Statistics and Applications {ISSN 2454-7395 (online)}
Volume 23, No. 2, 2025 (New Series), pp 1–20
http://www.ssca.org.in/journal

# Gumbel Distribution: Comparison of Bayesian Estimators under Various Loss Functions

**P. M. Safwana and C. Chandran**

*Department of Statistics, University of Calicut, Kerala, India*

Received: 20 February 2024; Revised: 06 August 2024; Accepted: 10 August 2024

**Abstract**

Bayesian estimation is carried out for the scale parameter of Gumbel distribution under various loss functions. A new class of loss functions is introduced and an extensive Monte Carlo simulation study is carried out for the comparison of estimators under loss functions. Two different prior distributions for scale parameter are used. The posterior distributions have no closed form hence Lindley approximation as well as Markov Chain Monte Carlo methods are used for deriving estimates. The study is illustrated through a real life data.

*Key words:* Generalized extreme value distribution; Gumbel distribution; Bayesian estimation; Loss functions; Markov Chain Monte Carlo; Lindley approximation.

**AMS Subject Classifications:** 62K05, 05B05

## 1. Introduction

Extreme value distributions are the limiting distributions of maximum (minimum) of a large collection of independent and identically distributed random variables from an arbitrary distribution. The non-degenerate asymptotic distribution of maximum $M_n$ must belong to one of the three possible general families of distributions called Type I, Type II and Type III extreme value distributions, also known as Gumbel, Frechet, and Weibull distributions respectively. These three distributions can be unified into a single family of distributions namely generalized extreme value distribution (GEVD), by incorporating an additional parameter. This unification work is done independently by Von Mises (1964) and Jenkinson (1955) and is a very helpful model for fitting purposes. The distribution function of GEVD has the form,

$$F(x;\alpha,\theta,\xi) = e^{-\left(1+\xi\left(\frac{x-\alpha}{\theta}\right)\right)^{-\frac{1}{\xi}}},$$

where $x > \alpha - \theta/\xi$, $\alpha$, $\theta$ and $\xi$ are the location, scale and shape parameters respectively. The Type I, Type II and Type III extreme value distributions correspond to the cases

Corresponding Author: P. M. Safwana
Email: safwanamath@gmail.com

$\xi = 0$, $\xi > 0$, and $\xi < 0$, respectively. The Gumbel case, the case $\xi = 0$ is interpreted as $\lim_{\xi \to 0} F(x; \alpha, \theta, \xi)$ and the distribution function corresponding to this can be derived as,

$$F(x; \alpha, \theta) = e^{-e^{-\left(\frac{x-\alpha}{\theta}\right)}}, \quad -\infty < x < \infty; \tag{1}$$

where $-\infty < \alpha < \infty$ and $\theta > 0$. Detailed literature on extreme value distributions and other related issues of GEVD can be seen in Leadbetter *et al.* (1983), Embrechts *et al.* (2013) and Coles (2001). The parameter estimation of extreme value distributions and GEVD are studied by many authors in a series of papers. The maximum likelihood (ML) estimation of GEVD are studied by Prescott and Walden (1983), Hosking (1985) and Smith (1985). Since the ML estimators do not exist in some cases other traditional methods of estimation were employed to estimate parameters of extreme value distributions and GEVD. Moment estimators can be derived in explicit form for location and scale parameters, however, they do not perform well in many situations. The probability weighted moments (PWM) estimators of Gumbel parameters were being studied by Landwehr *et al.* (1979). Mahdi and Cenac (2005) compared it with moment estimators and ML estimators and were seen to be favorable with the moment estimators and ML estimators. The PWM estimators of the parameters and quantiles of the GEVD were considered by Hosking *et al.* (1985). Another method of estimation similar to PWM was introduced by Hosking (1990) called method of L-moments. Hosking derived L-moment estimators for generalized extreme value distribution and showed that they are reliably efficient with those of ML estimators. Comparison of moment estimators, ML estimators, maximum entropy estimators, and PWM estimators of Gumbel distribution were studied by Phien (1987). Phien summarized that the moment estimators do not perform very satisfactorily as compared to the remaining three estimators also PWM is the best choice when considering unbiased estimators. The moment estimators have almost the same performance as the ML estimators.

Bayesian estimation techniques for estimating location and scale parameters of Gumbel distribution were also used by several authors. Effect of prior assumptions on the posterior distributions were examined by Coles and Tawn (1996). Coles and Powell (1996) discussed the modeling ideas using approximation techniques as well as exploring the feasibility of performing Bayesian analysis when ML method fails. A conjugate prior distribution for Gumbel parameters were considered by Chechile (2001) and provided the solution for posterior normalization constant along with the marginal distribution for scale parameter. Record value based analysis were studied by Mousa *et al.* (2002) for estimation of location and scale parameters of Gumbel distribution. The estimators of parameters, reliability, and failure rate functions of Gumbel distribution were derived by Al-Aboud (2009) using type II censoring and Lindley approximation. The estimators are derived under both symmetric and asymmetric loss functions such as squared error loss, linex loss, and entropy loss. They observed that linex loss and entropy loss functions are sensitive to parameter values. A non-informative prior distribution for the location parameter and three different prior distributions for the scale parameter were considered in Vidal (2014). Under these conditions the posterior distribution of location and scale parameters, posterior modes, expected values, quantiles, and credibility intervals are also derived.

Recently, Yılmaz *et al.* (2021) compared various estimation techniques to determine the best estimators of Gumbel parameters including classical and Bayesian methods. They

considered maximum likelihood, moments, least square, weighted least square, ordinary least square, percentile, L-moments, trimmed L-moments, and Bain and Engelhardt's method of estimation. Bayesian parameter estimation was considered under squared error loss function and assumed normal prior for location and gamma prior for scale parameters. To obtain Bayes estimators they used Lindley's approximation technique and Markov Chain Monte Carlo (MCMC) method. In accordance with them, the Bayes method of estimation has better performance among all other classical methods, and suggest Lindley's approximation techniques over MCMC. It is remarkable that most of the Bayesian inference approaches have been developed under squared error loss function which is symmetric in nature, and the posterior mean will be the corresponding estimate. However, squared error loss may be inappropriate in many cases. The aim of this work is to obtain the Bayes estimators of scale parameter of Gumbel distribution under a class of loss functions including the squared error loss. The Bayes estimators under various losses and various priors are also compared. The rest of the paper is arranged as follows. In Section 2, we list the various loss functions and their corresponding Bayes estimators that we come across in the later sections of this paper. A new class of loss function is introduced in this section and related estimators are derived. Bayesian estimation procedures using Lindley's approximation method and MCMC method are presented in Section 3. A Monte Carlo simulation study is performed and included in Section 4. The findings in Section 3 are also applied to a real life data in Section 5 and a brief conclusion is given in Section 6.

## 2. Bayesian estimators under various loss functions

The Bayesian method of estimation is an old method of estimation when more additional information about the parameter $\theta$ is known prior to observing data. This assumes that $\theta$ has some probability distribution over the parameter space $\Theta$, known as prior distribution, $\pi(\theta)$. When data $\mathbf{X} = \mathbf{x}$ is observed, the prior distribution can be updated with the sample information to obtain posterior distribution $\pi(\theta \mid \mathbf{x})$ using Bayes rule as follows,

$$\pi(\theta \mid \mathbf{x}) = \frac{L(\mathbf{x};\theta)\pi(\theta)}{m(\mathbf{x})},$$

where $m(\mathbf{x}) = \int_\Theta L(\mathbf{x};\theta)\pi(\theta)$ is the marginal distribution of $\mathbf{x}$ and $L(\mathbf{x};\theta)$ is the likelihood function. Another additional information about parameter $\theta$ is the consequences of our decision known as the loss function, $L(\theta,\hat{\theta})$, which is a function that assigns a penalty when $\hat{\theta}$ is an estimator for true parameter $\theta$. Then the Bayes estimator under the loss $L(\theta,\hat{\theta})$ is that $\hat{\theta}$ which minimizes the Bayes risk function defined by,

$$r(\theta,\hat{\theta}) = \int_\Theta \int_{\mathbf{X}} L(\theta,\hat{\theta})L(\mathbf{x};\theta)\pi(\theta)d\mathbf{x}d\theta.$$

Since $L(\mathbf{x};\theta)\pi(\theta) = \pi(\theta \mid \mathbf{x})m(\mathbf{x})$,

$$r(\theta,\hat{\theta}) = \int_\Theta \int_{\mathbf{X}} L(\theta,\hat{\theta})\pi(\theta \mid \mathbf{x})m(\mathbf{x})d\mathbf{x}d\theta.$$

Minimum of this integral is the same as the minimum of the integral,

$$\int_\Theta L(\theta,\hat{\theta})\pi(\theta \mid \mathbf{x})d\theta, \tag{2}$$

which is known as the posterior expected loss. When doing Bayesian analysis, a number of symmetric and asymmetric loss functions are used in the Bayesian literature. The symmetric loss functions are useful when the positive and negative errors or the over estimation and under estimation are treated equally. In contrast, the asymmetric loss functions gives different weights to positive error and negative error of the same magnitude. Compared with symmetric loss, asymmetric losses have more applications in real life. See for more details, Zellner (1986), Varian (1975), Calabria and Pulcini (1994), Norstrom (1996), and Berger (2013). Some of the loss functions relevant to this paper and corresponding Bayes estimators that minimize the posterior expected loss given in (2), are given in the next section.

