



IJIRCCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 8, Issue 10, October 2020

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 7.488

 9940 572 462

 6381 907 438

 ijircce@gmail.com

 www.ijircce.com

Estimation and Prediction of the GPW Model under Bayesian Framework on Real Data Set

Ramesh Kumar Joshi¹, Ashwini Kumar Srivastava^{2*}, Vijay Kumar³

Associate Professor, Dept. of Statistics, Trichandra Multiple Campus, Tribhuvan University, Kathmandu, Nepal

Assistant Professor, Dept. of Computer Application, Shivharsh.Kisan P.G. College, Basti, U.P., India

Professor, Dept. of Mathematics & Statistics, D.D.U. Gorakhpur University, Gorakhpur, U.P., India

*Corresponding Author

ABSTRACT: In this paper, we consider the estimation and prediction of the Generalized Power Weibull, i.e., GPW model under Bayesian framework. The computation of MLEs, approximate variances, and confidence intervals of the parameters for the GPW model based on a complete sample are performed. For the Bayesian analysis we have assumed the uniform prior for the scale parameter and gamma priors for shape parameters. Then the Gibbs sampling technique is applied to obtain the posterior samples using OpenBUGS software. Bayesian estimators of the parameters, posterior variances, and credible intervals are obtained using posterior samples. We have obtained the Bayes estimates of the hazard and reliability functions, and their probability intervals are also presented. We have applied the predictive check method to discuss the issue of model compatibility. All computational tools are developed in OpenBUGS and R software. A real data set is considered for illustration of the proposed Bayesian approach.

KEYWORDS: GPW model, pdf, cdf, Bayesian framework, MCMC, OpenBUGS, MLE.

I. INTRODUCTION

The Power Generalize Weibull family is another extension of the Weibull family. This family which presented at first by Bagdonavicius and Nikulin [1] contains four shapes of the hazard function and is most used in reliability and survival analysis domain. This family is used often for constructing Accelerated Failures Times (AFT) models, describing dependence of the lifetime distribution on the explanatory variables. It does not only contain models with unimodal and bathtub hazard shape, but also allows for a broader class of monotone hazard rate. The GPW family can be used as a possible alternative to the Exponentiated Weibull family for modeling lifetime data. Nikulin and Haghghi [2] discussed its suitability for the cancer censored data. Lai [3] correctly named the model as Generalized Power Weibull. Lai [3] and Nikulin and Wu [4] provide an excellent review of the Generalized Power Weibull (GPW) model.

II. RELATED WORK

In this paper, we consider the estimation and prediction of the Generalized Power Weibull (GPW) model under Bayesian framework. The computation of MLEs, approximate variances, and confidence intervals of the parameters for the GPW model based on a complete sample are performed. Then the Gibbs sampling technique is applied to construct the Bayesian estimators of the parameters, posterior variances, and credible intervals based on the assumption that the prior of the scale parameter is uniform density function and priors of the shape parameters are gamma density functions. We have obtained the Bayes estimates of the hazard and reliability functions, and their probability intervals are also presented. We have applied the predictive check method to discuss the issue of model compatibility. All computational tools are developed in OpenBUGS and R software. A real data set is considered for illustration of the proposed Bayesian approach.

III. MODEL ANALYSIS

A. Cumulative distribution function:

The cumulative distribution function of GPW model with three parameters is given by

$$F(x; \alpha, \beta, \theta) = 1 - \exp \left\{ 1 - \left[1 + \left(\frac{x}{\beta} \right)^\alpha \right]^\theta \right\}; \alpha > 0, \beta > 0, \theta > 0, x \geq 0, \quad (3.1)$$

where α , β and θ are the parameters, [3]. The GPW distribution will be denoted by $GPW(\alpha, \beta, \theta)$.

The particular cases of the Power Generalized Weibull model are:

- $\theta = 1$: the family of Weibull model;

- $\theta = 1$ and $\alpha = 1$: the family of Exponential model.

B. Probability density function:

The probability density function(PDF) of GPW model with three parameters is given by

$$f(x; \alpha, \beta, \theta) = \frac{\alpha\theta}{\beta} \left(\frac{x}{\beta}\right)^{\alpha-1} \left\{1 + \left(\frac{x}{\beta}\right)^\alpha\right\}^{\theta-1} \exp\left\{1 - \left[1 + \left(\frac{x}{\beta}\right)^\alpha\right]^\theta\right\}, \tag{3.2}$$

where $(\alpha, \beta, \theta) > 0$ and $x \geq 0$.

Differentiating $f(x)$ with respect to x , gives

$$f'(x) = \left\{ (\alpha-1)x^{-1} + \frac{\alpha}{\beta} \left(\frac{x}{\beta}\right)^{\alpha-1} \frac{(\theta-1) - \theta \left(1 + \left(\frac{x}{\beta}\right)^\alpha\right)^\theta}{1 + \left(\frac{x}{\beta}\right)^\alpha} \right\} f(x).$$

If $\alpha \leq 1$ then the $f'(x)$ is non positive implying that the p.d.f's of all GPW model with $\alpha \leq 1$ are monotone decreasing. On the other hand if $\alpha > 1$ then $f''(x) \leq 0$. Hence for $\alpha > 1$, the power Generalized Weibull p.d.f's are unimodal. The shapes of the tails of the densities of the power Generalized Weibull family can be understood by examining the limit of $f(x)$ as $x \rightarrow 0$ and $x \rightarrow \infty$. We have the following:

- (i) If $\alpha > 1$, then $f(x) \rightarrow 0$ as $x \rightarrow 0$ and $x \rightarrow \infty$.
- (ii) If $\alpha = 1$, then $f(x) \rightarrow \frac{\theta}{\beta}$ as $x \rightarrow 0$, and $f(x) \rightarrow 0$ as $x \rightarrow \infty$, hence the p.d.f. is high-tailed at the left end, similar to that of the Exponential model.
- (iii) If $\alpha < 1$, then $f(x) \rightarrow \infty$ as $x \rightarrow 0$, and $f(x) \rightarrow 0$ as $x \rightarrow \infty$, hence the family has a high left tail asymptote at $x = 0$.

