Hacettepe Journal of Mathematics & Statistics

Hacet. J. Math. Stat.
Volume 53 (1) (2024), 251 – 268
DOI : 10.15672/hujms.1053432

Research Article

Non-informative Bayesian estimation in dispersion models

Ibrahim Sadok*[1], Mourad Zribi[2], Afif Masmoudi[3]

[1] *Department of Mathematics and Computer Science, Faculty of Exact Sciences, University of Bechar, Bechar, Algeria*

[2] *Laboratoire d'Informatique Signal et Image de la Côte d'Opale (LISIC- EA 4491), ULCO, 50 Rue Ferdinand Buisson BP 719, 62228 Calais Cedex, France*

[3] *Laboratory of Probability and Statistics, Sfax Faculty of Sciences, University of Sfax, Sfax, Tunisia*

Abstract

The estimation of parameters for a distribution function is a significant and prominent field within statistical inference. This particular problem holds great relevance in various domains, including industries, stock markets, image processing, and reliability studies. There are two recognized approaches to estimation: point estimation and interval estimation, also known as confidence intervals. In this study, our primary focus lies in the point estimation of parameters associated with an exponential dispersion distribution function. In this process, we consider one of the parameters as a random variable that requires estimation. To tackle this, we adopt a Bayesian inference approach utilizing a one-parameter dispersion distribution. We explore non-informative priors, such as uniform and Jeffrey's priors, and provide evidence of the effectiveness of our method through simulation studies.

Mathematics Subject Classification (2020). 62F15, 62F10

Keywords. Bayesian estimation, exponential dispersion distribution, non-informative prior, uniform prior, Jeffrey's prior

1. Introduction

The exponential dispersion model (EDM) represents a family of distributions characterized by two parameters, encompassing a linear exponential family along with an additional dispersion parameter. These models hold significant importance within the field of statistics, primarily as the response distributions for generalized linear models [21]. Over time, EDMs have emerged as a distinct area of study, extensively explored in terms of their properties by [15, 16]. The role of EDMs extends to probability and statistics, with notable applications in image processing [24]. One of the key advantages of EDMs lies in their versatility, offering a diverse range of probability distributions that find practical utility, such as Gaussian, Gamma, inverse Gaussian, and more. The determination of exponential dispersion distribution (EDD) parameters serves as a pivotal topic within the realm of statistical methodology, as evidenced by [23]. The adoption of Bayesian approaches in

*Corresponding Author.
Email addresses: ibrahim.sadok@univ-bechar.dz (I. Sadok), mourad.zribi@univ-littoral.fr (M. Zribi), afif.masmoudi@fss.rnu.tn (A. Masmoudi)
Received: 06.01.2022; Accepted: 11.08.2023

contemporary statistical analysis has gained significant traction, finding practical utility across diverse scientific domains.

By integrating prior knowledge into the analysis, the Bayesian methodology enables the process of refining prior beliefs in light of current data. The parameters to be estimated, denoted as θ, are characterized, in part, by expressing prior beliefs through the measure $\pi(\theta)$. In the presence of an observed sample y, the posterior density of θ can be represented as $p(\theta|y) \propto p(y|\theta)\pi(\theta)$, wherein the term $p(y|\theta)$ denotes the likelihood function of the econometric model under estimation. Employing the principles of Bayes' method, this equation facilitates the generation of the posterior distribution $p(\theta|y)$ for the parameter θ taking into account both the prior information and the econometric model at hand. In contemporary practice, the Bayesian approach has gained extensive popularity as a means of parameter estimation. Particularly, its application in the analysis of failure time data has garnered considerable attention. By integrating prior knowledge regarding the parameters and effectively assimilating the available data, this method allows for a comprehensive analysis. When there exists informative prior knowledge pertaining to the parameter, its utilization becomes appropriate and well-founded. Nevertheless, when confronted with scenarios lacking any prior knowledge about the parameter and inaccessible vital information from experts, a non-informative prior emerges as a viable substitute [7, 10].

Within our research, our primary focus lies in the domain of Bayesian inference, specifically targeting the estimation of unknown parameters belonging to the EDD. In the realm of Bayesian inference, the selection of appropriate priors has been a matter of great importance and extensive discourse. The concept of a prior distribution proves valuable when we possess prior knowledge concerning the unknown parameter. However, in numerous practical scenarios, our prior information is severely constrained. Hence, it becomes prudent to contemplate the utilization of a non-informative prior.

The uniform distribution stands out as one of the frequently employed non-informative priors, particularly when confronting parameters that possess a bounded support (e.g., $\pi(\theta) = \mathbb{I}_{[a,b]}(\theta)$ as expounded in [11]). However, this uniform prior lacks the desirable attribute of invariance under smooth one-to-one transformations. To tackle this concern, H. Jeffreys proposed an alternative prior, as elucidated in [14]. This alternative prior is derived by calculating the square root of the determinant of the Fisher information matrix $I(\theta)$ and is expressed as $\pi(\theta) = \sqrt{\det I(\theta)}$. This new prior has the desirable property of being invariant under smooth one-to-one transformations. But, it has some limitations when there are nuisance parameters present [1].

The main objective of this paper is to propose non-informative priors with a uniform distribution for a parameter of EDD with unbounded support $(\pi(\lambda) = \mathbb{I}(0, +\infty)(\lambda))$. Our results show that the posterior distribution of the dispersion parameter exists and converges towards the Gamma distribution $(\lambda|y_1, ..., y_N \sim Ga\left(\frac{N}{2}+1, \frac{1}{2}\sum_{i=1}^{N}\frac{(y_i-\mu_i)^2}{V(\mu_i)}\right))$, given the condition $\left(\int_0^{+\infty}\lambda^s\prod_{i=1}^{N}f(y_i;\lambda,\mu)\,d\lambda < +\infty\right)$. Additionally, we discuss the use of non-informative priors and Jeffrey's priors in light of our study. The performance of the proposed prior distributions is evaluated through a simulation study, and their performance for EDD is compared under different loss functions [7, 19].

The rest of the paper is arranged as follows: Section 2 exhibits an overview of exponential dispersion models. A concise description of the Bayesian estimator with informative prior is displayed in Section 3. Section 4 provides Bayesian analysis of the dispersion parameter under non-informative priors. Unbounded uniform and Jeffreys priors are discussed. The results are illustrated simulated data specific models in Section 5. Lastly, some conclusions and prospects are revealed in Section 6.

2. Exponential dispersion distributions

In the following section, we shall provide a brief review of important characteristics of exponential dispersion models according to [15].