### 2.1. Various loss functions and corresponding estimators

Squared error loss function (sqlf) is an extensively used loss function in Bayesian analysis because of its simplicity and does not yield any extensive calculations. The sqlf is given in the form,

$$L_s(\theta, \hat{\theta}) = (\theta - \hat{\theta})^2. \tag{3}$$

Under this loss function mean of the posterior distribution is the Bayes estimator of $\theta$, i.e,

$$\hat{\theta}_s = E_{\pi(\theta|\mathbf{x})}(\theta \mid \mathbf{x}) = E(\theta \mid \mathbf{x}). \tag{4}$$

The squared error loss can be generalized to the weighted squared error loss and has the form,

$$L_{ws}(\theta, \hat{\theta}) = w(\theta)(\theta - \hat{\theta})^2,$$

where $w(\theta)$ is the weight function. The Bayes estimate $\hat{\theta}_{ws}$ under this loss is given by,

$$\hat{\theta}_{ws} = \frac{E(\theta w(\theta) \mid \mathbf{x})}{E(w(\theta) \mid \mathbf{x})}.$$

Both of these loss functions are symmetric in nature. As an alternative to symmetric loss, Norstrom (1996) defined a specific class of asymmetric loss functions called precautionary loss functions (pqlf) and is given by,

$$L_p(\theta, \hat{\theta}) = \frac{(\theta - \hat{\theta})^2}{\hat{\theta}^k} w(\theta), \quad 0 \le k \le 2, w(\theta) > 0, \tag{5}$$

where $w(\theta)$ is an arbitrary weight function and the constant $k \le 2$ ensures that the cost increases as the difference $\theta - \hat{\theta}$ grows, so that $k$ is known as the precautionary index. When $k = 0$ and $w(\theta) = 1$, the precautionary loss reduces to the squared error loss and when $k = 0$ it reduces to the weighted squared error loss. Norstrom showed that the Bayes estimator, $\hat{\theta}_p$ corresponds to precautionary loss has the form,

$$\hat{\theta}_p = \frac{1}{\Psi_1(x)} \Big[ \bar{k} + \sqrt{\bar{k}^2 + \Psi_1(x)\Psi_2(x)} \, \Big], \tag{6}$$

where $\Psi_1(x) = (1+\bar{k})E(w(\theta) \mid \mathbf{x})/E(\theta w(\theta) \mid \mathbf{x})$, $\Psi_2(x) = (1-\bar{k})E(\theta^2 w(\theta) \mid \mathbf{x})/E(\theta w(\theta) \mid \mathbf{x})$, and $\bar{k} = 1 - k$. When $k = 1$ and $w(\theta) = 1$, the loss function in equation (5) will reduces to

$$L_p(\theta, \hat{\theta}) = \frac{(\theta - \hat{\theta})^2}{\hat{\theta}}, \tag{7}$$

and the Bayes estimator corresponds to this loss is,

$$\hat{\theta}_p = \sqrt{E(\theta^2 \mid \mathbf{x})}, \tag{8}$$

the square root of the second moment of the posterior distribution. Varian (1975) introduced another asymmetric loss function called linex loss and Zellner (1986) studied its properties and used it for estimating the mean of normal random variate. The linex loss has the form,

$$L_l(\theta, \hat{\theta}) = b(e^{a(\hat{\theta}-\theta)} - a(\hat{\theta} - \theta) - 1), \quad a \neq 0, b > 0. \tag{9}$$

This function rises approximately exponential on one side of zero and approximately linear on the other side. The linex loss reduces to squared error loss if $\mid a \mid \rightarrow 0$. Under the linex loss, the Bayes estimator, $\hat{\theta}_l$ is given by,

$$\hat{\theta}_l = -\frac{1}{a} \log E(e^{-a\theta} \mid \mathbf{x}).$$

For $a = 1$,

$$\hat{\theta}_l = -\log E(e^{-\theta} \mid \mathbf{x}). \tag{10}$$

As an alternative to linex loss, Calabria and Pulcini (1994) considered a loss function called general entropy loss and has the form,

$$L_e(\theta, \hat{\theta}) = \left(\frac{\hat{\theta}}{\theta}\right)^p - p \log\left(\frac{\hat{\theta}}{\theta}\right) - 1, \tag{11}$$

where $p \neq 0$ is the shape parameter. This is a generalization of the entropy loss used by Dey *et al.* (1986), when the shape parameter $p$ in equation (11) is equal to 1. When $p > 0$, a positive error causes more serious consequences than a negative error and vice versa. The Bayes estimator, $\hat{\theta}_e$ under this loss is,

$$\hat{\theta}_e = \left[E(\theta^{-p} \mid \mathbf{x})\right]^{-1/p},$$

and for $p = 1$,

$$\hat{\theta}_e = \left[E(\theta^{-1} \mid \mathbf{x})\right]^{-1}. \tag{12}$$

That is, $\hat{\theta}$ is the harmonic mean of the posterior distribution, and the Bayes estimator under this loss is the same as that of weighted squared error loss with $w(\theta) = 1/\theta$ and when $p = -1$ the estimator coincides with the estimator under squared error loss function given in equation (3). El-Sayyad (1967) considered another loss function in the form,

$$L_{el}(\theta, \hat{\theta}) = w(\theta)(\log \hat{\theta} - \log \theta)^2, \tag{13}$$

the Bayes estimator under this loss, $\hat{\theta}_{el}$ is given by,

$$\hat{\theta}_{el} = e^{\frac{E(w(\theta) \log \theta \mid \mathbf{x})}{E(w(\theta) \mid \mathbf{x})}},$$

and for $w(\theta) = 1$,

$$\hat{\theta}_{el} = e^{E(\log \theta \mid \mathbf{x})}. \tag{14}$$

In the discussion above we saw that mean and the square root of the second order moment of the posterior distribution are the Bayes estimator under the loss functions sqlf and pqlf respectively. In the next section we introduce a new class of loss function which is a function of sqlf and pqlf and try to get Bayes estimator in terms of the first and second order moments of the posterior distribution.

### 2.2. A generalized class of loss functions

Generally loss functions are non-negative functions that increase to infinity when the distance between $\theta$ and $\hat{\theta}$ increases and decrease to zero when the distance between $\theta$ and $\hat{\theta}$ decreases. Further most of them are smooth convex functions. Here we define a new class of loss functions which is a function of squared error loss defined in equation (3) and precautionary loss defined in equation (7) discussed in the Section 2.1. Let $g(.)$ be a real-valued monotone function, the new class of loss functions $L_g(\theta, \hat{\theta})$ is defined as,

$$L_g(\theta, \hat{\theta}) = g^{-1}(\lambda g(L_s(\theta, \hat{\theta})) + (1-\lambda) g(L_p(\theta, \hat{\theta}))), \quad \lambda \in [0, 1], \tag{15}$$

where $L_s(\theta, \hat{\theta})$ is the sqlf in (3) and $L_p(\theta, \hat{\theta})$ is the pqlf in (7). Note that when $\lambda = 0$, $L_g(\theta, \hat{\theta})$ reduces to $L_p(\theta, \hat{\theta})$ in (7) and when $\lambda = 1$ it reduces to $L_s(\theta, \hat{\theta})$ in (3). The Bayes estimators under these losses are discussed in Section 2.1. Hence, here onward we consider only the case $\lambda \in (0, 1)$ in equation (15). We consider three cases where $g(.)$ is $g_1(z) = \log z$, $g_2(z) = 1/z$ and $g_3(z) = z$ and discuss these in the following three cases.

**Case 1**

Let $g(z) = g_1(z) = \log z$, then $g_1^{-1}(z) = e^z$ and $L_g(\theta, \hat{\theta})$ in (15) becomes,

$$\begin{aligned} L_{g_1}(\theta, \hat{\theta}) &= g_1^{-1}(\lambda g_1(L_s(\theta, \hat{\theta})) + (1-\lambda) g_1(L_p(\theta, \hat{\theta}))) \\ &= \frac{(\theta - \hat{\theta})^2}{\hat{\theta}^{1-\lambda}}. \end{aligned}$$

This is a special case of precautionary loss function defined in equation (5), where $k = 1 - \lambda$ and $w(\theta) = 1$. The Bayes estimator corresponds to this loss can be derive from (6) by substituting $k = 1 - \lambda$ and $w(\theta) = 1$ as,

$$\hat{\theta}_{g_1} = \frac{1}{\Psi_1(x)} \Big[ \lambda + \sqrt{\lambda^2 + \Psi_1(x)\Psi_2(x)} \Big],$$

with $\Psi_1(x) = (1+\lambda)/E(\theta \mid \mathbf{x})$ and $\Psi_2(x) = (1-\lambda)E(\theta^2 \mid \mathbf{x})/E(\theta \mid \mathbf{x})$. That is,

$$\hat{\theta}_{g_1} = \frac{\lambda E(\theta \mid \mathbf{x}) + \sqrt{E(\theta^2 \mid \mathbf{x}) + \lambda^2((E(\theta \mid \mathbf{x}))^2 - E(\theta^2 \mid \mathbf{x}))}}{\lambda + 1}. \tag{16}$$

That is $\hat{\theta}_{g_1}$ can be expressed as the combinations of first order and second order moments of the posterior distribution. For comparison purpose and simplicity we denote these posterior moments as $u = E(\theta \mid \mathbf{x})$ and $v = E(\theta^2 \mid \mathbf{x})$ and the Bayes estimator in equation (16) can be rewritten as,

$$\hat{\theta}_{g_1} = \frac{\lambda u + \sqrt{v + \lambda^2(u^2 - v)}}{\lambda + 1}. \tag{17}$$

**Case 2**

Let $g(z) = g_2(z) = 1/z$, then $g_2^{-1}(z) = 1/z$ and $L_g(\theta, \hat{\theta})$ in (15) becomes,

$$\begin{aligned} L_{g_2}(\theta, \hat{\theta}) &= g_2^{-1}(\lambda g_2(L_s(\theta, \hat{\theta})) + (1-\lambda) g_2(L_p(\theta, \hat{\theta}))) \\ &= \frac{(\theta - \hat{\theta})^2}{\lambda + (1-\lambda)\hat{\theta}}. \end{aligned}$$

The loss $L_{g_2}(\theta, \hat{\theta})$ does not correspond to any existing form of losses discussed earlier, the Bayes estimator under this loss can be derived by minimizing posterior expected loss given in equation (2) and can be derived as,

$$\hat{\theta}_{g_2} = \frac{\lambda - \sqrt{v + 2u\lambda - 2v\lambda + \lambda^2 - 2u\lambda^2 + v\lambda^2}}{\lambda - 1}, \tag{18}$$

where $u = E(\theta \mid \mathbf{x})$ and $v = E(\theta^2 \mid \mathbf{x})$.