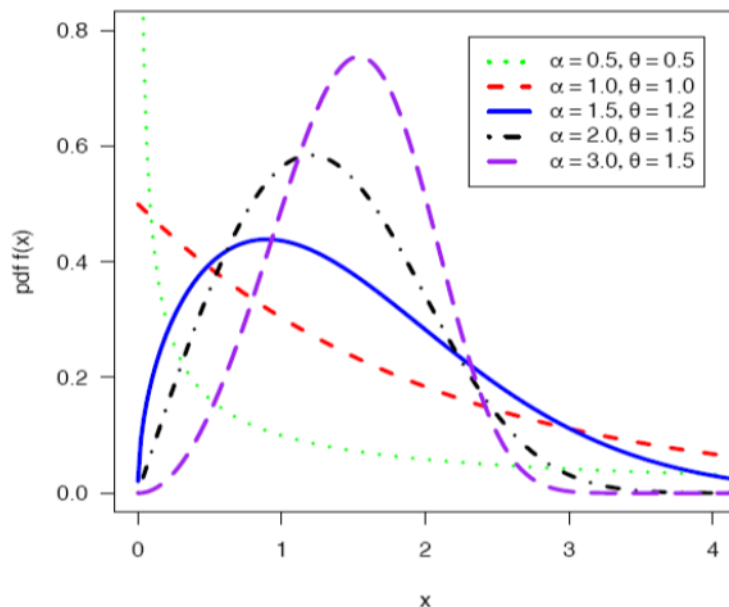


Figure 1 The probability density function of GPW model for $\beta = 1$ and different values of α and θ

The R functions `dgen.pow.weib()` and `pgen.pow.weib()` given in [6] can be used for the computation of pdf and cdf respectively. Some of the typical GPW density functions for different values of α and θ are depicted in Figure 1. which is clearly shows that the density function of the GPW model can take different shapes.

C. The Reliability/Survival function:

The reliability/survival function of GPW model with three parameters is

$$R(x) = \exp \left\{ 1 - \left[1 + \left(\frac{x}{\beta} \right)^\alpha \right]^\theta \right\}; \alpha > 0, \beta > 0, \theta > 0, x \geq 0. \tag{3.3}$$

The R function `sgen.pow.weib()` given in [6] which computes the reliability/ survival function.

D. The hazard rate function:

The hazard rate function (HRF) of GPW model with three parameters is

$$h(x; \alpha, \beta, \theta) = \frac{\alpha\theta}{\beta} \left(\frac{x}{\beta} \right)^{\alpha-1} \left\{ 1 + \left(\frac{x}{\beta} \right)^\alpha \right\}^{\theta-1}. \tag{3.4}$$

The hazard rate has nice and flexible properties[2]. Depending on the values of the parameters, $h(x)$ can be constant, monotone (increasing, decreasing), unimodal or bathtub shaped.

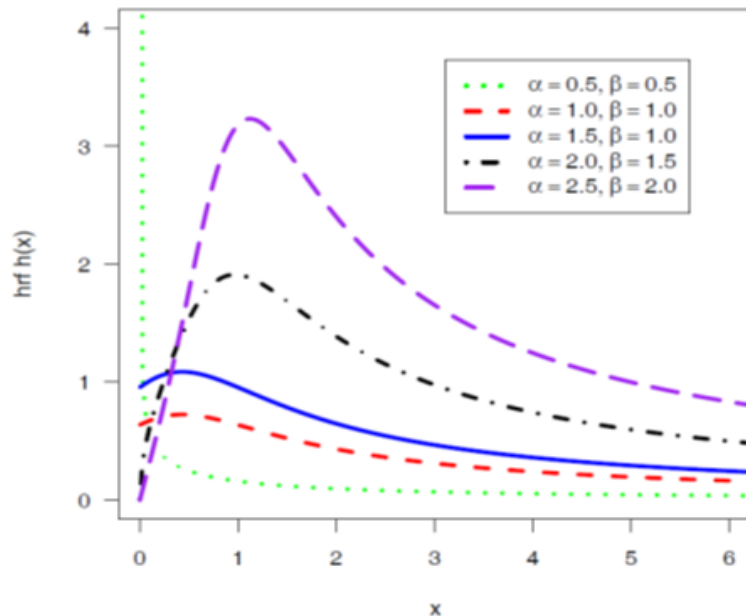


Figure 2 The hazard rate function of GPW model for $\theta = 1$ and different values of α and β .

More specifically, it has been shown that the hazard rate curve is

- (a) monotone increasing if either $\alpha > 1$ and $\alpha\theta > 1$ or $\alpha = 1$ and $\theta > 1$;
- (b) monotone decreasing if either $0 < \alpha < 1$ and $\alpha\theta < 1$ or $0 < \alpha < 1$ and $\alpha\theta = 1$;
- (c) unimodal (inverted bathtub shaped) if $\alpha > 1$ and $0 < \alpha\theta < 1$;
- (d) bathtub shaped $0 < \alpha < 1$ and $\alpha\theta > 1$.

Some of the typical GPW hazard functions for different values of α and θ for $\beta = 1$ are depicted in Figure 5.2. It is clear from the Figure 5.2 that the hazard function of the GPW model can take different shapes including the unimodal (upside down bathtub). The associated R function `hgen.pow.weib()` given in [6]. Figure 5.2 exhibits the different hazard rate functions for GPW model.