Exponential Dispersion Models are based on Natural Exponential Families [24]. They are a general class of models, defined by their probability density function

$$f(y;\theta,\lambda)=e^{\lambda[\theta y-K_\nu(\theta)]}c(y,\lambda),\quad y\in\mathbb{R} \tag{2.1}$$

where λ is the dispersion parameter and θ is called the canonical parameter, with domain $(\lambda,\theta)\in\Lambda\times\Theta\subseteq\mathbb{R}_+\times\mathbb{R}$. $K_\nu(\theta)=\log\int_{\mathbb{R}}e^{\theta x}\nu(dx)$ is a known function called the cumulant function of a generating probability measure ν (not Dirac) and $c(y,\lambda)$ is a constant that ensures equation (2.1) is a probability function.

For EDMs we have some well-known relations, if $Y\sim f(.;\theta,\lambda)$, then $\mu=\mathbb{E}(Y)=K'_\nu(\theta)$ is the expectation of Equation (2.1) due to the relationship or map between θ and μ. The variance of Equation (2.1) is $Var(Y)=\dfrac{1}{\lambda}V(\mu)$ and $V(\mu)$ being the variance function which uniquely corresponds to an exponential dispersion model [23]. Let defined $\psi_\nu(\mu)=(K'_\nu(\theta))^{-1}$ and $V(\mu)=K''_\nu(\psi_\nu(\mu))$. It can also be shown that when the functions $K_\nu(.)$ and $c(.,.)$ as well as ψ_ν are fixed, the subfamily arising from taking different θ consists of elements that are all Esscher-transforms of each other. A family with K_ν, c and θ fixed and varying ψ_ν can be generated by the operation of taking sample means. For further information, we refer the reader to [18].

In what follows, we assume that the generating measure ν is infinitely divisible (*i.e.*, $\Lambda(\nu)=(0,+\infty)$) and absolutely continuous with density $c(y,1)$ with respect to the Lebesgue measure [15].

In Table 1, we present necessary details of absolutely continuous PDFs of the EDM family specifying the normalizing constant ($c(y,\lambda)$), the cumulant function (K_ν), canonical parameter (θ), dispersion parameter (λ), mean (K'_ν), inverse function of the mean (ψ_ν) and variance function (V) of each distribution.

Table 1. Examples of some absolutely continuous PDF of EDMs.

	Gaussian	**Gamma**	**Inverse Gaussian**	**Laplace**
$c(y,\lambda)$	$\frac{\sqrt{\lambda}}{\sqrt{2\pi}}e^{-\frac{\lambda y^2}{2}}$	$\frac{\lambda^\lambda y^{\lambda-1}}{\Gamma(\lambda)}$	$\frac{\sqrt{\lambda}}{\sqrt{2\pi}}y^{-\frac{3}{2}}e^{-\frac{\lambda}{2y}}$	$\frac{\lambda e^{\lambda y}}{\Gamma(\lambda)^2}\int_{\lambda y}^{+\infty}e^{-2t}t^{\lambda-1}(t-\lambda y)^{\lambda-1}dt$
K_ν	$\frac{\theta^2}{2}$	$-\log(-\theta)$	$-\sqrt{-2\theta}$	$-\log(1-\theta^2)$
K'_ν	θ	$-\frac{1}{\theta}$	$(-2\theta)^{-1/2}$	$\frac{2\theta}{1-\theta^2}$
ψ_ν	μ	$-\frac{1}{\mu}$	$-\frac{1}{2\mu^2}$	$\frac{\left\lvert\sqrt{1+\mu^2}-1\right\rvert}{\mu}$
V	1	μ^2	μ^3	$\frac{\left\lvert\sqrt{1+\mu^2}-1\right\rvert}{\mu^2\sqrt{1+\mu^2}}$

In this stage, we can introduce an important approximation of the probability density function which is called the saddlepoint approximation [15] for dispersion models for the distribution of y that is significant and useful in the asymptotic theory of the general linear model.

Let us consider a continuous reproductive exponential dispersion model $Y\sim f(.;\theta,\lambda)$, we thus obtain the approximation for the density of Y, for some λ large by the following formula:

$$f(y;\mu,\lambda)=\lambda^{\frac{1}{2}}(2\pi V(y))^{-\frac{1}{2}}e^{\lambda[\psi(\mu)y-K(\psi(\mu))]}\quad\text{when}\quad\lambda\mapsto+\infty. \tag{2.2}$$

3. Informative Bayesian estimation of the EDD

There exist two distinct types of estimation procedures, namely point estimation and confidence interval estimation [1,3,25]. In our analysis, we will direct our attention toward point estimation, which pertains to the process of approximating a parameter associated with a distribution function by utilizing sample data governed by a specific probability distribution.

Suppose the observed data $Y = (y_1, y_2, \ldots, y_N)$ follows $f(.;\mu,\lambda)$, where μ and λ are unknown and N is the sample size. There are two approaches for obtaining a point estimator for unknown parameters: the classical method and the decision-theoretic approach. In this section, we will focus on estimating parameter λ using Pearson's method.

Our present emphasis will be directed towards the estimation of vector parameters (μ,λ) within the context of the exponential dispersion distribution. Consider a set of N independent and identically distributed observations $(y_1,\ldots,y_N)$ drawn from the probability density function $f(.;\mu,\lambda)$. The likelihood function L_N of $(y_1,\ldots,y_N)$ is given by

$$L_N\left(y_1,...,y_N;\mu,\lambda\right) = \prod_{i=1}^{N}\left[c\left(y_i,\lambda\right)e^{\lambda(\psi(\mu)y_i - K(\psi(\mu)))}\right]. \tag{3.1}$$

Due to the necessity of transforming multiplication into addition, the log-likelihood l_N is expressed as follows:

$$\begin{aligned} l_N\left(y_1,...,y_N;\mu,\lambda\right) &= \log L_N\left(y_1,...,y_N;\Theta\right) \\ &= \sum_{i=1}^{N}\left[\log\left(c\left(y_i,\lambda\right)\right) + \lambda\left(\psi(\mu)y_i - K(\psi(\mu))\right)\right]. \end{aligned}$$

In order to calculate the estimator of μ, we need to solve the following equation:

$$\begin{aligned} \frac{\partial l_N\left(y_1,...,y_N;\mu,\lambda\right)}{\partial\mu} &= \sum_{i=1}^{N}\left(\psi'(\mu)y_i - \mu\psi'(\mu)\right) \\ &= \sum_{i=1}^{N}\psi'(\mu)\left(y_i-\mu\right). \end{aligned}$$

By making it equal to zero, we obtain

$$\widehat{\mu} = \frac{1}{N}\sum_{i=1}^{N} y_i = \overline{Y}. \tag{3.2}$$

Since the log-likelihood function is strictly concave with respect to μ, then $\widehat{\mu}$ the maximum likelihood estimate for the mean μ. In general cases, the maximum likelihood estimate of $\lambda = \frac{1}{\sigma^2}$ does not exist and it is estimated by the Pearson estimator given by

$$\widehat{\lambda} = \frac{N-q}{\displaystyle\sum_{i=1}^{N}\frac{\left(y_i-\widehat{\mu}\right)^2}{V(\widehat{\mu})}}, \tag{3.3}$$

where $\widehat{\mu}$ is the estimation of the mean μ and q is the total number of unknown parameters. For more details see [15,16].

