance, and MLEs ˆλ, ˆθ in our attempt to specify a_0 and a1 through the prior mean and variance for inverse gamma distribution; (ii) It is known that the mean of the reciprocal of τ_(Y) is 1/μ +1/λ. The sample value of τ^{-1} _ (y) together with the MLEs ˆμ,ˆλ, ˆθ can be used to determine b_i and c_i ,i = 0, 1. The non-informative distribution with $a_i = b_i = c_i = 0$, i = 0, 1 can be chosen if no prior knowledge is available. The joint posterior distribution of the parameters (μ, λ, θ) with the sample data y is given by

$$\pi(\mu, \lambda, \theta|\boldsymbol{y}) \propto L(\mu, \lambda, \theta|\boldsymbol{y}) \times \pi(\mu)\pi(\lambda|\mu)\pi(\theta). \tag{18}$$

It follows that the full conditional posteriors are

$$\pi(\mu|\lambda, \theta, \boldsymbol{y}) \propto \mu^{-(a_0+b_0+1)} \exp\left\{-\left[\frac{\lambda[b_1 + \sum_{i=1}^n \tau_\theta(y_i)]}{2\mu^2} + \frac{a_1 - n\lambda}{\mu}\right]\right\}, \tag{19}$$

$$\lambda|(\mu, \theta, \boldsymbol{y}) \sim \text{Gamma}\left(\frac{\nu_0}{2}, \frac{\nu_1}{2\mu^2}\right), \ \nu_0 = b_0 + n, \ \nu_1 = b_1 + \sum_{i=1}^n [\tau_\theta(y_i) - \mu]^2/\tau_\theta(y_i), \tag{20}$$

$$\pi(\theta|\mu, \lambda, \boldsymbol{y}) \propto \theta^{c_0-1} \prod_{i=1}^n [(1+\theta y_i)(\tau_\theta(y_i))^{-\frac{3}{2}}] \exp\left\{-\left[\left(c_1 + \frac{\lambda}{4\mu^2}\sum_{i=1}^n y_i^2\right)\theta + \frac{\lambda}{2}\sum_{i=1}^n \tau_\theta^{-1}(y_i)\right]\right\}. \tag{21}$$

We implement a Gibbs sampling procedure (a Markov chain Monte Carlo (MCMC) algorithm, see, for example, Casella and George (1992)) to draw posterior samples from their full conditional posterior distributions. First, we take MLEs of μ, λ and θ as the initial values to make the algorithm converge more quickly, and then repeat the following steps M times. Given the values at the mth iteration, the (m + 1)th iteration is as follows:

(i) Draw μ_{m+1} from $\pi(\mu|\lambda_m, \theta_m, y)$ in (19) using a Metropolis-Hastings (MH) procedure (see, Chib and Greenberg (1995)). To make algorithm efficiency, a resembled proposal distribution of the conditional posterior of μ is considered here. We generate a proposed random variate μp from a lognormal distribution with centered at the previous value, i.e. log $N(\log \mu_m, r_\mu \sigma_m^2)$, where $r_\mu > 0$, where $r_\mu > 0$ is a tuning parameter, and the variance term σ_m^2: evaluated at values $(\mu_m, \lambda_m, \theta_m)$, is specified as the reciprocal of Fisher information of the conditional posterior of log μ, whose log density is log $\pi(\mu|\lambda_m, \theta_m, y)$ + log |J| with the Jacobian term J = μ due to a log transformation on μ.