E. The Quantile function:

The quantile function of GPW model with three parameters is given by

$$x_p = \beta \left\{ 1 - [\log(1-p) - 1]^{1/\theta} \right\}^{1/\alpha}; 0 < p < 1. \tag{3.5}$$

For the computation of quantiles the R function `qgen.pow.weib()`, given in [6] can be used.

F. The random deviate generation:

The random deviate can be generated from $GPW(\alpha, \beta, \theta)$ by

$$x = \beta \left\{ 1 - [\log(1-u) - 1]^{1/\theta} \right\}^{1/\alpha}; 0 < u < 1, \tag{3.6}$$

where u has the $U(0,1)$ model. The R function `rgen.pow.weib()`, given in [6] the random deviate from $GPW(\alpha, \beta, \theta)$. For model choice based on information criterion, the values of AIC and BIC can be computed using the R function `abic.gen.pow.weib()` given in [6].

IV. MLE AND INFORMATION MATRIX

In this section, we discuss the maximum likelihood estimators (MLE's) of the GPW model and their asymptotic properties to obtain approximate confidence intervals based on MLE's.

Let $\underline{x} = (x_1, \dots, x_n)$ be a random sample of size n from $GPW(\alpha, \beta, \theta)$, then the log-likelihood function $\ell(\alpha, \beta, \theta | \underline{x})$ can be written as;

$$\begin{aligned} \ell(\alpha, \beta, \theta | \underline{x}) = & n \log \alpha + n \log \theta - n \alpha \log \beta + (\alpha - 1) \sum_{i=1}^n \log x_i \\ & + (\theta - 1) \sum_{i=1}^n \log [1 + (x_i/\beta)^\alpha] + n - \sum_{i=1}^n \left\{ 1 + (x_i/\beta)^\alpha \right\}^\theta \end{aligned} \tag{4.1}$$

Differentiating with respect to α, β and θ , we have

$$\frac{\partial \ell}{\partial \alpha} = \frac{n}{\alpha} - n \log \beta + \sum_{i=1}^n \log x_i + (\theta - 1) \sum_{i=1}^n \frac{(x_i/\beta)^\alpha \log(x_i/\beta)}{\left\{ 1 + (x_i/\beta)^\alpha \right\}} - \theta \sum_{i=1}^n \left\{ 1 + (x_i/\beta)^\alpha \right\}^{\theta-1} (x_i/\beta)^\alpha \log(x_i/\beta) = 0$$

$$\frac{\partial \ell}{\partial \beta} = -\frac{n\alpha}{\beta} - \frac{\alpha(\theta - 1)}{\beta} \sum_{i=1}^n \frac{(x_i/\beta)^\alpha}{\left\{ 1 + (x_i/\beta)^\alpha \right\}} + \frac{\alpha\theta}{\beta} \sum_{i=1}^n (x_i/\beta)^\alpha \left\{ 1 + (x_i/\beta)^\alpha \right\}^{\theta-1} = 0$$

and

$$\frac{\partial \ell}{\partial \theta} = \frac{n}{\theta} + \sum_{i=1}^n \log \left\{ 1 + (x_i/\beta)^\alpha \right\} - \sum_{i=1}^n \left\{ 1 + (x_i/\beta)^\alpha \right\}^\theta \log \left\{ 1 + (x_i/\beta)^\alpha \right\} = 0$$

Therefore, to obtain the MLE's of α, β and θ , we can maximize (4.1) directly with respect to α, β and θ or setting these equations to zero and solving them simultaneously yield the maximum likelihood estimates (MLEs) of the model parameters. Numerical methods can be used to obtain the ML estimates of the parameters. For example, the Newton-Raphson iterative technique could be applied to solve the likelihood equations numerically.

Let us denote the parameter vector by $\underline{\varphi} = (\alpha, \beta, \theta)$ and the corresponding MLE of $\underline{\varphi}$ as $\hat{\underline{\varphi}} = (\hat{\alpha}, \hat{\beta}, \hat{\theta})$, then the asymptotic normality results in

$$\left(\hat{\underline{\varphi}} - \underline{\varphi} \right) \rightarrow N_3 \left(0, \left(I(\underline{\varphi}) \right)^{-1} \right) \tag{4.2}$$

where $I(\underline{\varphi})$ is the Fisher's information matrix given by

$$I(\underline{\varphi}) = - \begin{pmatrix} E\left(\frac{\partial^2 \ell}{\partial \alpha^2}\right) & E\left(\frac{\partial^2 \ell}{\partial \alpha \partial \beta}\right) & E\left(\frac{\partial^2 \ell}{\partial \alpha \partial \theta}\right) \\ E\left(\frac{\partial^2 \ell}{\partial \beta \partial \alpha}\right) & E\left(\frac{\partial^2 \ell}{\partial \beta^2}\right) & E\left(\frac{\partial^2 \ell}{\partial \beta \partial \theta}\right) \\ E\left(\frac{\partial^2 \ell}{\partial \theta \partial \alpha}\right) & E\left(\frac{\partial^2 \ell}{\partial \theta \partial \beta}\right) & E\left(\frac{\partial^2 \ell}{\partial \theta^2}\right) \end{pmatrix} \quad (4.3)$$