4. Non-informative Bayesian estimation of the EDD

Bayesian inference holds immense significance in contemporary statistics, particularly within the realm of mathematical statistics. Its applicability extends beyond and finds utility in diverse fields such as engineering, medicine, accounting, and image processing [3]. By employing Bayes' rule, Bayesian inference recalibrates the probability estimation of a hypothesis in light of fresh evidence. This approach incorporates both pre-existing knowledge regarding the parameters and the accessible data [1]. In the absence of prior knowledge about a parameter and inability to obtain information from experts, a non-informative prior is a suitable alternative [5, 10]. Non-informative priors have minimal impact on the posterior distribution.

The objective of this research is to examine the utilization of a non-informative prior to estimate the parameter λ of EDD, while the scale parameter μ remains known. In Bayesian analysis, λ is treated as a stochastic variable and is assigned a prior distribution. The adoption of a non-informative prior is preferred due to its versatility and capability to encompass a range of distributions.

Bayesian analysis combines prior information $\pi(\lambda)$ and sample information $(y_1,\dots,y_N)$ to form the posterior distribution of λ, given $Y = (y_1,\dots,y_N)$, from which decisions and inferences are made. $f(\lambda|Y)$ represents updated beliefs λ after observing the sample Y. This text will derive the estimation of the unknown parameter λ of EDD using non-informative priors. We will prove the existence of the posterior distribution and the expectation a posteriori (EAP) estimator of λ. The following results will be established.

4.1. Proposed method for estimating using unbounded uniform Prior

The uniform prior distribution is often used in Bayesian analysis because it produces non-informative priors and appropriate posterior distributions. This prior assigns equal weight to all possible values [11]. In this section, we propose using the uniform prior with unbounded support of the λ. Let $Y = (y_1, ..., y_N)$ be a sample of reproductive exponential dispersion distribution, the likelihood function of λ using $f(., \mu, \lambda)$, is given by

$$l\left(Y;\mu,\lambda\right)=\prod_{i=1}^{N} f\left(y_i;\mu,\lambda\right).$$

Assume that the non-informative prior for λ follows an improper unbounded uniform distribution $\pi(\lambda) = \mathbb{I}_{(0,+\infty)}(\lambda)$ and the posterior distribution of the parameter λ exists, if and only if, the integral

$$\int_0^{+\infty} l\left(Y;\mu,\lambda\right)d\lambda<+\infty$$

converges almost surely. In this case, the posterior density function $f\left(\lambda;\mu,y_{1:n}\right)$ of the parameter λ, is defined by

$$f\left(\lambda;\mu,y_{1:N}\right)=\frac{\displaystyle\prod_{i=1}^{N} f\left(y_i;\mu,\lambda\right)}{\displaystyle\int_0^{+\infty}\prod_{i=1}^{N} f\left(y_i;\mu,\lambda\right)d\lambda}. \tag{4.1}$$

If $\int_0^{+\infty}\lambda f\left(\lambda;\mu,y_{1:N}\right)d\lambda<+\infty$, then the non-informative Bayesian estimator $\hat\lambda$ of λ is given by

$$\hat\lambda=\mathbb{E}\left(\lambda|y_{1:N},\mu\right)=\int_0^{+\infty}\lambda f\left(\mu,\lambda;y_{1:N}\right)d\lambda. \tag{4.2}$$

Now, we put forward a sufficient condition such that the non-informative Bayesian estimator of the parameter λ exists. The results are as follows

Theorem 4.1. *Let $y_1, ..., y_N$ be N positive sample from a reproductive exponential dispersion model $f\left(y_{1:N};\mu,\lambda\right)$, then*

$$\int_0^{+\infty}\lambda^s\prod_{i=1}^{N} f\left(y_i;\mu,\lambda\right)d\lambda<+\infty$$

converges almost surely, for all $s \geqslant 0$. As a matter of fact, the non-informative Bayesian estimator $\hat\lambda$ of λ exists.

Proof. Before drawing in the proof of Theorem 4.1, we use Lemma 6.1 which goes back to [17] (See Appendix 6).

Since ν is concentrated on $\mathbb{R}_+$, then there exists $t_0 \in \Theta(\nu)$ such that, $(-\infty, t_0] \subset \Theta(\nu)$.

The likelihood function is given by

$$l\left(Y;\mu,\lambda\right)=\prod_{i=1}^{N}c\left(y_i,\lambda\right)e^{\lambda[\psi(\mu)y_i-K_\nu(\psi(\mu))]}.$$

In order to prove that the integral $\int_0^{+\infty}\lambda^s\prod_{i=1}^{N}f\left(y_i;\mu,\lambda\right)d\lambda$ converges, we shall firstly prove that the integral

$$I=\int_{\mathbb{R}_+^N}\int_0^{+\infty}\lambda^s e^{t\sum_{i=1}^N y_i}l\left(Y;\mu,\lambda\right)d\lambda dy_1...dy_N$$

converges for all $t \in (-\infty, \inf(t_0, 0))$. Indeed,

$$\begin{aligned} I &= \int_{\mathbb{R}_+^N}\int_0^{+\infty}\lambda^s e^{t\sum_{i=1}^N y_i}\left(\prod_{i=1}^{N}c\left(y_i,\lambda\right)e^{\lambda[\psi(\mu)y_i-K_\nu(\psi(\mu))]}\right)d\lambda dy_1...dy_N \\ &= \int_0^{+\infty}\lambda^s\left[\prod_{i=1}^{N}\int_0^{+\infty}e^{ty_i}c\left(y_i,\lambda\right)e^{\lambda[\psi(\mu)y_i-K_\nu(\psi(\mu))]}dy_i\right]d\lambda. \end{aligned}$$

Moreover, since

$$1=\int_{\mathbb{R}}e^{\lambda(\theta y-K(\theta))}c(y,\lambda)\xi(dy),$$

we notice that

$$e^{\lambda K(\theta)}=\int_{\mathbb{R}}e^{\lambda\theta y}c(y,\lambda)\xi(dy). \tag{4.3}$$

Using Equation (4.3), we get

$$\begin{aligned} \int_{\mathbb{R}}e^{ty_i}c\left(y_i,\lambda\right)e^{\lambda[\psi(\mu)y_i-K_\nu(\psi(\mu))]}dy_i &= \int_{\mathbb{R}}e^{ty_i\left[\frac{t}{\lambda}+\psi(\mu)\right]-K_\nu(\psi(\mu))}c\left(y_i,\lambda\right)dy_i \\ &= e^{\lambda\left[K_\nu\left(\frac{t}{\lambda}+\psi(\mu)\right)-K_\nu(\psi(\mu))\right]}. \end{aligned}$$