$$\sigma_m^2 = \left[-\text{E}\left\{\frac{\partial^2 (\log \pi(\mu|\lambda_m, \theta_m, \boldsymbol{y}) + \log|J|)}{\partial(\log\mu)^2}\right\}\right]^{-1}\Bigg|_{\mu=\mu_m, \lambda=\lambda_m, \theta=\theta_m}$$

$$= \left[-\mu_m^2 \text{E}\left(\frac{\partial^2 \ell}{\partial \mu_m^2}\right) + \frac{a_1}{\mu_m} + \frac{2b_1\lambda_m}{\mu_m^2}\right]^{-1} = \left[\mu_m^2 v_{\mu_m\mu_m} + \frac{a_1}{\mu_m} + \frac{2b_1\lambda_m}{\mu_m^2}\right]^{-1}, \tag{22}$$

where $v_{\mu_m\mu_m} = n\lambda_m/\mu_m^3$ from the equation (10). Then take $\mu_{k+1} = \mu_p$ with probability

$$\gamma_\mu = \min\left\{1, \frac{\pi(\mu_p|\lambda_m, \theta_m, \boldsymbol{y}) \times q_\mu(\mu_m|\mu_p)}{\pi(\mu_m|\lambda_m, \theta_m, \boldsymbol{y}) \times q_\mu(\mu_p|\mu_m)}\right\}, \tag{23}$$

where $q_\mu(\cdot)$ is the log-normal proposal density, and

$$\frac{q_\mu(\mu_m|\mu_p)}{q_\mu(\mu_p|\mu_m)} = \frac{\sigma_p\mu_p}{\sigma_m\mu_m} \times \frac{\exp\{-(\log\mu_m - \log\mu_p)^2/(2r_\mu\sigma_p^2)\}}{\exp\{-(\log\mu_p - \log\mu_m)^2/(2r_\mu\sigma_m^2)\}}. \tag{24}$$

with σ_p^2 being the same expression of σ_m^2 in (22) with the subscript m replaced by p.

(ii) Draw $\lambda_{m+1|(}\mu_{m+1,}\ \theta_m$, y) _ Gamma(v0/2, $\nu_1/(2\mu^2$ m+1)) with updated v0 and ν_1 in (20).
(iii) Draw θ_{m+1} from $\pi(\theta|\mu_{m+1}, \lambda_{m+1}, y)$ in (21) using a MH procedure. We first propose θ_p _ Gamma(r_θ_m, r_), where θ_m is the mean of the proposal and r_ > 0 is a tuning parameter to make algorithm fast, and then take θm+1 = θ_p with probability

$$\gamma_\theta = \min\left\{1, \frac{\pi(\theta_p|\mu_{m+1}\lambda_{m+1}, \boldsymbol{y}) \times q_\theta(\theta_m|\theta_p)}{\pi(\theta_m|\mu_{m+1}, \lambda_{m+1}, \boldsymbol{y}) \times q_\theta(\theta_p|\theta_m)}\right\}, \quad (25)$$

where $q_\theta(\cdot)$ is the Gamma proposal density, and

$$\frac{q_\theta(\theta_m|\theta_p)}{q_\theta(\theta_p|\theta_m)} = \frac{\Gamma(r_\theta\theta_m) r_\theta^{r_\theta\theta_p}}{\Gamma(r_\theta\theta_p) r_\theta^{r_\theta\theta_m}} \times \frac{\theta_m^{r_\theta\theta_p - 1}\exp\{-r_\theta\theta_m\}}{\theta_p^{r_\theta\theta_m - 1}\exp\{-r_\theta\theta_p\}}. \quad (26)$$

with the gamma function □(·).

The posterior inference of the parameters μ, λ and θ, such as posterior mean, credible interval (CI), etc. are made by their posterior samples (μ_m, λ_m, θ_m),m = 1, 2, ...,M.