In practice, it is useless that the MLE has asymptotic variance $(I(\underline{\varphi}))^{-1}$ because we do not know $\underline{\varphi}$. Hence, we approximate the asymptotic variance by “plugging in” the estimated value of the parameters. The common procedure is to use observed Fisher information matrix $O(\hat{\underline{\varphi}})$ (as an estimate of the information matrix $I(\underline{\varphi})$) given by

$$O(\hat{\underline{\varphi}}) = - \begin{pmatrix} \frac{\partial^2 \ell}{\partial \alpha^2} & \frac{\partial^2 \ell}{\partial \alpha \partial \beta} & \frac{\partial^2 \ell}{\partial \alpha \partial \theta} \\ \frac{\partial^2 \ell}{\partial \beta \partial \alpha} & \frac{\partial^2 \ell}{\partial \beta^2} & \frac{\partial^2 \ell}{\partial \beta \partial \theta} \\ \frac{\partial^2 \ell}{\partial \theta \partial \alpha} & \frac{\partial^2 \ell}{\partial \theta \partial \beta} & \frac{\partial^2 \ell}{\partial \theta^2} \end{pmatrix} \Big|_{(\hat{\alpha}, \hat{\beta}, \hat{\theta})} = -H(\underline{\varphi}) \Big|_{\underline{\varphi}=\hat{\underline{\varphi}}} \quad (4.4)$$

where H is the Hessian matrix, $\underline{\varphi} = (\alpha, \beta, \theta)$ and $\hat{\underline{\varphi}} = (\hat{\alpha}, \hat{\beta}, \hat{\theta})$. The Newton-Raphson algorithm to maximize the likelihood produces the observed information matrix. Therefore, the variance-covariance matrix is given by

$$\left(-H(\underline{\varphi}) \Big|_{\underline{\varphi}=\hat{\underline{\varphi}}}\right)^{-1} = \begin{pmatrix} \text{var}(\hat{\alpha}) & \text{cov}(\hat{\alpha}, \hat{\beta}) & \text{cov}(\hat{\alpha}, \hat{\theta}) \\ \text{cov}(\hat{\beta}, \hat{\alpha}) & \text{var}(\hat{\beta}) & \text{cov}(\hat{\beta}, \hat{\theta}) \\ \text{cov}(\hat{\theta}, \hat{\alpha}) & \text{cov}(\hat{\theta}, \hat{\beta}) & \text{var}(\hat{\theta}) \end{pmatrix} \quad (4.5)$$

Hence, from the asymptotic normality of MLEs, approximate $100(1-\eta)\%$ confidence intervals for α , β and θ can be constructed as

$$\hat{\alpha} \pm z_{\eta/2} \sqrt{\text{var}(\hat{\alpha})}, \quad \hat{\beta} \pm z_{\eta/2} \sqrt{\text{var}(\hat{\beta})} \quad \text{and} \quad \hat{\theta} \pm z_{\eta/2} \sqrt{\text{var}(\hat{\theta})} \quad (4.6)$$

where $z_{\eta/2}$ is the upper percentile of standard normal variate.

A. Bayesian Model Formulation

The Bayesian model is constructed by specifying the prior distributions for the model parameters α , β & θ and then multiplying with the likelihood function to obtain the posterior distribution function.

- Probability Model : $f(x | \alpha, \beta, \theta)$
- Prior distribution : $p(\alpha, \beta, \theta)$
- Data : $\underline{x} = (x_1, \dots, x_n)$

Given a set of data $\underline{x} = (x_1, \dots, x_n)$, the likelihood function is

$$L(\alpha, \beta, \theta | \underline{x}) = \alpha^n \theta^n \beta^{-n\alpha} \left(\prod_{i=1}^n x_i^{\alpha-1} \right) \left(\prod_{i=1}^n \left\{ 1 + \left(\frac{x_i}{\beta} \right)^\alpha \right\}^{\theta-1} \right) \exp \left\{ 1 - \left\{ 1 + \left(\frac{x_i}{\beta} \right)^\alpha \right\}^\theta \right\}$$

Denote the prior distribution of α , β and θ as $p(\alpha, \beta, \theta)$. The joint posterior is

$$p(\alpha, \beta, \theta | \underline{x}) = L(\alpha, \beta, \theta | \underline{x}) p(\alpha) p(\beta) p(\theta)$$

B. Prior distributions:

We assume the independent uniform priors for $\beta \sim U(a_2, b_2)$ and gamma priors for $\alpha \sim G(a_1, b_1)$ and $\theta \sim G(a_3, b_3)$ as

$$p(\alpha) = \frac{b_1^{a_1}}{\Gamma(a_1)} \alpha^{a_1-1} e^{-b_1 \alpha} \quad ; \alpha > 0, a_1 > 0, b_1 > 0 \quad , \quad p(\beta) = \frac{1}{b_2 - a_2} \quad ; \quad a_2 < \beta < b_2$$

and

$$p(\theta) = \frac{b_3^{a_3}}{\Gamma(a_3)} \theta^{a_3-1} e^{-b_3 \theta} \quad ; \quad \theta > 0, a_3 > 0, b_3 > 0.$$

C. Posterior distribution:

Combining the likelihood function with the prior via Bayes' theorem yields the posterior up to proportionality as

$$p(\alpha, \beta, \theta | \underline{x}) = \alpha^n \theta^n \beta^{-n\alpha} \left(\prod_{i=1}^n \{1 + (x_i / \beta)^\alpha\}^{\theta-1} \right) \left(\prod_{i=1}^n x_i^{\alpha-1} \right) \exp \left\{ 1 - \{1 + (x_i / \beta)^\alpha\}^\theta \right\} \left(\frac{1}{b_2 - a_2} \right) \left(\frac{b_1^{a_1}}{\Gamma(a_1)} \alpha^{a_1-1} e^{-b_1 \alpha} \right) \left(\frac{b_3^{a_3}}{\Gamma(a_3)} \theta^{a_3-1} e^{-b_3 \theta} \right)$$

or,

$$p(\alpha, \beta, \theta | \underline{x}) \propto \alpha^{n+a_1-1} \theta^{n+a_3-1} \beta^{-n\alpha} \left(\prod_{i=1}^n x_i^{\alpha-1} \right) \left(\prod_{i=1}^n \{1 + (x_i / \beta)^\alpha\}^{\theta-1} \right) \exp \left[1 - \{1 + (x_i / \beta)^\alpha\}^\theta - b_1 \alpha - b_3 \theta \right]$$