According to Lemma 6.1

$$\lim_{\lambda\longrightarrow 0^+}\frac{K_\nu\left(\frac{t}{\lambda}+\psi(\mu)\right)-K_\nu\left(\psi(\mu)\right)}{\frac{1}{\lambda}}=0, \forall t<\inf\left(t_0,0\right),$$

and note that we have

$$K_\nu\left(\frac{t}{\lambda}+\psi(\mu)\right)-K_\nu\left(\psi(\mu)\right)=K_{p(\psi(\mu),\nu)}\left(\frac{t}{\lambda}\right).$$

Therefore,

$$\lim_{\lambda\longrightarrow+\infty}\lambda^2\lambda^s e^{\lambda NK_{p(\psi(\mu),\nu)}\left(\frac{t}{\lambda}\right)}=\lim_{\lambda\longrightarrow+\infty}e^{\lambda\left[NK_{p(\psi(\mu),\nu)}\left(\frac{t}{\lambda}\right)+(s+2)\frac{\log\lambda}{\lambda}\right]}. \tag{4.4}$$

Since

$$K'_{p(\psi(\mu),\nu)}\left(\theta\right)=K'\left(\theta+\psi(\mu)\right),$$

we get

$$K'_{p(\psi(\mu),\nu)}\left(0\right)=K'\left(\psi(\mu)\right)=\mu.$$

By applying Taylor's formula of order 1, the Equation (4.4) becomes

$$\lim_{\lambda\longrightarrow+\infty}e^{tNK'_{p(\psi(\mu),\nu)}\left(\frac{\alpha(t,\lambda)t}{\lambda}\right)+(s+2)\frac{\log\lambda}{\lambda}}=e^{tN\mu}<+\infty,$$

where $0 < \alpha(t, \lambda) < 1$. Then, the integral I converges ν- a.s.

By the Fubini-Tonelli formula, we conclude that the integral

$$e^{t \sum_{i=1}^{N} y_i} \int_0^{+\infty} \lambda^s \prod_{i=1}^{N} f(y_i; \mu, \lambda)\, d\lambda < +\infty.$$

Finally, we deduce that $\int_0^{+\infty} \lambda^s \prod_{i=1}^{N} f(y_i; \mu, \lambda)\, d\lambda < +\infty$ ν- a.s and this stands for the desired result. □

Remark 4.2. If the observations y_i are not all positive, we suppose that $\lambda \in [\varepsilon, +\infty)$ and according to Equation (4.4), the integral I converges ν a.s.

Remark 4.3. Note that for $s = 0$, the non-informative Bayesian density function $f(\lambda; \mu, y_{1:N})$ exists. If $s = 1$, the non-informative Bayesian estimator $\hat{\lambda}$ exists. The conditional variance of the non-informative Bayesian estimator $V(\lambda | y_{1:N}, \mu)$ exists for $s = 2$.

According to Theorem 4.1, we have shown the existence of a parameter λ and to define this parameter we use the following proposition.

Proposition 4.4. *([15]) Let $Y = (y_1, ..., y_N) \sim f(.; \mu, \lambda)$ be a continuous exponential dispersion model renormalized saddlepoint approximation (2.2). Then,*

$$Y \xrightarrow{d} \mathcal{N}\left(\mu, \frac{V(\mu)}{\lambda}\right) \quad \textit{when} \quad \lambda \mapsto +\infty,$$

where $\xrightarrow{d}$ denotes convergence in distribution.

Theorem 4.5. *Let $y_1, ..., y_N$ be observations from $f(.; \mu, \lambda)$. Then*

(1) *The posterior distribution of λ is given by*

$$\lambda | y_1, ..., y_N \sim Ga\left(\frac{N}{2} + 1, \frac{1}{2} \sum_{i=1}^{N} \frac{(y_i - \mu_i)^2}{V(\mu_i)}\right).$$

(2) *The non-informative Bayesian estimator of λ is* $\hat{\lambda} = \dfrac{N+2}{\sum_{i=1}^{N} \frac{(y_i - \mu_i)^2}{V(\mu_i)}}$.

Proof. According to Theorem 4.1 and Proposition 4.4, the non-informative Bayesian density function $f(\mu, \lambda; y_1, ..., y_N)$ (represented by the equation (4.1)) of the parameter λ can be evaluated as

$$\begin{aligned} f(\lambda; \mu, y_1, ..., y_N) &\propto \prod_{i=1}^{N} f(y_i; \mu, \lambda) \\ &\propto \prod_{i=1}^{N} \mathcal{N}\left(\mu_i, \frac{V(\mu_i)}{\lambda}\right) \\ &\propto \lambda^{\frac{N}{2}} e^{-\frac{\lambda}{2} \sum_{i=1}^{N} \frac{(y_i - \mu_i)^2}{V(\mu_i)}} \\ &\sim Ga\left(\frac{N}{2} + 1, \frac{1}{2} \sum_{i=1}^{N} \frac{(y_i - \mu_i)^2}{V(\mu_i)}\right). \end{aligned}$$

Therefore, for λ is large enough $\lambda \gg \lambda_0$, the non-informative Bayesian density function $f(\mu, \lambda; Y)$ is only the Gamma distribution with a shape parameter $\frac{N}{2} + 1$ and a scale one

$\frac{1}{2}\sum_{i=1}^{N}\frac{(y_i-\mu_i)^2}{V(\mu_i)}$. Consequently, the non-informative Bayesian estimator $\hat{\lambda}$ of λ represented in Equation (4.2), is obtained as $\hat{\lambda} = \frac{N+2}{\sum_{i=1}^{N}\frac{(y_i-\mu_i)^2}{V(\mu_i)}}$.

□

4.2. Proposed method for estimating Jeffrey's prior

Jeffrey's prior, as discussed in the work by [5], is derived from the observed Fisher information matrix. This prior is characterized by its local uniformity, rendering it a non-informative choice. Its value lies in the fact that it remains relatively stable within regions where the likelihood holds significance, while maintaining limited influence beyond that range due to its local uniformity property. The justification for employing Jeffrey's prior stems from its invariance under parametrization, as emphasized by [26] in their study.

4.2.1. Information matrices. Jeffrey's proposed a non-informative prior, which is commonly used in cases where the parameters are poorly informed or unknown. It defines the density of the parameters as proportional to the positive square root of the Fisher information matrix.