**Case 3**

Let $g(z) = g_3(z) = z$ then $g_3^{-1}(z) = z$ and $L_g(\theta, \hat{\theta})$ in (15) becomes,

$$\begin{aligned} L_{g_3}(\theta, \hat{\theta}) &= g_3^{-1}(\lambda g_3(L_s(\theta, \hat{\theta})) + (1-\lambda) g_3(L_p(\theta, \hat{\theta}))) \\ &= \lambda(\theta - \hat{\theta})^2 + (1-\lambda)\frac{(\theta - \hat{\theta})^2}{\hat{\theta}}. \end{aligned}$$

Again by minimizing posterior expected loss given in equation (2) we get the Bayes estimator $\hat{\theta}_{g_3}$ as,

$$\begin{aligned} \hat{\theta}_{g3} = \frac{1 - \lambda - 2u\lambda}{6\lambda} + \frac{1 - \lambda - 2u\lambda^2}{3 \times 2^{2/3}\lambda(H_1(u,v,\lambda) + \sqrt{H_2(u,v,\lambda) + (H_1(u,v,\lambda))^2})^{1/3}} + \\ \frac{1}{6 \times 2^{1/3}\lambda}(H_1(u,v,\lambda) + \sqrt{H_2(u,v,\lambda) + (H_1(u,v,\lambda))^2})^{1/3}, \end{aligned} \tag{19}$$

where $H_1(u,v,\lambda) = -2 + 6\lambda + 12u\lambda - 6\lambda^2 - 24u\lambda^2 - 24u^2\lambda^2 + 108v\lambda^2 + 2\lambda^3 + 12u\lambda^3 + 24u^2\lambda^3 + 16u^3\lambda^3 - 108v\lambda^3$, $H_2(u,v,\lambda) = -4(1 - \lambda - 2u\lambda)^6$, $u = E(\theta \mid \mathbf{x})$ and $v = E(\theta^2 \mid \mathbf{x})$.

The estimators $\hat{\theta}_{g_1}$, $\hat{\theta}_{g_2}$ have a simple form with the combination of first order moment and second order moment of posterior distribution. Although $\hat{\theta}_{g_3}$ has a lengthy expression of first order moment and second order moment, it can be calculated easily. The derivation of these estimators are included in the appended Annexures. The posterior risks under different loss functions are derived and given in Table 1, where $\hat{\theta}_{g_1}$, $\hat{\theta}_{g_2}$ and $\hat{\theta}_{g_3}$ have the expression given in equations (17), (18) and (19) respectively. In the next section we find the Bayes estimators of location and scale parameters of Gumbel distribution under these loss functions using two different prior information.

**Table 1: Posterior risk under various loss functions**

| Loss function | Posterior risk |
|---|---|
| $L_s(\theta,\hat{\theta})$ | $E(\theta^2 \mid x) - (E(\theta \mid x))^2$ |
| $L_p(\theta,\hat{\theta})$ | $2\Big(\sqrt{E(\theta^2 \mid x)} - E(\theta \mid x)\Big)$ |
| $L_l(\theta,\hat{\theta})$ | $\log E(e^{-\theta} \mid x) + E(\theta \mid x)$ |
| $L_e(\theta,\hat{\theta})$ | $E(\log\theta \mid x) + \log E(\theta^{-1} \mid x)$ |
| $L_{el}(\theta,\hat{\theta})$ | $E((\log\theta)^2) - (E(\log\theta))^2$ |
| $L_{g_1}(\theta,\hat{\theta})$ | $\hat{\theta}_{g_1}^{\lambda-1}E(\theta^2 \mid x) - 2\hat{\theta}_{g_1}^{\lambda}E(\theta \mid x) + \hat{\theta}_{g_1}^{\lambda+1}$ |
| $L_{g_2}(\theta,\hat{\theta})$ | $(E(\theta^2 \mid x) - 2\hat{\theta}_{g_2}E(\theta \mid x) + \hat{\theta}_{g_2}^2)/(\lambda + (1-\lambda)\hat{\theta}_{g_2})$ |
| $L_{g_3}(\theta,\hat{\theta})$ | $\lambda(E(\theta^2 \mid x) - 2\hat{\theta}_{g_3}E(\theta \mid x) + \hat{\theta}_{g_3}^2) + ((1-\lambda)(E(\theta \mid x) + \hat{\theta}_{g_3}^2))/(\hat{\theta}_{g_3})$ |

## 3. Bayes estimator of Gumbel scale parameter

We derive the Bayes estimator of the Gumbel scale parameter, with the location parameter set to zero, under various loss functions given in Section 2. The probability density function of Gumbel distribution with location zero has the form,

$$f(x;\theta) = \frac{1}{\theta} e^{-\left(\frac{x}{\theta}\right)} e^{-e^{-\left(\frac{x}{\theta}\right)}}, \quad -\infty < x < \infty,$$

where $\theta > 0$ is the scale parameter. Based on a sample, $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, the likelihood function is given by,

$$L(\mathbf{x};\theta) = \frac{1}{\theta^n}\, e^{-\sum_{i=1}^{n}\left(\frac{x_i}{\theta}\right)} e^{-\sum_{i=1}^{n} e^{-\left(\frac{x_i}{\theta}\right)}}. \tag{20}$$

For a prior distribution $\pi(\theta)$, the posterior distribution can be derived as,

$$\pi(\theta \mid \mathbf{x}) \propto L(\mathbf{x};\theta)\pi(\theta).$$

Here we assume two different prior distributions for $\theta$. As the first case, let $\pi(\theta) \propto \theta^{-1}$, a Jeffreys prior. It follows that the posterior distribution of $\theta$ is given by,

$$\pi(\theta \mid \mathbf{x}) \propto \frac{1}{\theta^{n+1}} e^{-\sum_{i=1}^{n}\left(\frac{x_i}{\theta}\right)} e^{-\sum_{i=1}^{n} e^{-\left(\frac{x_i}{\theta}\right)}}. \tag{21}$$

Now we assume gamma prior for $\theta$ as the second case. That is $\pi(\theta) \propto \theta^{c_1-1}e^{-c_2\theta}$, where $c_1$, and $c_2$ are the hyper parameters. Then the posterior distribution corresponds to this prior has the form,

$$\pi(\theta \mid \mathbf{x}) \propto \theta^{c_1-1}e^{-c_2\theta}\frac{1}{\theta^n}\, e^{-\sum_{i=1}^{n}\left(\frac{x_i}{\theta}\right)} e^{-\sum_{i=1}^{n} e^{-\left(\frac{x_i}{\theta}\right)}}. \tag{22}$$

In order to find the Bayes estimator of $\theta$, the scale parameter of Gumbel distribution, under the loss functions discussed in Section 2, we have to find the posterior expectations, $E(\theta \mid \mathbf{x})$, $E(\theta^2 \mid \mathbf{x})$ , $E(e^{-\theta} \mid \mathbf{x})$, $E(\theta^{-1} \mid \mathbf{x})$ and $E(\log\theta \mid \mathbf{x})$. Whereas the posterior distributions

under two assumed priors have no closed form, we use two approximation techniques to evaluate these expectations, following Lindley (1980) method and the Markov chain Monte Carlo method in this section. Yılmaz *et al.* (2021) applied these methods to acquire Bayes estimates for the parameters of the Gumbel distribution under the squared error loss function and Abbas *et al.* (2019) used to derive Bayes estimates of power Lindley distribution. A brief overview of these two methods are provided, along with the derivation of the posterior expectations for the Gumbel scale parameter in the following subsections.