The posterior is obviously complicated and no close form inferences appear possible. We, therefore, propose to consider MCMC methods to simulate samples from the posterior so that sample-based inferences can be easily drawn. Markov chain Monte Carlo draws samples by running a cleverly constructed Markov chain that eventually converges to the target distribution (called stationary or equilibrium) which, in our case, is the posterior distribution $p(\alpha, \beta, \theta | \underline{x})$. There are many ways of constructing these chains, but all of them, including the Gibbs sampler [7, 8, 9, 10]

D. Gibbs Sampler : Algorithm

For Gibbs sampler implementation, the full conditionals for α , β and θ upto proportionality can be specified as

(i) Full conditional distribution of the parameter α for given β , θ and \underline{x}

$$p(\alpha | \beta, \theta, \underline{x}) \propto \alpha^{n+a_1-1} \left(\prod_{i=1}^n x_i^{\alpha-1} \right) \left(\prod_{i=1}^n \{1 + (x_i / \beta)^\alpha\}^{\theta-1} \right) \exp \left[1 - \{1 + (x_i / \beta)^\alpha\}^\theta - b_1 \alpha \right]$$

(ii) Full conditional distribution of the parameter β for given α , θ and \underline{x}

$$p(\beta | \alpha, \theta, \underline{x}) \propto \beta^{-n\alpha} \left(\prod_{i=1}^n \{1 + (x_i / \beta)^\alpha\}^{\theta-1} \right) \exp \left[1 - \{1 + (x_i / \beta)^\alpha\}^\theta \right]$$

(iii) Full conditional distribution of the parameter θ for given α , β and \underline{x}

$$p(\theta | \alpha, \beta, \underline{x}) \propto \theta^{n+a_3-1} \left(\prod_{i=1}^n \{1 + (x_i / \beta)^\alpha\}^{\theta-1} \right) \exp \left[1 - \{1 + (x_i / \beta)^\alpha\}^\theta - b_3 \theta \right]$$

We shall use OpenBUGS software to obtain posterior samples. As the GPW model is not available in OpenBUGS, it requires incorporation of a module in *ReliaBUGS* [12] and [13] subsystem of OpenBUGS GPW model. A module `dgen.pow.weib(alpha, beta, theta)` is written in Component Pascal for GPW model, given [12], to perform full Bayesian analysis in OpenBUGS using the method described in [14] and [15]. It is important to note that this module can be used for any set of suitable priors of the model parameters. Almost all aspects of the model in Bayesian framework can be studied using the developed module `dgen.pow.weib(alpha, beta, theta)`

E. Gibbs Sampler : Implementation

1. Select an initial value $\underline{\varphi}^{(0)} = (\alpha^{(0)}, \beta^{(0)}, \theta^{(0)})$ to start the chain.
2. Suppose at the i^{th} -step, $\underline{\varphi} = (\alpha, \beta, \theta)$ takes the value $\underline{\varphi}^{(i)} = (\alpha^{(i)}, \beta^{(i)}, \theta^{(i)})$ then from full conditionals, we generate

$$\alpha^{(i+1)} \text{ from } p(\alpha | \beta^{(i)}, \theta^{(i)}, \underline{x})$$

$\beta^{(i+1)}$ from $p(\beta | \alpha^{(i+1)}, \theta^{(i)}, \underline{x})$ and

$\theta^{(i+1)}$ from $p(\theta | \alpha^{(i+1)}, \beta^{(i+1)}, \underline{x})$.

3. This completes a transition from $\underline{\varphi}^{(i)}$ to $\underline{\varphi}^{(i+1)}$
4. Repeat Step 2, N times.

F. MCMC output : Posterior sample

It is well known that rapid convergence is facilitated by choosing appropriate starting values. In order to guarantee the convergence and to remove the effect of the selection of initial value, the first ‘ B ’ simulated variates are discarded [11]. Also to reduce the effect of autocorrelation, select a sampling lag $L > 1$ after which the corresponding autocorrelation are low. Consider $(\underline{\varphi}^{(1)}, \dots, \underline{\varphi}^{(j)}, \dots, \underline{\varphi}^{(M)})$ as the MCMC output (posterior sample) for the posterior analysis

$$\underline{\varphi}^{(j)} = (\alpha^{(j)}, \beta^{(j)}, \theta^{(j)}); j = 1, 2, \dots, M.$$

Thus, MCMC output is referred as the sample after removing the initial iterations (produced during the burn-in period) and considering the appropriate lag, which can be used to develop the Bayesian inference.

The Bayes estimates of $\underline{\varphi} = (\alpha, \beta, \theta)$, under the square error loss (SEL) function, are given by

$$\hat{\alpha} = \frac{1}{M} \sum_{j=1}^M \alpha^{(j)}; \hat{\beta} = \frac{1}{M} \sum_{j=1}^M \beta^{(j)}; \hat{\theta} = \frac{1}{M} \sum_{j=1}^M \theta^{(j)}$$

The Bayes estimates under absolute and zero-one loss functions are posterior median and mode, respectively.

V. DATA ANALYSIS

A real data set is considered for illustration of the proposed methodology for GPW model. The data below are from an accelerated life test of 59 conductors. Failure times are in hours, and there are no censored observations [16].

A. Classical Analysis

A.1 Exploratory data analysis (EDA)

The aim of data analysis is to gain information from the data. The modern statistical data analysis tools include the exploratory data analysis. Exploratory data analysis is a set of methods to display and summarize the data:

- Displaying the data in a graph that shows overall patterns and unusual observations (boxplot, histogram, density curve etc.)
- Computing descriptive statistics that summarize specific aspects of the data (centre, spread etc.).