4 Simulation Study

We carry out a simulation study to assess performance of parameter estimation by ML and the Bayesian methods. We take four settings of parameter values as (μ, λ, θ) = (0.5, 1.5, 1.0), (1.0, 1.5, 1.5), (2.0, 2.0, 2.0), and generate 10,000 data sets for each of these parameter settings with three sample sizes n = 20, 30, 50. For the Bayesian analysis, we choose “flat” or “less informative” prior distributions with the hyperparameter values $a_0 = a_1 = b_0 = b_1 = c_0 = c_1 = 0$ to reflect little prior knowledge about the parameters. With each simulated data, we find that the the tuning parameter values rμ = 10, r= 2.2 is adequate in ensuring the acceptance rates around 35% - 40%, and we run five MCMC chains with fairly different starting values and each 10,000 iterations with the first 2,000 as a burn-in period. The scale reduction factor estimate $\sqrt{\widehat{RF}} = \sqrt{Var(\psi)/W}$ is used to monitor convergence of MCMC simulations (Gelman et al., 2004), where ψ is the estimand of parameters such as μ, λ and θ, and V ar(ψ) = (N − 1)W/N + B/N with the iteration number N = 10, 000 for each chain, the between- and within-chain variances B and W. The scale factors for the sequences of μ, λ and θ are within 1.01-1.03 for all five MCMC chains, indicating their convergence.

The remaining 8,000 samples are used to compute the average biases, mean squared error (MSE) of the estimates, average lengths (AL) of the 95% credible intervals (CI) (formed by the lower 2.5th and upper 97.5th percentiles), and coverage probability (CP) for the parameters. The results are displayed in Table 1 along with these estimates from the ML method for the purpose of comparison. The main findings are outlined as follows: (i) the bias of estimates, MSE and AL of 95% CI decrease, and CP is closer to the nominal level as sample size n increases for all cases; (ii) the estimation of all parameters from the Bayesian method is much better than from the ML approach as to smaller biases and MSEs, narrower CIs and higher CPs. The MLE ˆθ dose not perform well for the small to moderate sample sizes (n = 20, 30). (iii) relatively, both methods produce much more accurate estimation of λ, less precise for μ and least for θ. With the larger sample size (n = 50), both methods perform similarly for the estimation of the parameter λ; (iv) it is observed that the estimates of μ and θ have a smaller bias and MSE under smaller true values of μ and θ, whereas the estimate of λ has a smaller MSE for both larger and small true values of θ. In summary, the Bayesian inference method outperforms ML approach for all parameter settings, particularly under small sample sizes.

5 Real Data Applications

This section presents real data analysis to further illustrate the usefulness of the proposed method in parameter inference. The first real data example was given in Nelson (2004) about the breakdown time in an accelerated test employed a pair of parallel disk electrodes immersed in an insulating oil. Voltage V (stress) across the pair was increase linearly with time y at a specified rate R in volts/sec and breakdown time was recorded at one squared inch electrode.