### 3.1. Lindley approximation

The work of Lindley (1980) provides an asymptotic solution for the ratio of two integrals commonly encountered while evaluating the posterior expectations. Let $q(\theta)$ be any function of $\theta$ then,

$$E(q(\theta) \mid x) \approx \frac{\int q(\theta)L(\mathbf{x};\theta)\pi(\theta)d\theta}{\int L(\mathbf{x};\theta)\pi(\theta)d\theta}, \tag{23}$$

where $\pi(\theta)$ is the prior distribution and $L(\mathbf{x};\theta)$ is the likelihood function. According to Lindley (1980), this ratio of integrals can be approximated by,

$$E(q(\theta) \mid x) \approx q(\tilde{\theta}) + 0.5(\tilde{q}_{11} + 2\tilde{q}_1\tilde{\pi}_1)\tilde{\sigma}_{11} + 0.5(\tilde{q}_1\tilde{L}_{111}\tilde{\sigma}_{11}^2), \tag{24}$$

where $\tilde{\theta}$ is the ML estimate of $\theta$,

$$\tilde{q}_1 = \frac{\partial q(\tilde{\theta})}{\partial\tilde{\theta}}, \quad \tilde{q}_{11} = \frac{\partial^2 q(\tilde{\theta})}{\partial\tilde{\theta}^2}, \quad \tilde{\pi}_1 = \frac{\partial \log\pi(\tilde{\theta})}{\partial\tilde{\theta}},$$

$$\tilde{L}_{111} = \frac{\partial^3 \log L(\mathbf{x},\tilde{\theta})}{\partial\tilde{\theta}^3}, \quad \tilde{L}_{11} = \frac{\partial^2 \log L(\mathbf{x},\tilde{\theta})}{\partial\tilde{\theta}^2}, \quad \tilde{\sigma}_{11} = \frac{-1}{\tilde{L}_{11}}.$$

The approximation of posterior expectations given in Section 2 for Jeffreys prior can be achieved by utilizing equation (24) in the following manner.

$$E(\theta \mid \mathbf{x}) \approx \tilde{\theta} + \tilde{\pi}_1\tilde{\sigma}_{11} + 0.5(\tilde{L}_{111}\tilde{\sigma}_{11}^2)$$

$$E(\theta^2 \mid \mathbf{x}) \approx \tilde{\theta}^2 + (1 + 2\tilde{\theta}\tilde{\pi}_1)\tilde{\sigma}_{11} + \tilde{\theta}\tilde{L}_{111}\tilde{\sigma}_{11}^2$$

$$E(e^{-\theta} \mid \mathbf{x}) \approx e^{-\tilde{\theta}} + 0.5(e^{-\tilde{\theta}} - 2e^{-\tilde{\theta}}\tilde{\pi}_1)\tilde{\sigma}_{11} - 0.5(e^{-\tilde{\theta}}\tilde{L}_{111}\tilde{\sigma}_{11}^2)$$

$$E(\theta^{-1} \mid \mathbf{x}) \approx \tilde{\theta}^{-1} + (\tilde{\theta}^{-3} - \tilde{\theta}^{-2}\tilde{\pi}_1)\tilde{\sigma}_{11} - 0.5(\tilde{\theta}^{-2}\tilde{L}_{111}\tilde{\sigma}_{11}^2)$$

$$E(\log\theta \mid \mathbf{x}) \approx \tilde{\theta}^{-1} + 0.5(-\tilde{\theta}^{-2} + 2\tilde{\theta}^{-1}\tilde{\pi}_1)\tilde{\sigma}_{11} - 0.5(\tilde{\theta}^{-1}\tilde{L}_{111}\tilde{\sigma}_{11}^2),$$

where,

$$\begin{aligned}
\tilde{L}_{111} &= \frac{-2n}{\tilde{\theta}^3} + 6\sum_{i=1}^{n}\frac{x_i}{\tilde{\theta}^4} + 6\sum_{i=1}^{n}e^{-\frac{x_i}{\tilde{\theta}}}\frac{x_i}{\tilde{\theta}^4} - 6\sum_{i=1}^{n}e^{-\frac{x_i}{\tilde{\theta}}}\frac{x_i^2}{\tilde{\theta}^5} + \sum_{i=1}^{n}e^{-\frac{x_i}{\tilde{\theta}}}\frac{x_i^3}{\tilde{\theta}^6},\\
\tilde{L}_{11} &= \frac{n}{\tilde{\theta}^2} - 2\sum_{i=1}^{n}\frac{x_i}{\tilde{\theta}^3} - 2\sum_{i=1}^{n}e^{-\frac{x_i}{\tilde{\theta}}}\frac{x_i}{\tilde{\theta}^3} + \sum_{i=1}^{n}e^{-\frac{x_i}{\tilde{\theta}}}\frac{x_i^2}{\tilde{\theta}^3},\\
\tilde{\sigma}_{11} &= \frac{-1}{\tilde{L}_{11}}, \quad \tilde{\pi}_1 = \frac{-1}{\tilde{\theta}}.
\end{aligned}$$

The Bayes estimate corresponds to each loss function can now be evaluated by plug in these approximations of posterior expectations in equations given in Section 2 accordingly. Similarly one can evaluate estimates for scale parameter $\theta$ under gamma prior.

### 3.2. MCMC method

MCMC method was first introduced by Metropolis *et al.* (1953) and Hastings (1970). The Metropolis-Hastings algorithm is among the straightforward MCMC techniques used to generate samples from distributions that might otherwise be challenging to sample from. Here we use the Metropolis-Hasting algorithm to generate sample from posterior distribution. Let $\pi(. \mid \mathbf{x})$ be the posterior distribution and consider normal distribution as a proposal distribution, the algorithm can be framed as follows,

1. Initialize $\theta_1 = \tilde{\theta}$ and set $t = 1$
2. Repeat : $t = 1$ to $N$
   (a) Generate $y$ from $I(\theta > 0)N(\theta_t, 0.1)$
   (b) Generate U from Uniform (0,1) and if

$$U \leq \frac{\pi(y \mid \mathbf{x})}{\pi(\theta_t \mid \mathbf{x})}$$

   accept $y$ and set $\theta_{t+1} = y$ otherwise set $\theta_{t+1} = \theta_t$
   (c) increment t.

Using this algorithm we can generate observations for $\theta$ from posteriors in equation (21) curresponds to Jeffreys prior and in equation (22) for gamma prior. Thus we can calculate the posterior expectations after discarding first b (burn in period) observations by the approximation,

$$E(q(\theta)|\mathbf{x}) = \frac{1}{N-b}\sum_{t=1}^{N-b} q(\theta_t).$$

Substituting $q(\theta)$ as $\theta$, $\theta^2$, $e^{-\theta}$, $\theta^{-1}$ and $\log\theta$ the posterior expectations can be obtained. The Bayes estimates corresponding to each loss function can now be evaluated by substituting these approximations of posterior expectations into the equations provided in Section 2 accordingly.

## 4. Simulation and calculations

Extensive analysis is conducted through Monte Carlo simulation to estimate and compare the scale parameter of the Gumbel distribution under various loss functions and priors. We simulate 1000 samples of sizes 30, 50, and 100 using the inverse transformation method. For the gamma prior, we set the hyper parameters as $c_1 = c_2 = 1$. The estimators under the loss functions $L_{g_1}(\theta, \hat{\theta})$, $L_{g_2}(\theta, \hat{\theta})$ and $L_{g_3}(\theta, \hat{\theta})$ were derived using Mathematica software. To evaluate the estimates $\hat{\theta}_{g_1}$, $\hat{\theta}_{g_2}$ and $\hat{\theta}_{g_3}$, we assume some arbitrary values for $\lambda$ ranging from 0 to 1, specifically tabulated for $\lambda = \{0.25, 0.5, 0.75\}$ for simplicity. Also we compute the Bayes estimates based on 20000 MCMC samples and burn in period b is set as 5000. The maximum likelihood estimate and all other computations are done using the R software. The simulated estimates, posterior risk and absolute bias, for $\theta = 1$ under Jeffreys prior and gamma prior using Lindley approximation are presented in Table 2 and

**Table 2: The Bayesian estimates and respective simulated absolute bias and posterior risk for $\theta = 1$ under Jeffreys prior and gamma prior using Lindleys method**