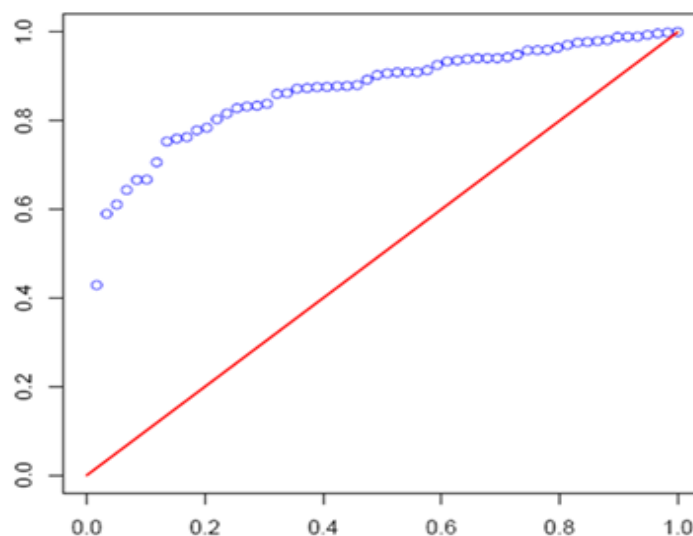


Figure 3 The empirical scaled TTT transform of the data set

In the above Figure 3, we have plotted the empirical version of the scaled TTT transform of the data set in Figure3. Since the empirical version of the scaled TTT transform is concave, it indicates that the hazard function is increasing.

A.2 MLE Computation of GPW model

The maximum likelihood estimates (MLEs) are obtained by direct maximization of the log-likelihood function $\ell(\alpha, \beta, \theta)$ given in (3.1). The advantage of this procedure is that it runs immediately using existing statistical packages such as R[17]. We consider the software R through the Quasi-Newton algorithm [18] to compute the MLEs.

Table 1 MLE, standard error and 95% confidence interval

Parameter	MLE	Std. Error	95%ConfidenceInterval
alpha	6.7548	1.6249	(3.5700, 9.9396)
beta	6.0953	0.5879	(4.9430, 7.2476)
theta	0.4344	0.1817	(0.0783, 0.7905)

From Figure 3 we show that the likelihood equations have a unique solution. The Table 1 shows the ML estimates, standard error (SE) and 95 % confidence Intervals for parameters α , β and θ . The maximized value of loglikelihood is

$$\ell(\hat{\alpha}, \hat{\beta}, \hat{\theta}) = -111.1907.$$

A.3 Process for inspecting the validity of GPW model

To check the validity of the model we compute the Kolmogorov-Smirnov (KS) distance between the empirical distribution function and the fitted distribution function when the parameters are obtained by method of maximum likelihood is 0.0619 and the corresponding p-value is 0.9669.

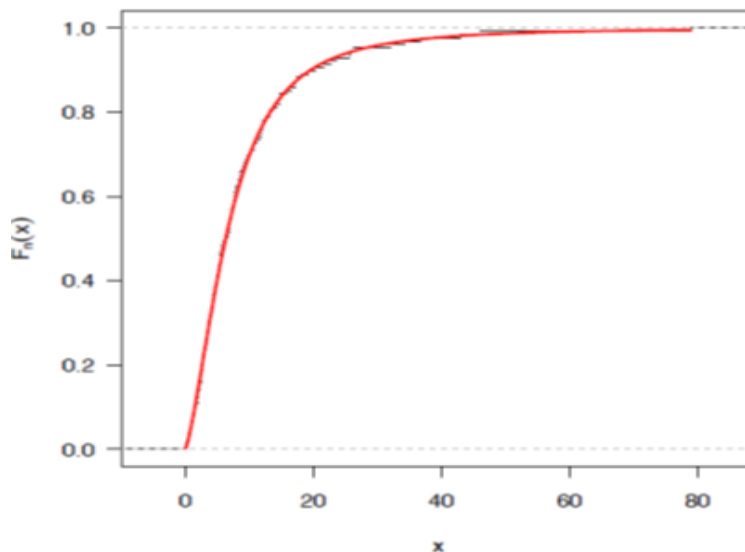


Figure 4. The graph of empirical and fitted model function.

We have used the R function `ks.gen.pow.weib()`, given in [12]. We have plotted the empirical distribution function and the fitted distribution function in Figure 4 which shows very clear that the fitted GPW model provides nice to the given data.

B. Bayesian Analysis

B.1 OpenBUGS script for the Bayesian analysis of GPW model

```

model
{
  for( i in 1 : N )
  {
    x[i] ~ dgen.pow.weib(alpha, beta, theta)
    reliability[i] <- R(x[i], x[i])
  }
}
    
```

```
f[i] <- density(x[i], x[i])
hrf[i] <- hrf(x[i], x[i])
ep[i] <- (i - 0.5) / N
x.new[i] <- beta*pow(pow(1.0-log(1.0-ep[i]),1/theta) -1, 1/alpha)
}
# Prior distributions of the model parameters
alpha ~ dgamma(0.001, 0.001)
theta ~ dgamma(0.001, 0.001)
beta~ dunif(0, 20)
}
```

Data

```
list(N=59, x =c(2.997, 4.137, ..., 10.491, 11.038))
```

Initial values

```
list(alpha=1.0, beta=0.5, theta=0.1) # Chain1
list(alpha=5.0, beta= 5.5, theta=0.5) # Chain2
```

We assume the independent uniform prior for $\beta \sim U(a_2, b_2)$ and gamma priors for $\alpha \sim G(a_1, b_1)$ and $\theta \sim G(a_3, b_3)$ with hyper parameter values

$$(a_1 = 0.001, b_1 = 0.001), (a_2 = 0, b_2 = 20.0) \text{ and } (a_3 = 0.001, b_3 = 0.001).$$

We run the model to generate two Markov Chains at the length of 30,000 with different starting points of the parameters [19]. We have chosen initial values $(\alpha = 1.0, \beta = 0.5, \theta = 0.1)$ for the first chain and $(\alpha = 5.0, \beta = 5.5, \theta = 0.5)$ for the second chain. The convergence is monitored using trace and ergodic mean plots, we find that the Markov Chain converge together after approximately 3000 observations. Therefore, burn-in of 5000 samples is more than enough to erase the effect of starting point(initial values). Finally, samples of size 5000 are formed from the posterior by picking up equally spaced every fifth outcome (to minimize the auto correlation among the generated deviates.), i.e. thin=5, starting from 5001 [20].