Table 1: TIG Estimation Results for Simulated Data

n		ML Method				Bayesian			
		Bias	MSE	AL	CP(%)	Bias	MSE	AL	CP(%)
				$\mu = 0.5, \lambda = 1.0, \theta = 1.0$					
20	μ	0.1103	0.4831	1.1258	92.45	0.0941	0.3889	0.6493	93.11
	λ	0.1212	0.4582	0.5518	92.27	0.1007	0.2244	0.4723	93.19
	θ	0.6673	1.1736	1.3905	89.87	0.2831	0.9122	1.0322	90.74
30		0.1023	0.3663	0.9852	93.17	0.0892	0.2267	0.5529	94.82
		0.1153	0.3053	0.4546	93.21	0.0931	0.1872	0.3676	94.65
		0.4199	0.8568	1.1118	91.29	0.2227	0.7216	0.8209	92.08
50		-0.0718	0.2039	0.7767	94.43	0.0574	0.1313	0.4044	94.94
		0.0970	0.1551	0.4033	94.67	0.0621	0.1268	0.2823	95.15
		0.2368	0.4140	0.7252	94.05	0.1311	0.3416	0.5054	94.80
				$\mu = 1.0, \lambda = 1.5, \theta = 1.5$					
20	μ	0.1504	0.5004	1.0141	92.37	0.1272	0.3884	0.7345	92.87
	λ	0.1829	0.4473	0.4094	92.51	0.1615	0.3316	0.3441	93.10
	θ	0.7914	0.9840	1.2081	89.85	-0.6114	0.9220	1.1099	90.71
30		0.1251	0.3331	0.8386	94.52	0.1105	0.1425	0.5687	93.84
		0.1381	0.2963	0.3457	94.40	0.1147	0.0905	0.3186	94.77
		0.4158	0.8774	0.9422	92.88	0.2783	0.6613	0.7646	92.14
50		0.1009	0.1549	0.5503	94.48	0.0902	0.0822	0.3789	95.18
		0.1492	0.1238	0.2128	94.53	0.1027	0.0778	0.1832	95.22
		0.2280	0.6145	0.6957	93.92	0.2223	0.4207	0.4819	94.29
				$\mu = 2.0, \lambda = 2.0, \theta = 2.0$					
20	μ	0.3992	0.4357	1.0778	91.94	0.1620	0.4012	0.7107	92.72
	λ	0.2742	0.3433	0.6218	92.19	0.1432	0.1493	0.4865	92.89
	θ	0.8871	1.1067	1.2128	88.86	0.5272	0.8531	1.1191	90.05
30		0.3528	0.3830	0.9018	93.52	0.1464	0.3108	0.6109	94.28
		0.2254	0.2038	0.5896	93.88	0.1294	0.1247	0.4033	94.29
		-0.3232	0.9028	0.9480	94.26	-0.3105	0.6424	0.8222	94.89
50		0.1455	0.2503	0.4135	93.49	0.1174	0.1142	0.3278	94.88
		0.1255	0.1243	0.3371	94.31	0.1036	0.1046	0.2883	94.92
		0.2638	0.6316	0.7098	93.03	0.1994	0.4810	0.5144	94.26
				$\mu = 5.0, \lambda = 4.0, \theta = 5.0$					
20	μ	0.4832	0.6303	1.2650	92.08	0.3276	0.5363	0.8058	92.41
	λ	0.3382	0.4777	0.8674	92.40	0.2887	0.3214	0.7195	92.84
	θ	-1.5126	1.2432	1.5216	88.92	1.1411	0.8933	1.2852	90.17
30		0.3683	0.4298	0.9260	93.35	0.2634	0.3891	0.6779	93.71
		0.2909	0.3456	0.7744	93.73	-0.1939	0.2742	0.6438	94.16
		-1.0423	0.9061	1.1238	92.33	0.9320	0.6820	0.8961	93.86
50		0.2068	0.3091	0.7047	94.38	0.1245	0.2128	0.5682	95.03
		-0.1664	0.1636	0.5209	94.73	0.1103	0.1455	0.4105	95.17
		-0.7809	0.6674	0.8675	93.92	0.4106	0.5010	0.5884	94.30

Table 2: Oil breakdown time (seconds) in an accelerated test employed in an insulating oil

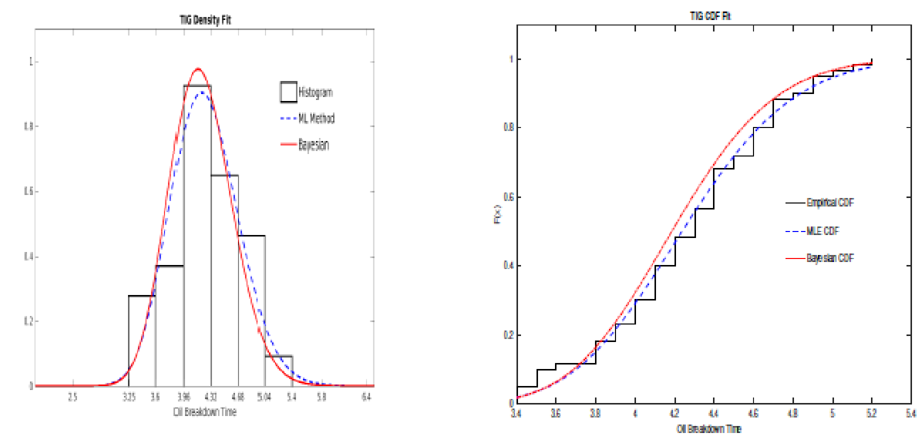
3.4	3.4	3.4	3.5	3.5	3.5	3.6	3.8	3.8	3.8	3.8	3.9	3.9	3.9	4.0
4.0	4.0	4.0	4.1	4.1	4.1	4.1	4.1	4.1	4.2	4.2	4.2	4.2	4.2	4.3
4.3	4.3	4.3	4.3	4.4	4.4	4.4	4.4	4.4	4.4	4.4	4.5	4.5	4.6	4.6
4.6	4.6	4.6	4.7	4.7	4.7	4.7	4.7	4.8	4.9	4.9	4.9	5.0	5.1	5.2