| n | Estimator | Jeffreys prior | | | Gamma prior | | |
|---|---|---|---|---|---|---|---|
| | | Estimate | Posterior risk | Absolute bias | Estimate | Posterior risk | Absolute bias |
| 30 | $\hat{\theta}_s$ | 1.084661 | 0.009769 | 0.084661 | 0.974197 | 0.021422 | 0.025803 |
| | $\hat{\theta}_p$ | 1.089181 | 0.009039 | 0.089181 | 0.985135 | 0.021876 | 0.014865 |
| | $\hat{\theta}_l$ | 1.079026 | 0.004773 | 0.079026 | 0.963569 | 0.001344 | 0.036431 |
| | $\hat{\theta}_e$ | 1.071072 | 0.082695 | 0.071072 | 0.952852 | 0.147326 | 0.047149 |
| | $\hat{\theta}_{el}$ | 0.986032 | 0.067908 | 0.013967 | 1.104104 | 0.063667 | 0.104104 |
| | $\hat{\theta}_{g_{1,.25}}$ | 1.088051 | 0.009220 | 0.088051 | 0.982404 | 0.021784 | 0.017596 |
| | $\hat{\theta}_{g_{1,.5}}$ | 1.086922 | 0.009402 | 0.086922 | 0.979674 | 0.021677 | 0.020326 |
| | $\hat{\theta}_{g_{1,.75}}$ | 1.085793 | 0.009585 | 0.085793 | 0.976940 | 0.021557 | 0.023060 |
| | $\hat{\theta}_{g_{2,.25}}$ | 1.018206 | 0.020074 | 0.018206 | 0.982366 | 0.021783 | 0.017634 |
| | $\hat{\theta}_{g_{2,.5}}$ | 1.015716 | 0.020152 | 0.015716 | 0.979616 | 0.021675 | 0.020384 |
| | $\hat{\theta}_{g_{2,.75}}$ | 1.013205 | 0.020218 | 0.013205 | 0.976892 | 0.021555 | 0.023108 |
| | $\hat{\theta}_{g_{3,.25}}$ | 1.087983 | 0.009226 | 0.087983 | 0.983520 | 0.021823 | 0.016481 |
| | $\hat{\theta}_{g_{3,.5}}$ | 1.086835 | 0.009410 | 0.086835 | 0.979731 | 0.021679 | 0.020269 |
| | $\hat{\theta}_{g_{3,.75}}$ | 1.085729 | 0.009591 | 0.085729 | 0.976990 | 0.021558 | 0.023010 |
| 50 | $\hat{\theta}_s$ | 1.048778 | 0.009065 | 0.048778 | 0.983729 | 0.012326 | 0.016271 |
| | $\hat{\theta}_p$ | 1.053094 | 0.008630 | 0.053094 | 1.009107 | 0.012302 | 0.009107 |
| | $\hat{\theta}_l$ | 1.043957 | 0.008190 | 0.043956 | 0.980286 | 0.009294 | 0.019718 |
| | $\hat{\theta}_e$ | 1.038437 | 0.044702 | 0.038437 | 0.974213 | 0.082056 | 0.025787 |
| | $\hat{\theta}_{el}$ | 0.993037 | 0.025964 | 0.006963 | 1.057525 | 0.024007 | 0.057525 |
| | $\hat{\theta}_{g_{1,.25}}$ | 1.052015 | 0.008741 | 0.052015 | 0.991104 | 0.012470 | 0.008896 |
| | $\hat{\theta}_{g_{1,.5}}$ | 1.050937 | 0.008850 | 0.050937 | 0.989543 | 0.012440 | 0.010457 |
| | $\hat{\theta}_{g_{1,.75}}$ | 1.049858 | 0.008958 | 0.049858 | 0.987982 | 0.012405 | 0.012018 |
| | $\hat{\theta}_{g_{2,.25}}$ | 1.011043 | 0.012073 | 0.011043 | 0.991093 | 0.012470 | 0.008907 |
| | $\hat{\theta}_{g_{2,.5}}$ | 1.009542 | 0.012103 | 0.009542 | 0.989527 | 0.012440 | 0.010473 |
| | $\hat{\theta}_{g_{2,.75}}$ | 1.008032 | 0.012129 | 0.008032 | 0.987968 | 0.012404 | 0.012032 |
| | $\hat{\theta}_{g_{3,.25}}$ | 1.051974 | 0.008743 | 0.051974 | 0.991114 | 0.012470 | 0.008886 |
| | $\hat{\theta}_{g_{3,.5}}$ | 1.050884 | 0.008852 | 0.050884 | 0.989560 | 0.012440 | 0.010440 |
| | $\hat{\theta}_{g_{3,.75}}$ | 1.049820 | 0.008960 | 0.049820 | 0.987996 | 0.012405 | 0.012004 |
| 100 | $\hat{\theta}_s$ | 1.021974 | 0.005365 | 0.021974 | 0.991169 | 0.006144 | 0.008831 |
| | $\hat{\theta}_p$ | 1.024596 | 0.005243 | 0.024596 | 0.994263 | 0.006189 | 0.005737 |
| | $\hat{\theta}_l$ | 1.019216 | 0.002123 | 0.019216 | 0.988110 | 0.002443 | 0.011890 |
| | $\hat{\theta}_e$ | 1.016277 | 0.021730 | 0.016277 | 0.985059 | 0.040463 | 0.014941 |
| | $\hat{\theta}_{el}$ | 0.994431 | 0.018080 | 0.005569 | 1.025735 | 0.017843 | 0.025735 |
| | $\hat{\theta}_{g_{1,.25}}$ | 1.023941 | 0.005274 | 0.023941 | 0.993490 | 0.006180 | 0.006510 |
| | $\hat{\theta}_{g_{1,.5}}$ | 1.023286 | 0.005305 | 0.023286 | 0.992716 | 0.006169 | 0.007284 |
| | $\hat{\theta}_{g_{1,.75}}$ | 1.022630 | 0.005336 | 0.022630 | 0.991943 | 0.006157 | 0.008057 |
| | $\hat{\theta}_{g_{2,.25}}$ | 1.003335 | 0.006091 | 0.003335 | 0.993486 | 0.006180 | 0.006514 |
| | $\hat{\theta}_{g_{2,.5}}$ | 1.002574 | 0.006095 | 0.002574 | 0.992711 | 0.006169 | 0.007289 |
| | $\hat{\theta}_{g_{2,.75}}$ | 1.001813 | 0.006099 | 0.001813 | 0.991938 | 0.006157 | 0.008062 |
| | $\hat{\theta}_{g_{3,.25}}$ | 1.023929 | 0.005275 | 0.023929 | 0.993493 | 0.006180 | 0.006507 |
| | $\hat{\theta}_{g_{3,.5}}$ | 1.023271 | 0.005306 | 0.023271 | 0.992722 | 0.006169 | 0.007278 |
| | $\hat{\theta}_{g_{3,.75}}$ | 1.022619 | 0.005336 | 0.022619 | 0.991948 | 0.006157 | 0.008052 |

using MCMC method in Table 3. From these tables, we can observe that the posterior risk and absolute bias under Lindley method are lesser than that of MCMC method. Hence, we can conclude that Lindley approximation performs slightly better than MCMC method,

**Table 3: The Bayesian estimates and respective simulated absolute bias and posterior risk for $\theta = 1$ under Jeffreys prior and gamma prior using MCMC method**

| n | Estimator | Jeffreys prior | | | Gamma prior | | |
|---|---|---|---|---|---|---|---|
| | | Estimate | Posterior risk | Absolute bias | Estimate | Posterior risk | Absolute bias |
| 30 | $\hat{\theta}_s$ | 1.024823 | 0.023535 | 0.024823 | 1.080640 | 0.029746 | 0.080640 |
| | $\hat{\theta}_p$ | 1.036099 | 0.022551 | 0.036099 | 1.094007 | 0.026734 | 0.094007 |
| | $\hat{\theta}_l$ | 1.013464 | 0.011360 | 0.013464 | 1.055327 | 0.014247 | 0.055327 |
| | $\hat{\theta}_e$ | 1.003360 | 0.010495 | 0.003360 | 1.066392 | 0.011725 | 0.066392 |
| | $\hat{\theta}_{el}$ | 1.013918 | 0.021217 | 0.013918 | 1.067756 | 0.023725 | 0.067756 |
| | $\hat{\theta}_{g_{1,.25}}$ | 1.033283 | 0.022785 | 0.033283 | 1.090669 | 0.027430 | 0.090669 |
| | $\hat{\theta}_{g_{1,.5}}$ | 1.030469 | 0.023027 | 0.030469 | 1.087332 | 0.028163 | 0.087332 |
| | $\hat{\theta}_{g_{1,.75}}$ | 1.027651 | 0.023277 | 0.027651 | 1.083991 | 0.028932 | 0.083991 |
| | $\hat{\theta}_{g_{2,.25}}$ | 1.033354 | 0.022749 | 0.033354 | 1.090895 | 0.027353 | 0.090895 |
| | $\hat{\theta}_{g_{2,.5}}$ | 1.030568 | 0.022978 | 0.030568 | 1.087652 | 0.028043 | 0.087652 |
| | $\hat{\theta}_{g_{2,.75}}$ | 1.027728 | 0.023240 | 0.027728 | 1.084244 | 0.028836 | 0.084244 |
| | $\hat{\theta}_{g_{3,.25}}$ | 1.033194 | 0.022821 | 0.033194 | 1.090403 | 0.027520 | 0.090403 |
| | $\hat{\theta}_{g_{3,.5}}$ | 1.030370 | 0.023075 | 0.030370 | 1.087019 | 0.028284 | 0.087019 |
| | $\hat{\theta}_{g_{3,.75}}$ | 1.027591 | 0.023313 | 0.027591 | 1.083782 | 0.029025 | 0.083782 |
| 50 | $\hat{\theta}_s$ | 1.000461 | 0.012736 | 0.000461 | 1.058818 | 0.015640 | 0.058818 |
| | $\hat{\theta}_p$ | 1.006735 | 0.012548 | 0.006735 | 1.066092 | 0.014547 | 0.066092 |
| | $\hat{\theta}_l$ | 0.994210 | 0.006252 | 0.005791 | 1.044764 | 0.007634 | 0.044764 |
| | $\hat{\theta}_e$ | 0.988231 | 0.006116 | 0.011769 | 1.051185 | 0.006640 | 0.051185 |
| | $\hat{\theta}_{el}$ | 0.994293 | 0.012300 | 0.005707 | 1.051711 | 0.013385 | 0.051711 |
| | $\hat{\theta}_{g_{1,.25}}$ | 1.005168 | 0.012589 | 0.005168 | 1.064275 | 0.014805 | 0.064275 |
| | $\hat{\theta}_{g_{1,.5}}$ | 1.003601 | 0.012634 | 0.003601 | 1.062458 | 0.015073 | 0.062458 |
| | $\hat{\theta}_{g_{1,.75}}$ | 1.002032 | 0.012683 | 0.002032 | 1.060640 | 0.015351 | 0.060640 |
| | $\hat{\theta}_{g_{2,.25}}$ | 1.005177 | 0.012576 | 0.005177 | 1.064361 | 0.014780 | 0.064361 |
| | $\hat{\theta}_{g_{2,.5}}$ | 1.003615 | 0.012616 | 0.003615 | 1.062578 | 0.015038 | 0.062578 |
| | $\hat{\theta}_{g_{2,.75}}$ | 1.002045 | 0.012669 | 0.002045 | 1.060734 | 0.015324 | 0.060734 |
| | $\hat{\theta}_{g_{3,.25}}$ | 1.005151 | 0.012603 | 0.005151 | 1.064177 | 0.014830 | 0.064177 |
| | $\hat{\theta}_{g_{3,.5}}$ | 1.003586 | 0.012652 | 0.003586 | 1.062339 | 0.015107 | 0.062339 |
| | $\hat{\theta}_{g_{3,.75}}$ | 1.002027 | 0.012697 | 0.002027 | 1.060559 | 0.015376 | 0.060559 |
| 100 | $\hat{\theta}_s$ | 0.999136 | 0.006171 | 0.000864 | 1.026637 | 0.006858 | 0.026637 |
| | $\hat{\theta}_p$ | 1.002999 | 0.006126 | 0.002999 | 1.029954 | 0.006634 | 0.029954 |
| | $\hat{\theta}_l$ | 0.996880 | 0.003055 | 0.003119 | 1.020079 | 0.003399 | 0.020079 |
| | $\hat{\theta}_e$ | 0.993901 | 0.003018 | 0.006099 | 1.023238 | 0.003197 | 0.023238 |
| | $\hat{\theta}_{el}$ | 0.996901 | 0.006054 | 0.003097 | 1.023345 | 0.006409 | 0.023345 |
| | $\hat{\theta}_{g_{1,.25}}$ | 1.002234 | 0.006136 | 0.002234 | 1.029125 | 0.006688 | 0.029125 |
| | $\hat{\theta}_{g_{1,.5}}$ | 1.001468 | 0.006146 | 0.001468 | 1.028296 | 0.006743 | 0.028296 |
| | $\hat{\theta}_{g_{1,.75}}$ | 1.000702 | 0.006158 | 0.000702 | 1.027467 | 0.006800 | 0.027467 |
| | $\hat{\theta}_{g_{2,.25}}$ | 1.002236 | 0.006132 | 0.002236 | 1.029144 | 0.006684 | 0.029144 |
| | $\hat{\theta}_{g_{2,.5}}$ | 1.001471 | 0.006142 | 0.001471 | 1.028322 | 0.006738 | 0.028322 |
| | $\hat{\theta}_{g_{2,.75}}$ | 1.000705 | 0.006154 | 0.000705 | 1.027487 | 0.006796 | 0.027487 |
| | $\hat{\theta}_{g_{3,.25}}$ | 1.002230 | 0.006140 | 0.002230 | 1.029105 | 0.006692 | 0.029105 |
| | $\hat{\theta}_{g_{3,.5}}$ | 1.001465 | 0.006151 | 0.001465 | 1.028271 | 0.006749 | 0.028271 |
| | $\hat{\theta}_{g_{3,.75}}$ | 1.000701 | 0.006161 | 0.000701 | 1.027450 | 0.006804 | 0.027449 |