The general class of probability density $f(y;\mu,\lambda)$ in Equation (2.1) proposed by [16] can be also written as

$$f(y;\mu,\lambda) = c(y,\lambda)e^{\lambda t(y,\mu)},$$

where $t(y,\mu) = \theta y - K_\nu(\theta)$. For a density in the dispersion model family, we have

$$\log f(y;\mu,\lambda) = \log c(y,\lambda) + \lambda t(y,\mu).$$

Since

$$0 = \mathbb{E}\left(\frac{\partial \log f(y|\theta)}{\partial\mu}\right) = \lambda\mathbb{E}\left(\frac{\partial t(y,\mu)}{\partial\mu}\right),$$

where $\theta = (\mu,\lambda)$, we have that

$$\mathbb{E}\left(\frac{\partial^2 \log f(y|\theta)}{\partial\mu\partial\lambda}\right) = \mathbb{E}\left(\frac{\partial t(y,\mu)}{\partial\mu}\right) = 0.$$

Therefore, the information matrix for dispersion models can be given by

$$I(\mu,\lambda) = \begin{pmatrix} -\lambda\mathbb{E}\left(\frac{\partial^2 t(y,\mu)}{\partial\mu^2}\right) & 0 \\ 0 & -\mathbb{E}\left(\frac{\partial^2 \log c(y,\lambda)}{\partial\lambda^2}\right)\end{pmatrix}.$$

In fact, we know more about the $I(\mu,\lambda)$ shape in the case of exponential dispersion models. From Equation (2.1), we obtain

$$\begin{aligned}\frac{\partial \log f(y|\theta,\lambda)}{\partial\mu} &= \frac{\partial \log f(y|\theta,\lambda)}{\partial\theta}\frac{\partial\theta}{\partial\mu} \\ &= \lambda(y-\mu)\left(\frac{\partial}{\partial\mu}\psi_\nu(\mu)\right).\end{aligned}$$

Hence,

$$-\mathbb{E}\left(\frac{\partial^2 \log f(y|\theta)}{\partial\mu^2}\right) = \lambda\frac{\partial}{\partial\mu}\psi_\nu(\mu) = \frac{\lambda}{V(\mu)}.$$

On the other hand, since

$$\frac{\partial \log f(y|\theta,\lambda)}{\partial\lambda} = \frac{\partial \log c(y,\lambda)}{\partial\lambda} + (\theta y - K_\nu(\theta)),$$

we get

$$-\mathbb{E}\left(\frac{\partial^2 \log f(y|\theta,\lambda)}{\partial\lambda^2}\right) = -\mathbb{E}\left(\frac{\partial^2 \log c(y,\lambda)}{\partial\lambda^2}\right).$$

Thus, the information matrix for an EDM can be written as

$$I(\mu,\lambda) = \begin{pmatrix} \dfrac{\lambda}{V(\mu)} & 0 \\ 0 & -\mathbb{E}\left(\dfrac{\partial^2 \log c(y,\lambda)}{\partial\lambda^2}\right) \end{pmatrix}.$$

It should be to note that Jeffreys prior is proportional to the square root of the determinant of the information matrix, *i.e.*, $\pi(\mu,\lambda) \propto |I_{11}I_{22}|^{\frac{1}{2}}$, where $I_{11} = \dfrac{\lambda}{V(\mu)}$ and $I_{22} = -\mathbb{E}\left(\dfrac{\partial^2 \log c(y,\lambda)}{\partial\lambda^2}\right)$.

Proposition 4.6. *Let μ or λ is considered the parameter of interest, while the other is regarded as a nuisance parameter. Suppose I_{11} and I_{22}, can be factored into functions of μ and λ. Say,*

$$I_{11} \propto h_{11}(\mu)h_{12}(\lambda),$$
$$I_{22} \propto h_{21}(\lambda)h_{22}(\mu),$$

where we assume the $h_{ij} > 0$, $i,j = \{1,2\}$. Then, Jeffreys prior can expressed as

$$\pi(\mu,\lambda) \propto h_{11}^{\frac{1}{2}}(\mu)h_{12}^{\frac{1}{2}}(\lambda)h_{21}^{\frac{1}{2}}(\lambda)h_{22}^{\frac{1}{2}}(\mu). \tag{4.5}$$

Proof. The proof of the theorem relies on the direct application of a result presented by [5], taking advantage of the block diagonal structure of the information matrix. □

Proposition 4.7. *The prior of the dispersion model with parameters μ and λ is given by*

$$\pi(\mu,\lambda) \propto \sqrt{\frac{\lambda}{V(\mu)}h_{21}(\lambda)h_{22}(\mu)}.$$

Proof. We can obtain the result directly by substituting the values of h_{11} and h_{12} into the Equation (4.5) from Proposition 4.6. □

Table 2 gives a review of the outcomes for some typical members of the dispersion family (Gaussian, Gamma and inverse Gaussian). The table incorporates also component densities for some members of the dispersion model family ($t(y,\mu)$) and the associated Jeffreys priors. For each such distribution, we investigate the propriety of Jeffreys priors (I_{11}, I_{22}, h_{11}, h_{12}, h_{21}, h_{22}, $\pi(\mu,\lambda)$) and posterior distributions ($f(\mu,\lambda|y)$, $f(\lambda|y)$).

Remark 4.8. [8] showed that the posterior is proper under π for appropriate values of α and β.

4.2.2. Main results. In what follows, we consider $y \sim f(.;\mu,\lambda)$.

- **Case of Gaussian distribution**

  Theorem 4.9. *The density $f(\lambda|y)$ clearly is a Gamma distribution with respect to y i.e., $\lambda \sim Ga\left(\frac{N}{2}, \frac{1}{2}\sum_{i=1}^{N}(y_i-\mu)^2\right)$. Therefore, the estimator $\widehat{\lambda}$ of λ can be obtained as*

$$\widehat{\lambda} = \frac{N}{\sum_{i=1}^{N}(y_i-\mu)^2}. \tag{4.6}$$

Table 2. Summary of component densities for some dispersion models.