The data is presented in Table 2, consisting of 60 measured breakdown time (seconds). Fitting the TIG distribution with the Bayesian method, we choose the values of hyperparameters a0, a1 and c0, c1 such that the prior mean of μ and θ are close to the MLEs bμ and bθ calculated from the data, b0 and b1 such that the conditional prior mean E(λ|bμ) is close to the MLE bλ. Hence, we have the hyperparameter values as followings: a0 = 10, a_1 = 2.54, b_0 = 0.67, b_1 = 0.16, c_0 = c_1 = 1. We run a chain of 20,000 iterations with a burn-in period of 5,000. To reduce the correlation among the samples, every 5^{th} sample of the remaining 15,000 samples are used for posterior inference. The results are tabulated in Table 3, where, due to relative large sample size (n = 60), the point estimates obtained by both ML and Bayesian methods are close each other. However, the 95% CIs from the Bayesian method are narrower than the ones of ML approach, especially for the intervals of μ and θ. Additionally, a smaller Chi-squared goodness of fit statistic of model fitting by the Bayesian (χ2 = 10.39) than ML method (χ2 = 20.24) indicates more accurate estimation in the Bayesian analysis. The data was analyzed in Nelson (2004), who fitted a Weibull distribution with an inverse power law on the cumulative decay rate β(y) = (x(0)/x(y))p with p = 7.8052, which is associated to a linear law we adopted here by a log transformation. This connection gives us an approximated relation for the rise rate of voltage R _ exp(pθ), leading to the estimated value of R computed, respectively from ML and Bayesian approaches, ˆR = 13.6734, 10.1488, which are close (especially from the Bayesian method) to the estimates obtained by Nelson (2004). Finally, for illustration, Figure 4 depicts the histogram of data and empirical cumulative distribution function (cdf) along with the fitted TIG density and cdf curves estimated by both methods. The second real data was presented in Balakrishna et al. (2009) on active repair times (in hours) for an airborne communications transceiver. To illustrate the estimation performance on a small sample size, we randomly select 20 repair times out of the total 46 observations to have following data: 0.3, 0.5, 0.6, 0.6, 0.7, 0.7, 0.8, 1.0, 1.3, 1.5, 1.5, 2.0, 2.2, 2.5, 4.0, 4.7, 5.0, 7.5, 8.8, 10.3.

Table 3: Breakdown Time: Estimation Results

	ML Method		Bayesian Method	
	Estimate	95% CI	Estimate	95% CI
μ	7.3251	(5.6298, 8.1543)	6.9564	(6.7135, 7.2238)
λ	332.1239	(270.1773, 403.8639)	370.1245	(327.6354, 380.5619)
θ	0.3351	(0.2766, 0.3845)	0.2969	(0.2744, 0.3239)