which supports the findings of Yilmaz et al. (2021). As the sample size increases, both methods exhibit nearly identical performance. When comparing Jeffreys prior and gamma prior, Jeffreys prior outperforms the gamma prior for the parameter $\theta$. The posterior risk

**Table 4: Table for estimate, posterior risk and absolute bias for various $\theta$ under Jeffreys prior and each loss functions**

| Loss function | Parameter | Estimate | Posterior Risk | Absolute Bias |
|---|---|---|---|---|
| $L_s(\theta,\hat{\theta})$ | 0.1 | 0.15271 | 0.00086 | 0.05273 |
| $L_p(\theta,\hat{\theta})$ | | 0.15542 | 0.00536 | 0.05542 |
| $L_l(\theta,\hat{\theta})$ | | 0.15231 | 0.00043 | 0.05231 |
| $L_e(\theta,\hat{\theta})$ | | 0.14731 | 0.01827 | 0.04731 |
| $L_{el}(\theta,\hat{\theta})$ | | 0.15002 | 0.03628 | 0.05002 |
| $L_{g1}(\theta,\hat{\theta})$ | | 0.15408 | 0.00214 | 0.05408 |
| $L_{g2}(\theta,\hat{\theta})$ | | 0.15310 | 0.00147 | 0.05310 |
| $L_{g3}(\theta,\hat{\theta})$ | | 0.15455 | 0.00199 | 0.05455 |
| $L_s(\theta,\hat{\theta})$ | 0.2 | 0.24213 | 0.00150 | 0.04213 |
| $L_p(\theta,\hat{\theta})$ | | 0.24502 | 0.00578 | 0.04502 |
| $L_l(\theta,\hat{\theta})$ | | 0.24138 | 0.00075 | 0.04138 |
| $L_e(\theta,\hat{\theta})$ | | 0.23618 | 0.01262 | 0.03618 |
| $L_{el}(\theta,\hat{\theta})$ | | 0.23917 | 0.02494 | 0.03917 |
| $L_{g1}(\theta,\hat{\theta})$ | | 0.24358 | 0.00293 | 0.04358 |
| $L_{g2}(\theta,\hat{\theta})$ | | 0.24272 | 0.00236 | 0.04272 |
| $L_{g3}(\theta,\hat{\theta})$ | | 0.24443 | 0.00365 | 0.04443 |
| $L_s(\theta,\hat{\theta})$ | 0.5 | 0.54204 | 0.00243 | 0.04204 |
| $L_p(\theta,\hat{\theta})$ | | 0.54435 | 0.00461 | 0.04435 |
| $L_l(\theta,\hat{\theta})$ | | 0.54083 | 0.00122 | 0.04083 |
| $L_e(\theta,\hat{\theta})$ | | 0.53738 | 0.00485 | 0.03738 |
| $L_{el}(\theta,\hat{\theta})$ | | 0.53971 | 0.00966 | 0.03971 |
| $L_{g1}(\theta,\hat{\theta})$ | | 0.54320 | 0.00332 | 0.04320 |
| $L_{g2}(\theta,\hat{\theta})$ | | 0.54283 | 0.00314 | 0.04283 |
| $L_{g3}(\theta,\hat{\theta})$ | | 0.54357 | 0.00353 | 0.04357 |
| $L_s(\theta,\hat{\theta})$ | 2 | 1.90446 | 0.00320 | 0.09554 |
| $L_p(\theta,\hat{\theta})$ | | 1.90536 | 0.00180 | 0.09464 |
| $L_l(\theta,\hat{\theta})$ | | 1.90286 | 0.00160 | 0.09714 |
| $L_e(\theta,\hat{\theta})$ | | 1.90266 | 0.00054 | 0.09734 |
| $L_{el}(\theta,\hat{\theta})$ | | 1.90356 | 0.00108 | 0.09644 |
| $L_{g1}(\theta,\hat{\theta})$ | | 1.90491 | 0.00238 | 0.09509 |
| $L_{g2}(\theta,\hat{\theta})$ | | 1.90503 | 0.00227 | 0.09497 |
| $L_{g3}(\theta,\hat{\theta})$ | | 1.90480 | 0.00250 | 0.09520 |
| $L_s(\theta,\hat{\theta})$ | 5 | 4.76435 | 0.00382 | 0.23565 |
| $L_p(\theta,\hat{\theta})$ | | 4.76476 | 0.00081 | 0.23524 |
| $L_l(\theta,\hat{\theta})$ | | 4.76244 | 0.00191 | 0.23756 |
| $L_e(\theta,\hat{\theta})$ | | 4.76354 | 0.00009 | 0.23646 |
| $L_{el}(\theta,\hat{\theta})$ | | 4.76394 | 0.00018 | 0.23606 |
| $L_{g1}(\theta,\hat{\theta})$ | | 4.76455 | 0.00175 | 0.23545 |
| $L_{g2}(\theta,\hat{\theta})$ | | 4.76468 | 0.00133 | 0.23532 |
| $L_{g3}(\theta,\hat{\theta})$ | | 4.76443 | 0.00231 | 0.23558 |

and bias under the newly defined generalized class of loss functions fall between the posterior risk and absolute bias of squared error and precautionary loss. As $\lambda$ increases from 0 to 1, the posterior risk varies from posterior risk of precautionary loss to that of squared error

loss. The bias under the new generalized class of loss function also behaves similarly. Consequently, it is observed that when squared error loss overestimates (underestimates) and precautionary loss underestimates (overestimates) the parameters, the new generalized class of loss functions outperforms squared error loss and precautionary loss in terms of absolute bias. For a comprehensive comparison of the loss functions outlined in Section 2, we computed Bayes estimates for various values of $\theta$ and determined their corresponding posterior risk and absolute bias under these loss functions. Since Jeffreys prior demonstrates superior performance compared to the gamma prior, the results are presented exclusively for Jeffreys prior under the Lindley method, as shown in Table 4. For smaller parameter values, sqlf and linex loss functions exhibit lower posterior risk, while linex loss shows a slightly smaller risk to a certain extent. As the parameter value increases, entropy loss and El-Sayyad loss functions display lower posterior risk; however, entropy loss outperforms the El-Sayyad loss function. Additionally, the posterior risk under sqlf increases with the increase of parameter values compared to other loss functions. It's noteworthy that we are not confined to selecting the squared error loss function or the posterior mean as an estimator.

## 5. Application

To demonstrate the practical implementation of the estimation techniques delineated in this paper, we analyze a dataset provided by Nelson (1982). The dataset pertains to a life-test experiment involving specimens of a specific electrical insulating fluid exposed to a constant voltage stress of 34 KV/minutes. Nelson, in his analysis, presumed a Weibull distribution for the breakdown times. Al-Aboud (2009) explored log-breakdowns times to derive Bayesian estimates for the parameters of the Gumbel distribution. The log-breakdown times are given by,

| | | | | | | |
|---|---|---|---|---|---|---|
| -1.66073 | -0.24846 | -0.04082 | 0.27003 | 1.02245 | 1.15057 | 1.42311 |
| 1.54116 | 1.57898 | 1.87180 | 1.99470 | 2.08069 | 2.11263 | 2.48989 |
| 3.45789 | 3.48186 | 3.52371 | 3.60305 | 4.28895 | | |

The Bayes estimates and posterior risk of the scale parameter $\theta$ of Gumbel distribution under different loss functions were computed and tabulated in Table 5. It is clear from the table that posterior risk is smaller for scale parameter $\theta$ under entropy loss function for both Jeffreys prior and gamma prior. Squared error loss function has greater posterior risk in both cases. The new loss functions $L_{g1}(\theta, \hat{\theta})$, $L_{g2}(\theta, \hat{\theta})$ and $L_{g3}(\theta, \hat{\theta})$ have posterior risk smaller than that of sqlf but greater than pqlf.