Density	Gaussian	Gamma	Inverse Gaussian
$f(y\|\mu,\lambda)$	$\sqrt{\frac{\lambda}{2\pi}}e^{-\frac{\lambda}{2}(y-\mu)^2}$	$\frac{\lambda^\lambda y^{-1}}{\Gamma(\lambda)}e^{-\lambda\left(\frac{y}{\mu}-\log\frac{y}{\mu}\right)}$	$\sqrt{\frac{\lambda}{2\pi}}y^{-\frac{3}{2}}e^{-\lambda\frac{(y-\mu)^2}{2\mu^2 y}}$
$t(y,\mu)$	$-\frac{(y-\mu)^2}{2}$	$-\left(\frac{y}{\mu}-\log\frac{y}{\mu}\right)$	$-\frac{(y-\mu)^2}{2\mu^2 y}$
$I_{11}(\mu,\lambda)$	λ	$\lambda\mu^{-2}$	$\frac{\lambda}{\mu^3}$
$I_{22}(\mu,\lambda)$	$\frac{1}{2\lambda^2}$	$\frac{d^2}{d\lambda^2}\log\Gamma(\lambda)-\frac{1}{\lambda}$	$\frac{1}{2\lambda^2}$
$h_{11}(\mu)$	1	$\frac{1}{\mu^2}$	$\frac{1}{\mu^3}$
$h_{12}(\lambda)$	λ	λ	λ
$h_{21}(\lambda)$	$\frac{1}{\lambda^2}$	$\frac{d^2}{d\lambda^2}\log\Gamma(\lambda)-\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
$h_{22}(\mu)$	1	1	1
$\pi(\mu,\lambda)$	$\lambda^{-\frac{1}{2}}$	$\lambda^\alpha\mu^{-1}\left(\frac{d^2}{d\lambda^2}\log\Gamma(\lambda)-\frac{1}{\lambda}\right)^\beta$	$\lambda^{-\frac{1}{2}}\mu^{-\frac{3}{2}}$
$f(\mu,\lambda\|y)$	$\lambda^{\frac{n-1}{2}}e^{-\frac{\lambda}{2}\sum_{i=1}^n(y_i-\mu)^2}$	$\lambda^{n\lambda+\alpha}\left(\frac{d^2}{d\lambda^2}\log\Gamma(\lambda)-\frac{1}{\lambda}\right)^\beta\Gamma^{-n}(\lambda)q^\lambda\mu^{-(n\lambda+1)}e^{-\frac{\lambda t}{\mu}}$	$\mu^{-\frac{3}{2}}\left(\lambda^{\frac{n+1}{2}-1}e^{\frac{\lambda s}{2\mu^2}}\right)$
$f(\lambda\|y)$	$\lambda^{\frac{n}{2}-1}e^{-\frac{\lambda}{2}\sum_{i=1}^n(y_i-\mu)^2}$	$\lambda^\alpha\left(\frac{d^2}{d\lambda^2}\log\Gamma(\lambda)-\frac{1}{\lambda}\right)^\beta\Gamma^{-n}(\lambda)\Gamma(n\lambda)q^\lambda t^{-n\lambda}$	$\mu^{-\frac{3}{2}}\lambda^{\frac{n-1}{2}}\left(\prod_{i=1}^n y_i^{-\frac{3}{2}}\right)e^{-\frac{\lambda s}{2\mu^2}}$

*Note that q, t and s denote $\prod_{i=1}^N y_i$, $\sum_{i=1}^N y_i$, $\sum_{i=1}^n \frac{(y_i-\mu)^2}{y_i}$, respectively.

- **Case of Gamma distribution**

Theorem 4.10. *With the use of Stirling's formula, we investigated the behaviour of $f(\lambda|y)$ for a large λ, and the results revealed that*

$$f(\lambda|y)\propto\lambda^{\alpha-2\beta+\frac{N+1}{2}}\exp\left[-\lambda\left(N\log\left|\sum_{i=1}^N y_i\right|-\log\left|\prod_{i=1}^N y_i\right|-N\log N\right)\right],$$

which is proportional to a Gamma density and propriety is obtained when $\alpha-2\beta+\frac{N+3}{2}>0$. Hence, the estimator $\widehat{\lambda}$ of λ can be expressed as

$$\widehat{\lambda}=\frac{\alpha-2\beta+\frac{N+3}{2}}{N\log\left(\sum_{i=1}^N y_i\right)-\log\left(\prod_{i=1}^N y_i\right)-N\log N}. \tag{4.7}$$

- **Case of inverse Gaussian distribution**

Theorem 4.11. *The density $f(\lambda|y)$ is a Gamma distribution i.e., λ can be defined as $\lambda\sim Ga\left(\frac{N+1}{2},\sum_{i=1}^N\frac{(y_i-\mu)^2}{2\mu^2 y_i}\right)$. Then,*

$$\widehat{\lambda} = \frac{N+1}{\sum_{i=1}^{N} \frac{(y_i-\mu)^2}{\mu^2 y_i}}. \tag{4.8}$$

Proof. The proofs of Theorem 4.9, Theorem 4.10, and Theorem 4.11 exhibit similarities to the proof of Theorem 4.5. □

5. Numerical illustration

5.1. Simulation study

Within this section, a numerical investigation is carried out to facilitate a comparison of our suggested estimation approach for λ. The assessment of various estimators is accomplished through the evaluation of the values derived from the Squared Error Loss Function, Entropy Loss Function, and Precautionary Loss Function. The explicit definitions of these loss functions are as follows:

- Squared Error loss function (SELF): The SELF is a commonly employed metric defined as $l(\hat{\lambda}, \lambda) = c(\hat{\lambda} - \lambda)^2$. Its popularity stems from its computational simplicity, as it does not require extensive numerical calculations. It is important to note that the SELF is a symmetrical loss function, assigning equal penalties to both overestimation and underestimation.
- Entropy Loss Function (ELF): The ELF is a widely employed asymmetric metric, given by $l(\delta^p) \propto [\delta^p - p\log(\delta) - 1]$, where $\delta = \frac{\hat{\lambda}}{\lambda}$ and $p > 0$. It can be observed that the ELF corresponds to an entropy-based measure of distance between the distributions represented by λ and $\hat{\lambda}$. Moreover, this loss function achieves its minimum value when $\hat{\lambda} = \lambda$
- Precautionary Loss Function (PLF): The PLF offers a robust and straightforward alternative as an asymmetric precautionary loss function, belonging to a broad category of such functions. It can be expressed as $l(\hat{\lambda}, \lambda) = \frac{(\hat{\lambda}-\lambda)^2}{\hat{\lambda}}$. This loss function gradually approaches the origin, thereby mitigating the possibility of underestimation and yielding conservative estimators, particularly when evaluating small failure levels.

In our simulation study, we opted for sample sizes of $n = 50, 100$, and 1000 to encompass small, medium, and large datasets, respectively. We then proceeded to estimate the dispersion parameter for various distributions (specifically Gaussian, Gamma, and Inverse Gaussian) employing two distinct methodologies: the Pearson estimator and a non-informative Bayesian approach utilizing unbounded uniform and Jeffrey's priors.

In the case of the Gaussian distribution, we have set the mean and variance parameters to fixed values of $\mu = 0.5$ and $\sigma^2 = 2$, respectively. Similarly, for the Gamma distribution, the shape and rate parameters are fixed at $\alpha = 1$ and $\beta = 2$. As for the inverse Gaussian distribution, we have considered fixed values of $\mu = 1$ for the mean parameter and $\gamma = 0.2$ for the shape parameter.

In the simulation analysis performed using the Matlab software, we conducted 10000 iterations to measure the dispersion parameter for each distribution under different approaches. We assessed the efficiency of the estimates for various sample sizes and compared them across different values of the loss parameters: $c = 0.5, 1.0$, and 1.5, and $p = 0.5$ and 1.0. The results of the simulation analysis, specifically pertaining to the SELF, ELF, and PLF, are presented in the form of tables and curves, allowing for a comprehensive evaluation and comparison.