Figure 4: Breakdown Data



(a) Histogram and fitted density curves by ML and Bayesian estimation methods

(b) Empirical CDF and estimated CDF by ML and Bayesian estimation methods

Modeling the data by the TIG distribution, we adopt the same procedure as discussed to specify the hyperparameter values by using the sample mean 2.825 of the data and MLEs. In summary, we choose the following hyperparameter values: a0 = 3.00, a1 = 6.93, b0 = 6.00, b1 = 44.68, c0 = 0.10, c1 = 30.30. A MCMC chain of 20,000 iterations with a burn-in period of 5,000 produces the estimation results of Bayesian approach, together with the results by ML method, tabulated in Table 4. For this data with relatively small sample size (n = 20), the produced 95% CIs are much narrower, as well as much smaller Chi-squared goodness of fit statistic value 10.87 by the Bayesian method than 26.53 of ML approach. Finally, for illustration, Figure 5 shows that the fitted TIG density and cdf curves estimated by the Bayesian method is better-fit to the histogram and empirical cdf than the ones by the ML approach. These outcomes demonstrate that the proposed Bayesian method produces much more accurate inference under a small sample size.

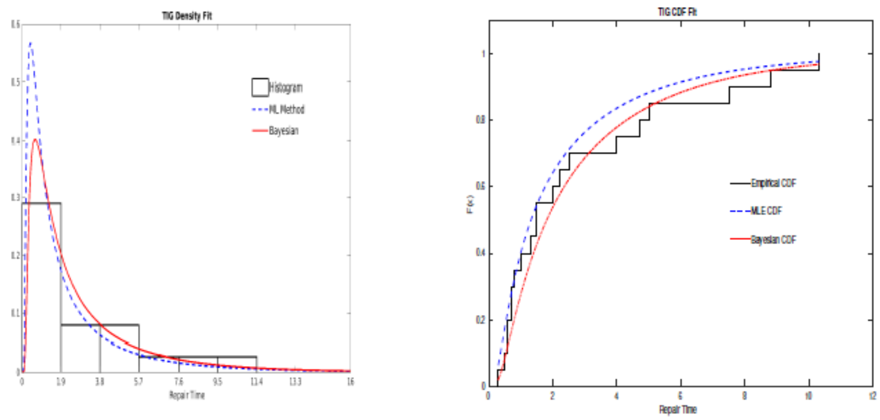
Table 4: Repair Time: Estimation Results

	ML Method		Bayesian Method	
	Estimate	95% CI	Estimate	95% CI
μ	3.0767	(2.8170, 5.6431)	4.3717	(3.9755, 4.8347)
λ	1.2567	(0.7396, 3.7531)	1.9391	(1.5261, 2.3845)
θ	0.1261	(0.0848, 0.7387)	0.1742	(0.1033, 0.4322)

6 Conclusion Remarks

We presented a statistical inference by ML and Bayesian approaches in time-varying-stress accelerated life testing with a Wiener decay process, which induced a transformed inverse Gaussian (TIG) distribution followed by the lifetime of products in the testing. We explored the properties of TIG and studied its Fisher's information used in the likelihood-based inference method, and propose a Bayesian procedure with carefully choosing prior distributions for parameter inference. The simulation study with various settings demonstrated that the Bayesian method outperformed the traditional likelihood-based approach especially for its efficient and impressive outcomes under small sample sizes. We have also illustrated, with two real data sets, that our Bayesian method can be readily applied for efficient, reliable and precise inference.

Figure 5: Repair Data



(a) Histogram and fitted density curves by ML and Bayesian estimation methods

(b) Empirical CDF and estimated CDF by ML and Bayesian estimation methods

Appendix

We present a detail derivation of the Fisher Information matrix for the TIG(μ, λ, θ). From the expressions of first partial derivatives in (8) with ∂τ_(y)/∂θ = $y^2/2$, the second partial derivatives of parameters for the log-likelihood function in (7) are in the followings

$$\frac{\partial^2\ell}{\partial\mu^2} = -\frac{\lambda}{\mu^4}\sum_{i=1}^{n}[3\tau_\theta(y_i)-2\mu],\ \frac{\partial^2\ell}{\partial\mu\partial\lambda} = \frac{1}{\mu^3}\sum_{i=1}^{n}[\tau_\theta(y_i)-\mu], \tag{27}$$