## 6. Conclusion

For a better choice of loss function while estimating the scale parameters of Gumbel distribution using Bayesian approach we consider different loss functions that are available in literature. It is concluded that entropy loss function has smaller posterior risk among other loss functions. El-Sayyad loss function also performs in similar way. In most of the research work squared error loss function is considered and posterior mean will be the estimator. However here we shows that squared error loss function has greater posterior risk when the parameter values are moderately large. The generalized class of loss function has smaller bias when squared error loss function overestimate(underestimate) and precautionary

**Table 5: The Bayesian estimates and posterior risk for $\theta$ under Jeffreys prior and gamma prior using for real data**

| | Jeffreys prior | | Gamma prior | |
|---|---|---|---|---|
| Estimator | Estimate | Posterior risk | Estimate | Posterior risk |
| $L_s(\theta, \hat{\theta})$ | 1.1572 | 0.1958 | 1.1436 | 0.2068 |
| $L_p(\theta, \hat{\theta})$ | 1.0692 | 0.1759 | 1.0492 | 0.1886 |
| $L_l(\theta, \hat{\theta})$ | 1.2336 | 0.0764 | 1.2238 | 0.0802 |
| $L_e(\theta, \hat{\theta})$ | 1.2538 | 0.0327 | 1.2450 | 0.0344 |
| $L_{el}(\theta, \hat{\theta})$ | 1.2134 | 0.0771 | 1.2028 | 0.0815 |
| $L_{g_1}(\theta, \hat{\theta})$ | 1.1136 | 0.1838 | 1.0969 | 0.1954 |
| $L_{g_2}(\theta, \hat{\theta})$ | 1.1113 | 0.1835 | 1.0948 | 0.1952 |
| $L_{g_3}(\theta, \hat{\theta})$ | 1.1161 | 0.1840 | 1.0992 | 0.1956 |

loss function underestimate(overestimate) the parameters. The class can be enlarged by considering other loss functions in the place of sqlf and pqlf.

## Acknowledgements

The authors would like to thank the Editor and reviewers for their valuable comments and suggestions, which helped improve the quality of this article.

## Conflict of interest

The authors do not have any financial or non-financial conflict of interest to declare for the research work included in this article.

## References

Abbas, P., Ghitany, M. E., and Mahmoudi, M. R. (2019). Bayesian inference on power Lindley distribution based on different loss functions. *Brazilian Journal of Probability and Statistics*, **33**, 894–914.

Al-Aboud, F. M. (2009). Bayesian estimations for the extreme value distribution using progressive censored data and asymmetric loss. *International Mathematical Forum*, **4**, 1603–1622.

Berger, J. O. (2013). *Statistical Decision Theory and Bayesian Analysis*. Springer.

Calabria, R. and Pulcini, G. (1994). An engineering approach to Bayes estimation for the Weibull distribution. *Microelectronics Reliability*, **34**, 789–802.

Chechile, R. A. (2001). Bayesian analysis of Gumbel distributed data. *Communications in Statistics-Theory and Methods*, **30**, 485–496.

Coles, S. (2001). *An Introduction to Statistical Modeling of Extreme Values*. Springer.

Coles, S. G. and Powell, E. A. (1996). Bayesian methods in extreme value modelling: a review and new developments. *International Statistical Review/Revue Internationale de Statistique*, **64**, 119–136.

Coles, S. G. and Tawn, J. A. (1996). A Bayesian analysis of extreme rainfall data. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, **45**, 463–478.

Dey, D. K., Ghosh, M., and Srinivasan, C. (1986). Simultaneous estimation of parameters under entropy loss. *Journal of Statistical Planning and Inference*, **15**, 347–363.

El-Sayyad, G. (1967). Estimation of the parameter of an exponential distribution. *Journal of the Royal Statistical Society: Series B (Methodological)*, **29**, 525–532.

Embrechts, P., Klüppelberg, C., and Mikosch, T. (2013). *Modelling Extremal Events: for Insurance and Finance*. Springer.

Hastings, W. K. (1970). Monte carlo sampling methods using markov chains and their applications. *Biometrica*, **57**, 97–109.

Hosking, J. R. (1985). Algorithm as 215: Maximum-likelihood estimation of the parameters of the generalized extreme-value distribution. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, **34**, 301–310.

Hosking, J. R. (1990). L-moments: Analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Society: Series B (Methodological)*, **52**, 105–124.

Hosking, J. R. M., Wallis, J. R., and Wood, E. F. (1985). Estimation of the generalized extreme-value distribution by the method of probability-weighted moments. *Technometrics*, **27**, 251–261.

Jenkinson, A. F. (1955). The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Quarterly Journal of the Royal Meteorological Society*, **81**, 158–171.

Landwehr, J. M., Matalas, N., and Wallis, J. (1979). Probability weighted moments compared with some traditional techniques in estimating Gumbel parameters and quantiles. *Water Resources Research*, **15**, 1055–1064.

Leadbetter, M. R., Lindgren, G., and Rootzén, H. (1983). *Extremes and Related Properties of Random Sequences and Processes*. Springer.

Lindley, D. V. (1980). Approximate bayesian methods. *Trabajos de Estadística y de Investigación Operativa*, **31**, 223–245.

Mahdi, S. and Cenac, M. (2005). Estimating parameters of gumbel distribution using the methods of moments, probability weighted moments and maximum likelihood. *Revista de Matemática: Teoría y Aplicaciones*, **12**, 151–156.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953). Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, **21**, 1087–1092.

Mousa, M. A., Jaheen, Z., and Ahmad, A. (2002). Bayesian estimation, prediction and characterization for the Gumbel model based on records. *Statistics: A Journal of Theoretical and Applied Statistics*, **36**, 65–74.

Nelson, W. B. (1982). *Applied Life Data Analysis*. John Wiley & Sons.

Norstrom, J. G. (1996). The use of precautionary loss functions in risk analysis. *IEEE Transactions on Reliability*, **45**, 400–403.

Phien, H. N. (1987). A review of methods of parameter estimation for the extreme value type-1 distribution. *Journal of Hydrology*, **90**, 251–268.

Prescott, P. and Walden, A. (1983). Maximum likeiihood estimation of the parameters of the three-parameter generalized extreme-value distribution from censored samples. *Journal of Statistical Computation and Simulation*, **16**, 241–250.

Smith, R. L. (1985). Maximum likelihood estimation in a class of nonregular cases. *Biometrika*, **72**, 67–90.

Varian, H. R. (1975). A Bayesian approach to real estate assessment. In *Studies in Bayesian Econometric and Statistics in Honor of Leonard J. Savage*, 195–208.

Vidal, I. (2014). A Bayesian analysis of the Gumbel distribution: an application to extreme rainfall data. *Stochastic Environmental Research and Risk Assessment*, **28**, 571–582.

Von Mises, R. (1964). La distribution de la plus grande de n valeurs. reproduced, selected papers of von mises, r. *American Mathematical Society*, **2**, 271–294.

Yılmaz, A., Kara, M., and Özdemir, O. (2021). Comparison of different estimation methods for extreme value distribution. *Journal of Applied Statistics*, **48**, 2259–2284.

Zellner, A. (1986). Bayesian estimation and prediction using asymmetric loss functions. *Journal of the American Statistical Association*, **81**, 446–451.

## ANNEXURE

### Annexure A. Derivation for the Bayesian estimator under the loss $\mathbf{L_{g1}}(\theta, \hat{\theta})$

Consider the case $g(z) = g_1(z) = \log z$ then the loss function has the form

$$L_{g_1}(\theta, \hat{\theta}) = \frac{(\theta - \hat{\theta})^2}{\hat{\theta}^{1-\lambda}}.$$

Here we derive the Bayes estimator under this loss by minimizing the posterior expected loss. The posterior expected loss is then given by,

$$E(L_{g_1}(\theta, \hat{\theta}) \mid \mathbf{x}) = \frac{1}{\hat{\theta}^{1-\lambda}} E(\theta^2 \mid \mathbf{x}) - 2\hat{\theta}^{\lambda} E(\theta \mid \mathbf{x}) + \hat{\theta}^{1+\lambda}.$$

Let $u = E(\theta \mid \mathbf{x})$ and $v = E(\theta^2 \mid \mathbf{x})$ then the above equation can be written as

$$E(L_{g_1}(\theta, \hat{\theta}) \mid \mathbf{x}) = \frac{1}{\hat{\theta}^{1-\lambda}} v - 2\hat{\theta}^{\lambda} u + \hat{\theta}^{1+\lambda}.$$

The Bayesian parameter estimator $\hat{\theta}$, which minimizes the posterior expected loss is then the solution of the equation,

$$\frac{d}{d\hat{\theta}} E(L_{g_1}(\theta, \hat{\theta}) \mid \mathbf{x}) = 0.$$

Implies that,

$$v(\lambda - 1)\hat{\theta}^{\lambda-2} - 2u\lambda\hat{\theta}^{\lambda-1} + (\lambda + 1)\hat{\theta}^{\lambda} = 0.$$