Table 3. Bayes Estimates of λ under Pearson estimation.

Size	Distribution	SELF			ELF		PLF
		c=0.5	c=1.0	c=1.5	p=0.5	p=1.0	
N=50	Gaussian	8.19×10^{-5}	3.55×10^{-4}	4.43×10^{-4}	2.15×10^{-4}	8.56×10^{-4}	5.44×10^{-4}
	Gamma	5.56×10^{-3}	8.06×10^{-2}	3.73×10^{-4}	5×10^{-4}	8.09×10^{-3}	6.92×10^{-3}
	inverse Gaussian	1.47×10^{-2}	8.64×10^{-2}	2.49×10^{-2}	3.20×10^{-2}	2.37×10^{-1}	6.59×10^{-2}
N=100	Gaussian	7.35×10^{-5}	8.96×10^{-5}	2.05×10^{-4}	5.18×10^{-5}	2.06×10^{-4}	2.50×10^{-4}
	Gamma	9.90×10^{-2}	5.60×10^{-6}	9.50×10^{-2}	4.81×10^{-3}	7.86×10^{-7}	9.82×10^{-2}
	inverse Gaussian	2.73×10^{-2}	1.24×10^{-2}	7.80×10^{-2}	5.40×10^{-2}	5.04×10^{-2}	2.35×10^{-1}
N=1000	Gaussian	5.04×10^{-7}	1.05×10^{-6}	1.48×10^{-6}	5.01×10^{-7}	2.00×10^{-6}	2.01×10^{-6}
	Gamma	1.67×10^{-3}	1×10^{-4}	1.25×10^{-3}	9.67×10^{-3}	1.17×10^{-3}	1.66×10^{-3}
	inverse Gaussian	1.59×10^{-3}	4.57×10^{-4}	2.13×10^{-3}	7.67×10^{-3}	4.92×10^{-3}	1.21×10^{-2}

Table 4. Bayes Estimates of λ under unbounded uniform prior.

Size	Distribution	SELF			ELF		PLF
		c=0.5	c=1.0	c=1.5	p=0.5	p=1.0	
N=50	Gaussian	5.70×10^{-4}	1.05×10^{-3}	9.13×10^{-4}	4.45×10^{-4}	1.8×10^{-3}	1.94×10^{-3}
	Gamma	2.45×10^{-4}	1.02×10^{-3}	2.63×10^{-3}	1.18×10^{-5}	2.57×10^{-4}	2.14×10^{-4}
	inverse Gaussian	3.36×10^{-3}	7.14×10^{-4}	2.75×10^{-4}	2.57×10^{-3}	2.72×10^{-3}	8.53×10^{-3}
N=100	Gaussian	1.74×10^{-4}	1.5×10^{-4}	2.65×10^{-4}	1.11×10^{-4}	4.5×10^{-4}	5.49×10^{-4}
	Gamma	1.31×10^{-4}	3.67×10^{-4}	1.24×10^{-4}	1.29×10^{-5}	8.60×10^{-5}	1.63×10^{-4}
	inverse Gaussian	1.99×10^{-5}	2.27×10^{-5}	1.04×10^{-3}	3.22×10^{-4}	1.39×10^{-4}	3.30×10^{-4}
N=1000	Gaussian	1.24×10^{-6}	2.63×10^{-6}	3.20×10^{-6}	1.12×10^{-6}	4.5×10^{-6}	4.71×10^{-6}
	Gamma	2.88×10^{-4}	2.66×10^{-5}	1.57×10^{-4}	1.61×10^{-5}	3.38×10^{-6}	2.74×10^{-4}
	inverse Gaussian	1.52×10^{-4}	6.73×10^{-5}	5.61×10^{-4}	1.33×10^{-3}	4.29×10^{-4}	1.69×10^{-3}

Table 5. Bayes Estimates of λ under Jeffreys prior.

Size	Distribution	SELF			ELF		PLF
		c=0.5	c=1.0	c=1.5	p=0.5	p=1.0	
N=50	Gaussian	4.38×10^{-5}	6.63×10^{-5}	2.24×10^{-4}	5.11×10^{-5}	2.05×10^{-4}	1.87×10^{-4}
	Gamma	4.10×10^{-4}	2.47×10^{-3}	1.97×10^{-4}	1.56×10^{-5}	1.70×10^{-6}	1.23×10^{-3}
	inverse Gaussian	6.67×10^{-5}	3.37×10^{-3}	3.28×10^{-4}	4.12×10^{-4}	9.77×10^{-4}	6.42×10^{-4}
N=100	Gaussian	9.62×10^{-6}	2.55×10^{-5}	6.56×10^{-5}	1.26×10^{-5}	5.06×10^{-5}	4.38×10^{-5}
	Gamma	5.03×10^{-4}	1.32×10^{-4}	2.27×10^{-4}	4.39×10^{-5}	1.44×10^{-5}	5.89×10^{-4}
	inverse Gaussian	3.62×10^{-7}	6.43×10^{-5}	2.81×10^{-4}	2.88×10^{-6}	5.92×10^{-5}	4.10×10^{-6}
N=1000	Gaussian	1.24×10^{-7}	2.04×10^{-7}	3.85×10^{-7}	1.25×10^{-7}	5×10^{-7}	4.98×10^{-7}
	Gamma	1.46×10^{-4}	5.33×10^{-4}	1.27×10^{-4}	1.17×10^{-5}	8.86×10^{-5}	1.66×10^{-4}
	inverse Gaussian	1.40×10^{-5}	8.56×10^{-5}	2.56×10^{-5}	7.73×10^{-5}	1.09×10^{-6}	1.33×10^{-4}

The results obtained from our analysis are presented in Tables 3–5, showcasing a range of parameter selections. Our findings consistently demonstrate that the SELF method tends to yield the smallest values across the majority of cases, particularly when the loss parameter $c = 0.5$. This conclusion is derived from a comprehensive comparison of the values obtained from the SELF method with those obtained from other methods, considering various conditions and parameter settings. Our analysis indicates that, overall, the SELF method consistently delivers the smallest values with a high degree of reliability, particularly when $c = 0.5$.

Within this simulation study, we performed a comparison of the posterior dispersion parameter λ under different loss functions using three distinct estimation methods: Pearson estimation, unbounded uniform prior, and Jeffrey's prior. Our analysis revealed that, within each loss function, Pearson estimation was the most suitable method for estimating the dispersion parameter of Gamma distributions. Conversely, for Gaussian distributions, the unbounded uniform prior proved to be the most appropriate estimation method. It is crucial to note that the choice of estimation method can significantly impact the accuracy of the results. Our conclusion regarding the suitability of Pearson estimation for Gamma distributions and the effectiveness of the unbounded uniform prior for Gaussian distributions is based on a comprehensive analysis of the data, considering the strengths and limitations of each method.