$$\frac{\partial^2\ell}{\partial\lambda^2} = -\frac{n}{2\lambda^2},\ \frac{\partial^2\ell}{\partial\lambda\partial\theta} = -\frac{1}{4\mu^2}\sum_{i=1}^{n}\left(1-\frac{\mu^2}{\tau_\theta^2(y_i)}\right)y_i^2,\ \frac{\partial^2\ell}{\partial\mu\partial\theta} = \frac{\lambda}{2\mu^3}\sum_{i=1}^{n}y_i^2, \tag{28}$$

$$\frac{\partial^2\ell}{\partial\theta^2} = -\sum_{i=1}^{n}\frac{y_i^2}{(1+\theta y_i)^2}+\frac{3}{8}\sum_{i=1}^{n}\frac{y_i^4}{\tau_\theta^2(y_i)}-\frac{\lambda}{4}\sum_{i=1}^{n}\frac{y_i^4}{\tau_\theta^3(y_i)}. \tag{29}$$

By applying $\tau_\theta(Y)\sim \mathrm{IG}(\mu,\lambda)$, it follows that $\mathrm{E}(\tau_\theta(Y))=\mu$, $\mathrm{Var}(\tau_\theta(Y))=\mu^3/\lambda$. Then the elements of Fisher's Information in (9) are

$$v_{\mu\mu} = -\mathrm{E}\left(\frac{\partial^2\ell}{\partial\mu^2}\right)=\frac{n\lambda}{\mu^3},\ v_{\mu\lambda}=v_{\lambda\mu}=-\mathrm{E}\left(\frac{\partial^2\ell}{\partial\mu\partial\lambda}\right)=0, \tag{30}$$

$$v_{\mu\theta} = v_{\theta\mu}=-\mathrm{E}\left(\frac{\partial^2\ell}{\partial\mu\partial\theta}\right)=-\frac{n\lambda}{2\mu^3}E(Y^2),\ v_{\lambda\lambda}=-\mathrm{E}\left(\frac{\partial^2\ell}{\partial\lambda^2}\right)=\frac{n}{2\lambda^2}, \tag{31}$$

$$v_{\lambda\theta} = v_{\theta\lambda}=-\mathrm{E}\left(\frac{\partial^2\ell}{\partial\lambda\partial\theta}\right)=\frac{n}{4\mu^2}\left[\mathrm{E}(Y^2)-\mu^2\mathrm{E}\left(\frac{Y^2}{\tau_\theta^2(Y)}\right)\right], \tag{32}$$

$$v_{\theta\theta} = -\mathrm{E}\left(\frac{\partial^2\ell}{\partial\theta^2}\right)=n\left[\mathrm{E}\left(\frac{Y^2}{(1+\theta Y)^2}\right)-\frac{3}{8}\mathrm{E}\left(\frac{Y^4}{\tau_\theta^2(Y)}\right)+\frac{\lambda}{4}\mathrm{E}\left(\frac{Y^4}{\tau_\theta^3(Y)}\right)\right]. \tag{33}$$

So the inverse Fisher's information matrix has the following forms

$$I^{-1}(\mu,\lambda,\theta) = \frac{1}{d^2}\begin{pmatrix} v_1^2 & v_{\mu\theta}v_{\lambda\theta} & -v_{\mu\theta}v_{\lambda\lambda} \\ v_{\mu\theta}v_{\lambda\theta} & v_2^2 & -v_{\mu\mu}v_{\lambda\theta} \\ -v_{\mu\theta}v_{\lambda\lambda} & -v_{\mu\mu}v_{\lambda\theta} & v_3^2 \end{pmatrix}, \tag{34}$$

with $v_1^2 = v_{\lambda\lambda}v_{\theta\theta} - v_{\lambda\theta}^2, v_2^2 = v_{\mu\mu}v_{\theta\theta} - v_{\mu\theta}^2, v_3^2 = v_{\mu\mu}v_{\lambda\lambda}, d^2 = v_{\mu\mu}v_1^2 - v_{\lambda\lambda}v_{\mu\theta}^2.$