Therefore, we have the posterior sample $(\alpha_1^{(j)}, \beta_1^{(j)}, \theta_1^{(j)}) ; j=1, \dots, 5000$ from chain 1 and $(\alpha_2^{(j)}, \beta_2^{(j)}, \theta_2^{(j)}) ; j=1, \dots, 5000$ from chain 2.

The given numerical summary is presented for both the chain 1 and chain 2 respectively [21].

B.2 Numerical Summary

We have considered various quantities of interest and their numerical values based on MCMC sample of posterior characteristics for GPW model. The MCMC results of the posterior mean, mode, standard deviation(SD), first quartile, median, third quartile, 2.5th percentile, 97.5th percentile, skewness, 95% symmetric and HPD credible intervals of the parameters α , β and θ are displayed in Table 2 and Table 3.

Table 2. Numerical summaries based on MCMC sample of posterior characteristics for GPW model

Characteristics	alpha	beta	theta
Mean	6.4377	6.5757	0.6551
Standard Deviation	1.6764	1.2831	0.7386
First Quartile (Q ₁)	5.2388	5.8430	0.3615
Median	6.2565	6.2910	0.4935
Third Quartile (Q ₃)	7.4630	6.8993	0.7027
Mode	5.7480	6.0112	0.3880
Skewness	0.6488	3.4364	8.7652
Kurtosis	0.5784	19.6397	113.9781

Table 3. 95% symmetric and HPD credible intervals

Parameter	Symmetric Credible Interval	HPD Credible Interval
alpha	(3.6448 10.24)	(3.32 9.616)
lambda	(5.2330, 9.9132)	(4.893 8.744)
theta	(0.2184, 2.1152)	(0.1452 1.473)

C. Comparison with MLE

We have used graphical method for the comparison of Bayes estimates with ML estimates. In Figure 5, the density functions $f(x; \hat{\alpha}, \hat{\beta}, \hat{\theta})$ using MLEs and Bayesian estimates (the posterior means), computed via MCMC samples [22], are plotted. It is evident from the above said figure that the MLEs and the Bayes estimates are quite close and fit the data very well.

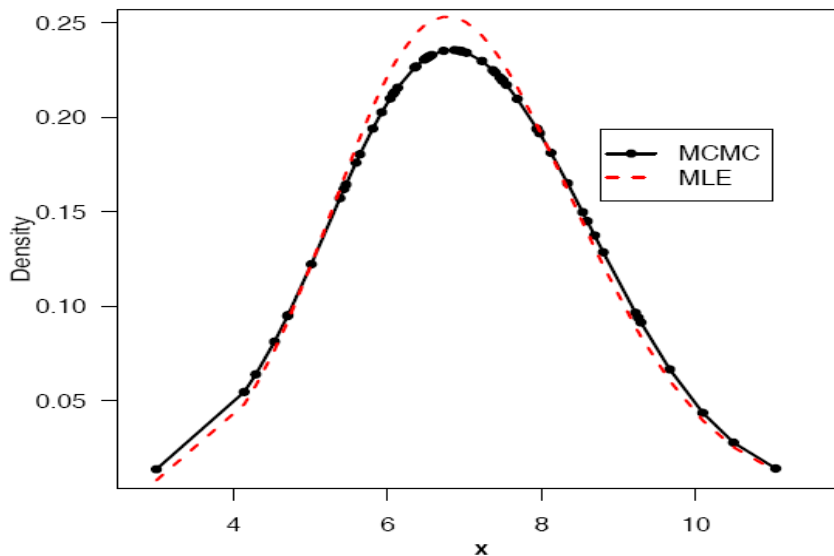


Figure 5 The density functions using ML and Bayesian estimates

D. Posterior Predictive Checks for Model compatibility

A natural way to assess the fit of a Bayesian model is to look at how well the predictions from the model agree with the observed data. We do this by comparing the posterior predictive simulations with the data [22]. The MCMC results of the posterior mean, median, mode of smallest and largest $(X_{(1)}, X_{(2)}, X_{(58)} \text{ and } X_{(59)})$ and $X_{(10)}$ are displayed in Table 4.

Table 4. Posterior characteristics

	Observed	Mode	Mean	Median
$X_{(1)}$	2.997	3.365	3.151	3.227
$X_{(2)}$	4.137	4.043	3.794	3.862
$X_{(10)}$	5.459	5.357	5.315	5.335
$X_{(58)}$	10.491	10.155	10.352	10.270
$X_{(59)}$	11.038	10.838	11.120	11.010

Figure 6 exhibits graphical posterior predictive check of the model adequacy, solid line (-) represents the posterior median and dashed lines (...) represent lower and upper bounds of 95% probability intervals, observed data is superimposed. The predictive data reflect the expected observations after replicating the experiment in future, having already observed x and assuming that the adopted model is true

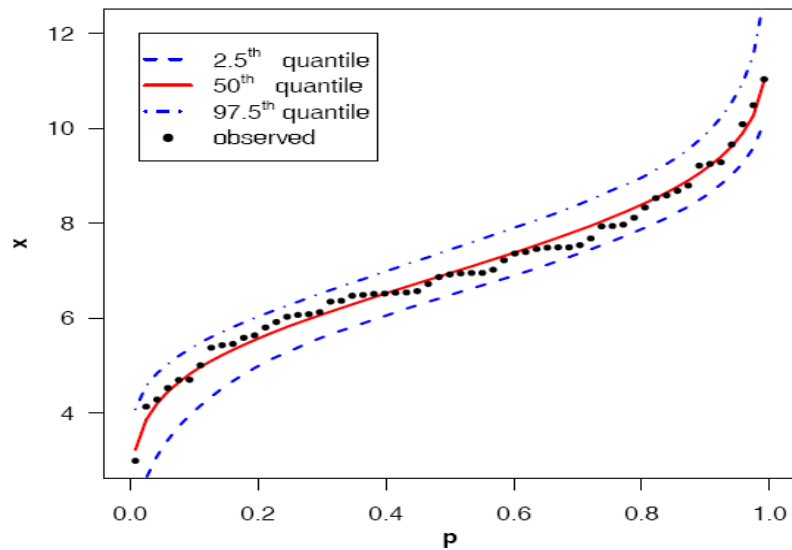


Figure 6. Graphical posterior predictive check of the model adequacy.