Multiplying by $\hat{\theta}^{2-\lambda}$,

$$v(\lambda - 1) - 2u\lambda\hat{\theta} + (\lambda + 1)\hat{\theta}^2 = 0$$

which is a quadratic equation in $\hat{\theta}$ and the two solutions for this equation are

$$\hat{\theta} = \frac{u\lambda + \sqrt{v + u^2\lambda^2 - v\lambda^2}}{\lambda + 1},$$

and,

$$\hat{\theta} = \frac{u\lambda - \sqrt{v + u^2\lambda^2 - v\lambda^2}}{\lambda + 1}.$$

For the Gumbel distribution, the second case is not feasible. So we consider the estimator $\hat{\theta}_{g_1}$ under the loss $L_{g_1}(\theta, \hat{\theta})$ as

$$\hat{\theta}_{g_1} = \frac{u\lambda + \sqrt{v + u^2\lambda^2 - v\lambda^2}}{\lambda + 1}.$$

**Annexure B. Derivation for the Bayesian estimator under the loss $\mathbf{L_{g_2}}(\theta, \hat{\theta})$**

For $g(z) = g_2(z) = 1/z$ and the loss function $L_{g_2}(\theta, \hat{\theta})$ is given as

$$L_{g_2}(\theta, \hat{\theta}) = \frac{(\theta - \hat{\theta})^2}{\hat{\theta}(1-\lambda) + \lambda}.$$

The posterior expected loss

$$E(L_{g_2}(\theta, \hat{\theta}) \mid \mathbf{x}) = \frac{1}{\hat{\theta}(1-\lambda) + \lambda} E(\theta^2 \mid x) - 2\hat{\theta}E(\theta \mid x) + \hat{\theta}^2.$$

Letting $u = E(\theta \mid x)$ and $v = E(\theta^2 \mid x)$, we have,

$$E(L_{g_2}(\theta, \hat{\theta}) \mid \mathbf{x}) = \frac{1}{\hat{\theta}(1-\lambda) + \lambda} v - 2\hat{\theta}u + \hat{\theta}^2.$$

Setting first derivative equal to zero

$$\frac{d}{d\hat{\theta}} E(L_{g_2}(\theta, \hat{\theta}) \mid \mathbf{x}) = 0.$$

Implies,

$$\frac{2\hat{\theta} - 2u}{\hat{\theta}(1-\lambda) + \lambda} - \frac{(v - 2\hat{\theta}u + \hat{\theta^2})(1-\lambda)}{(\hat{\theta}(1-\lambda) + \lambda)^2} = 0.$$

That is,

$$\hat{\theta}^2(1-\lambda) + 2\hat{\theta}\lambda + v(\lambda - 1) - 2u\lambda = 0.$$

This is again a quadratic equation in $\hat{\theta}$ and the two solutions are,

$$\hat{\theta} = \frac{\lambda + \sqrt{v + 2u\lambda - 2v\lambda + \lambda^2 - 2u\lambda^2 + v\lambda^2}}{\lambda - 1}$$

and

$$\hat{\theta} = \frac{\lambda - \sqrt{v + 2u\lambda - 2v\lambda + \lambda^2 - 2u\lambda^2 + v\lambda^2}}{\lambda - 1}.$$

Here the first case is not feasible for the Gumbel parameters. Then we consider the estimator $\hat{\theta_{g_2}}$ under the loss function $L_{g_2}(\theta, \hat{\theta})$ as

$$\hat{\theta_{g_2}} = \frac{\lambda - \sqrt{v + 2u\lambda - 2v\lambda + \lambda^2 - 2u\lambda^2 + v\lambda^2}}{\lambda - 1}.$$

**Annexure C. Derivation for the Bayesian estimator under the loss $\mathbf{L_{g_3}}(\theta, \hat{\theta})$**

Now for $g(z) = z$ and,

$$\begin{aligned} L_{g_3}(\theta, \hat{\theta}) &= \lambda(\theta - \hat{\theta})^2 + (1-\lambda)\frac{(\theta - \hat{\theta})^2}{\hat{\theta}} \\ &= \lambda\theta^2 - 2\lambda\theta\hat{\theta} + \lambda\hat{\theta}^2 + \frac{\theta^2}{\hat{\theta}} - 2\theta + \hat{\theta} - \frac{\lambda\theta^2}{\hat{\theta}} + 2\lambda\theta - \lambda\hat{\theta}. \end{aligned}$$

Then posterior expected loss can be find as,

$E(L_{g_3}(\theta, \hat{\theta}) \mid x) = \lambda E(\theta^2 \mid x) - 2\lambda E(\theta \mid x)\hat{\theta} + \lambda\hat{\theta}^2 + \frac{1}{\hat{\theta}}E(\theta^2 \mid x) - 2E(\theta \mid x) + \hat{\theta} - \frac{\lambda}{\hat{\theta}}E(\theta^2 \mid x) + 2\lambda E(\theta \mid x) - \lambda\hat{\theta}.$

Substituting $u = E(\theta \mid x)$ and $v = E(\theta^2 \mid x)$ we have,

$$E(L_{g_3}(\theta, \hat{\theta}) \mid x) = \lambda v - 2\lambda u\hat{\theta} + \lambda\hat{\theta}^2 + \frac{v}{\hat{\theta}} - 2u + \hat{\theta} - \frac{\lambda}{\hat{\theta}}v + 2\lambda u - \lambda\hat{\theta}.$$

To find the Bayes estimator under the loss function $L_{g_3}(\theta, \hat{\theta})$ that minimize the posterior expected loss we equate first derivative with respect to $\hat{\theta}$ equal to zero, i.e,

$$\frac{d}{d\hat{\theta}}E(L_{g_3}(\theta, \hat{\theta}) \mid x) = -2\lambda u + 2\lambda\hat{\theta} - \frac{v}{\hat{\theta}^2} + 1 + \frac{\lambda v}{\hat{\theta}^2} - \lambda = 0.$$

Multiplying both side by $\hat{\theta}^2$ we get a cubic polynomial on $\hat{\theta}$ as,

$$2\lambda\hat{\theta}^3 + (1 - \lambda - 2\lambda u)\hat{\theta}^2 + (\lambda - 1)v = 0.$$

The three solutions for $\hat{\theta}$ are given by,

$$\begin{aligned} \hat{\theta} = \frac{1-\lambda-2u\lambda}{6\lambda} + \frac{(1-\lambda-2u\lambda)^2}{3\times 2^{2/3}\lambda\Big[H_1(u,v,\lambda) + \sqrt{H_2(u,v,\lambda) + (H_1(u,v,\lambda))^2}\Big]^{1/3}} + \\ \frac{1}{6\times 2^{1/3}\lambda}\Big[H_1(u,v,\lambda) + \sqrt{H_2(u,v,\lambda) + (H_1(u,v,\lambda))^2}\Big]^{1/3}, \end{aligned}$$

or,

$$\begin{aligned} \hat{\theta} = \frac{1-\lambda-2u\lambda}{6\lambda} + \frac{(1+i\sqrt{3})(1-\lambda-2u\lambda)^2}{6\times 2^{2/3}\lambda\Big[H_1(u,v,\lambda) + \sqrt{H_2(u,v,\lambda) + (H_1(u,v,\lambda))^2}\Big]^{1/3}} - \\ \frac{1}{12\times 2^{1/3}\lambda}(1-i\sqrt{3})\Big[H_1(u,v,\lambda) + \sqrt{H_2(u,v,\lambda) + (H_1(u,v,\lambda))^2}\Big]^{1/3}, \end{aligned}$$

or,

$$\begin{aligned} \hat{\theta} = \frac{1-\lambda-2u\lambda}{6\lambda} + \frac{(1-i\sqrt{3})(1-\lambda-2u\lambda)^2}{6\times 2^{2/3}\lambda\Big[H_1(u,v,\lambda) + \sqrt{H_2(u,v,\lambda) + (H_1(u,v,\lambda))^2}\Big]^{1/3}} - \\ \frac{1}{12\times 2^{1/3}\lambda}(1+i\sqrt{3})\Big[H_1(u,v,\lambda) + \sqrt{H_2(u,v,\lambda) + (H_1(u,v,\lambda))^2}\Big]^{1/3}, \end{aligned}$$

where,

$$H_1(u,v,\lambda) = -2 + 6\lambda + 12u\lambda - 6\lambda^2 - 24u\lambda^2 - 24u^2\lambda^2 + 108v\lambda^2 + 2\lambda^3 + 12u\lambda^3 + 24u^2\lambda^3 + 16u^3\lambda^3 - 108v\lambda^3,$$

$$H_2(u,v,\lambda) = -4(1-\lambda-2u\lambda)^6.$$

Here two of these solutions are complex conjugates. Since our parameter space is restricted to real, we can only assume real values for $\hat{\theta}$ so the estimator $\hat{\theta_{g_3}}$ under the loss $L_{g_3}(\theta,\hat{\theta})$ can be taken as

$$\hat{\theta} = \frac{1-\lambda-2u\lambda}{6\lambda} + \frac{(1-\lambda-2u\lambda)^2}{3\times 2^{2/3}\lambda\Big[H_1(u,v,\lambda)+\sqrt{H_2(u,v,\lambda)+(H_1(u,v,\lambda))^2}\Big]^{1/3}} + \frac{1}{6\times 2^{1/3}\lambda}\Big[H_1(u,v,\lambda)+\sqrt{H_2(u,v,\lambda)+(H_1(u,v,\lambda))^2}\Big]^{1/3}.$$