Since there are no analytic forms of some moments in the elements, we estimate these quantities by the observational data to approximate $I^{-1}(\mu, \lambda, \theta)$ in the following

References

Bagdonavicius, V.B., Gerville-Reache, L. & Nikulin, M. (2002). Parametric Inference for step-stress models. *IEEE Transactions on Reliability*, 51, 27-31.

V.B., Gerville-Reache, L. and Nikulin, M. (2002). Parametric Inference for step-stress models. *IEEE Transactions on Reliability*, 51, 27-31.

Bai, D.S. & Chun, Y.R. (1991). Optimum simple step-stress accelerated life tests with competing causes of failure. *IEEE Transactions on Reliability*, 40, 622-627.

Balakrishnan, N., Leiva, V., Sanhueza, A. & Cabrera, E. (2009). Mixture inverse Gaussian distribution and its transformations, moments and applications. *Statistics*, 43, 91-104.

Bhattacharyya, G.K. & Fries, A. (1982). Fatigue Failure Models - Birnbaum-Saunders vs. Inverse Gaussian. *IEEE Transactions on Reliability*, 31(5), 439-441.

Casella, G. & George, E.I. (1992). Explaining the Gibbs sampler. *The American Statistician*, 46(3), 167-174.

Chhikara, R.S. & Folks, L. (1989). *The Inverse Gaussian Distribution: Theory, Methodology, and Applications*, New York: Marcel Dekker.

Chib, S. & Greenberg, E. (1995). Understanding the Metropolis-Hastings algorithm. *The American Statistician*, 49(4), 327-335.

Doksum, K.A. & Hoyland, A. (1992). Models for variable-stress accelerated life testing experiments based on Wiener processes and the inverse Gaussian distribution. *Technometrics*, 34(1), 74-82.

Gelman, A., Carlin, J.B., Stern, H.S. & Rubin, D.B. (2004). *Bayesian Data Analysis, 2nd edition*. Chapman & Hall, London, U.K.

Hu, C., Plante, R.D. & Tang, J. (2015). Equivalent step-stress accelerated life tests with log-location-scale lifetime distributions under type-I censoring. *IIE Transactions*, 47, 245-257.

Johnson, N.L., Kotz, S. & Balakrishnan, N. (1995). *Continuous Univariate Distributions, Vol 2*, New York: John Wiley & Sons.

Khamis, I.H. Optimum M-step step-stress test with K stress variables. *Communications in Statistics: Simulation and Computation*, 26(4), 1301-1313.

Nelson, W. (2004). *Accelerated Testing: Statistical Models, Test Plans, and Data Analysis*, New York: John Wiley & Sons, Inc. 507-509.

Nilsson, L. & Uvell, S. (1985). A model for accelerated tests with applications to optical fibers. *Statistical Research Report*, University of Umea, Department of Mathematical Statistics.

Padgett, W.J. (1981). Bayes estimation of reliability for the inverse Gaussian model. *IEEE Transactions on Reliability*, R-30, 384-385

Pandey, B.N. & Bandyopadhyay, P. (2012) Bayesian Estimation of Inverse Gaussian Distribution. *arXiv:1210.4524*.

Sha, N. & Pan, R. (2014). Bayesian analysis for step-stress accelerated life testing using weibull proportional hazard model. *Statistical Papers*, 55, 715-726.

Sinha, S.K. (1986). Bayesian estimation of the reliability function of the inverse Gaussian distribution. *Statistics and Probability Letters*, 4, 319-323.

Srivastava, P.W. & Shukla, R. (2008). A log-logistic step-stress model. *IEEE Transactions on Reliability*, 57(3), 431-434.