VI. CONCLUSION

We have proposed GPW model and discussed some of its properties using R software. We have obtained the MLE of the parameters and their asymptotic probability intervals. Then, we have discussed the Markov chain Monte Carlo (MCMC) method to compute the Bayesian estimates of the parameters, hazard and reliability functions of GPW model based on a complete sample. We have obtained the probability intervals for parameters, hazard and reliability functions. We have presented the model compatibility via the posterior predictive check method. We have applied the developed techniques on a real data set. Thus, the tools developed can be applied for full Bayesian analysis of GPW model.

VII. ACKNOWLEDGEMENT

The authors are thankful to the editor and the referees for their valuable suggestions, which improved the paper to a great extent.

REFERENCES

- [1]. Bagdonavicius V. & Nikulin M. S. (2001). Estimation in Degradation Models with Explanatory Variables, Lifetime Data Analysis volume 7, pages85–103(2001)
- [2]. Nikulin M. & Haghghi F. (2006). A Chi-Squared Test for the Generalized Power Weibull Family for the Head-and-Neck Cancer Censored Data, Journal of Mathematical Sciences 133(3):1333-1341.
- [3]. Lai, Chin-Diew. (2014). Generalized Weibull Distributions, SpringerBriefs in Statistics . eBook ISBN 978-3-642-39106-4
- [4]. Mikhail Nikulin M. & Wu I. (2016). The Cox Proportional Hazards Model, The Cox Model and Its Applications, pp.35-51.
- [5]. Srivastava, A.K. and Kumar V. (2011). Analysis of Software Reliability Data using Exponential Power Model. International Journal of Advanced Computer Science and Applications, Vol. 2, No. 2, February 2011, 38-45.
- [6]. Kumar, V. and Ligges, U. (2011). reliaR : A package for some probability distributions. <http://cran.r-project.org/web/packages/reliaR/index.html>.
- [7]. Geman, S. and Geman, D. (1984). Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. IEEE Trans. Pattern Anal. Mach. Intell., 6 721–741.
- [8]. Gelfand, A. E. and Smith, A. F. M. (1990), “Sampling-based approaches to calculating marginal densities,” Journal of the American Statistical Association, 85, 398–409.
- [9]. Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A., and Teller, E. (1953), “Equations of state calculations by fast computing machines,” J. Chem. Phys., 21, 1087–1091.
- [10]. Hastings, W. (1970), “Monte Carlo sampling methods using Markov chains and their applications,” Biometrika, 57, 97–109.
- [11]. Srivastava, A.K. (2019). “Bayesian Approach in Estimation of Shape and Scale Parameter of Log-Weibull model”. International Journal of Innovative Science, Engineering & Technology, Vol. 6 Issue 12, December 2019, 185-196.

- [12]. Kumar, V., Ligges, U., & Thomas, A. (2010). ReliaBUGS User Manual : A subsystem in OpenBUGS for some statistical models. Version 1.0, OpenBUGS 3.2.1, Retrieved from <http://openbugs.info/w/Downloads/>
- [13]. Lunn, D.J., Jackson, C., Best, N., Andrew, A., & Spiegelhalter, D. (2013). The BUGS Book: A Practical Introduction to Bayesian Analysis. London, UK: Chapman & Hall/CRC.
- [14]. Thomas, A., O'Hara, B., Ligges, U. and Sturtz, S., "Making BUGS Open," R News, 6, 12–17, URL <http://mathstat.helsinki.fi/openbugs/>, (2006).
- [15]. Thomas, A., "OpenBUGS Developer Manual," version 3.1.2, <http://www.openbugs.info/>, (2010).
- [16]. Lawless, J. F., (2003). Statistical Models and Methods for Lifetime Data, 2nd ed., John Wiley and Sons, New York.
- [17]. R Development Core Team (2016), "R: A language and environment for statistical computing," R Foundation for Statistical Computing, Vienna, Austria.
- [18]. Lange, K. (1999). Numerical analysis for statisticians. Springer, New York.
- [19]. Gelman, A. (2003). A Bayesian Formulation of Exploratory Data Analysis and Goodness-of-fit Testing. International Statistical Review, 71(2), 369-382.
- [20]. Gelman, A., J. Carlin, H. Stern and D. Rubin. 2004. Bayesian data analysis. Second Edition, London, Chapman & Hall.
- [21]. Joshi, R.K. and Kumar V. (2020). "New Lindley-Rayleigh Distribution with Statistical properties and Applications" International Journal of Mathematics Trends and Technology, Vol. 66, No. 9, September 2020, 197-208.
- [22]. Joshi, R.K. and Kumar V. (2020). "Lindley inverse Weibull distribution: Theory and Application" Bulletin of Mathematics and Statistics Research, Vol. 08, No. 3, July- September 2020, 32-46. DOI:10.33329/bomsr.8.3.32



INNO  SPACE
SJIF Scientific Journal Impact Factor

Impact Factor:
7.488

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

 9940 572 462  6381 907 438  ijircce@gmail.com



www.ijircce.com

Scan to save the contact details