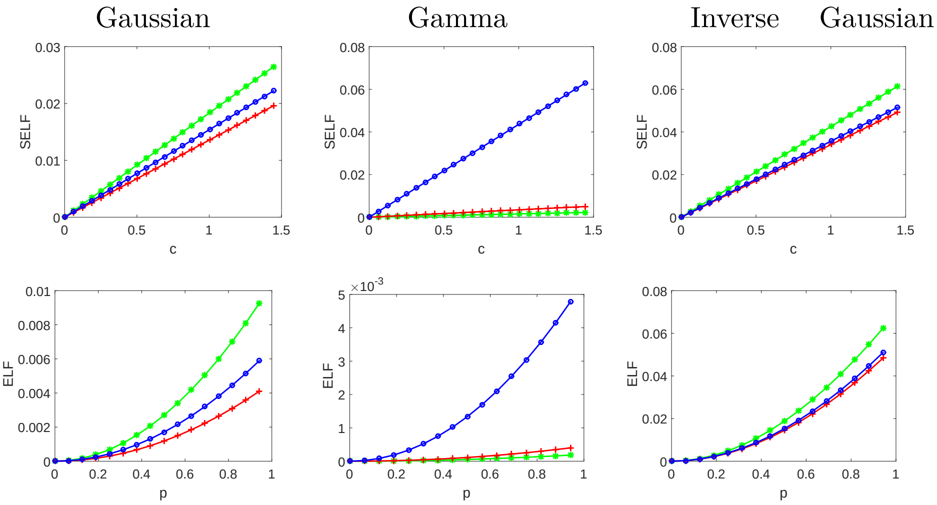

Figure 1. Graphical representation of the SELF and ELF values as a function of its parameters c and p, respectively.

Figure 1 displays the estimations of the dispersion parameter using the SELF and ELF loss functions, employing the Pearson method and non-informative priors (unbounded uniform prior and Jeffrey's prior). Our findings demonstrate that, for both the Gaussian distribution and the inverse Gaussian distribution, Bayes estimators with the unbounded uniform prior exhibit superior performance compared to those derived through the Pearson method and Jeffrey's prior, irrespective of the chosen loss functions. However, when considering the Gamma distribution, the Pearson estimator using the SELF and ELF loss functions outperforms the estimators based on the non-informative prior.

Furthermore, it is worth noting that when an unbounded uniform prior is employed, the resulting values tend to be the smallest. Additionally, we have observed that as the values of the parameters c and p for the SELF and ELF loss functions, respectively, increase, the values obtained from the unbounded uniform prior remain relatively small.

5.2. Data analysis

In this section, we examine the performance of the proposed estimator for the EDD by utilizing a real dataset and conducting an analysis using Matlab software. The dataset in question comprises 100 observations pertaining to the breaking stress of carbon fibers (measured in Gba) and is presented in Table 6. Previous studies conducted by [22] have examined this dataset. Furthermore, Fatima and Ahmad [7] have analyzed the same dataset using the transmuted exponentiated Pareto distribution and compared it with the transmuted Pareto, exponentiated Pareto, and Pareto distributions. In order to analyse this dataset, we assume an EDD with the density defined in Equation (2.1).

Table 6. Breaking stress of carbon fibers data.

3.70	2.74	2.73	2.50	3.60	3.11	3.27	2.87	1.47	3.11
4.42	2.41	3.19	3.22	1.69	3.28	3.09	1.87	3.15	4.90
3.75	2.43	2.95	2.97	3.39	2.96	2.53	2.67	2.93	3.22
3.39	2.81	4.20	3.33	2.55	3.31	3.31	2.85	2.56	3.56
3.15	2.35	2.55	2.59	2.38	2.81	2.77	2.17	2.83	1.92
1.41	3.68	2.97	1.36	0.98	2.76	4.91	3.68	1.84	1.59
3.19	1.57	0.81	5.56	1.73	1.59	2.00	1.22	1.12	1.71
2.17	1.17	5.08	2.48	1.18	3.51	2.17	1.69	1.25	4.38
1.84	0.39	3.68	2.48	0.85	1.61	2.79	4.70	2.03	1.80
1.57	1.08	2.03	1.61	2.12	1.89	2.88	2.82	2.05	3.65

Table 7 presents the estimated posterior values of the dispersion parameter λ assuming various prior distributions and utilizing the Pearson method. By incorporating different prior distributions, the table provides a comprehensive view of the variability of the estimated posterior values of λ based on the choice of prior distribution. The results in the table enable readers to make informed decisions regarding the selection of a suitable prior distribution and to understand the impact of prior assumptions about the estimated posterior values of λ.

Table 7. Parameter dispersion estimates of breaking stress data.

Estimate	**Pearson**			**Unbounded Uniform prior**			**Jeffreys prior**		
	Gaussian	Gamma	Inverse Gaussian	Gaussian	Gamma	Inverse Gaussian	Gaussian	Gamma	Inverse Gaussian
$\widehat{\lambda}$	0.9728	6.6848	17.5236	1.0023	6.8874	18.0546	0.9728	5.9647	2.2562

From Table 7, we have noted that when assuming different prior distributions for λ, the inference results are comparable only in the case of the Inverse Gaussian distribution. This suggests that the choice of prior has a significant impact on the inference results, but only when a different distribution other than the Inverse Gaussian is used. This highlights the importance of carefully considering the choice of prior in Bayesian analysis, especially when working with exponential family models.

In order to choose the most suitable prion distribution for analysing the breaking stress data performed by the EDD, it is important to assess several selection criteria. One approach is to consider criteria based on information, such as the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and the corrected AIC (AICc). These criteria are widely employed in the field of statistics and offer a measure of how well different models fit the data, considering both the accuracy of the fit and the number of parameters used in the model. The results of these criteria are typically presented in Table 8 for easy comparison, enabling the researcher to select the model that strikes the best balance between fit and parsimony.

Table 8. Goodness of fit for various models fitted for breaking stress data.

Model	**Best estimator**	**AIC**	**BIC**	**AICc**
Gaussian	Jeffreys prior	−283.5457	−280.9405	−283.5049
Gamma	Unbounded Uniform prior	−284.4673	−281.8621	−284.4264
Inverse Gaussian	Jeffreys prior	−299.4561	−296.8509	−299.4153

In Figure 2, a comprehensive analysis is presented, comparing the actual histogram and empirical cumulative distribution function (CDF) of the breaking stress data with their corresponding fitted probability density functions and cumulative distribution functions. The left side showcases the breaking stress data's histogram, accompanied by the fitted probability density functions that estimate the underlying data distribution. Meanwhile, the right side depicts the empirical CDF of the breaking stress data, overlaid with the fitted cumulative distribution functions that accurately represent the distribution's shape. This graphical depiction serves as a valuable tool for comprehending the breaking stress data's distribution and drawing meaningful conclusions about its characteristics.

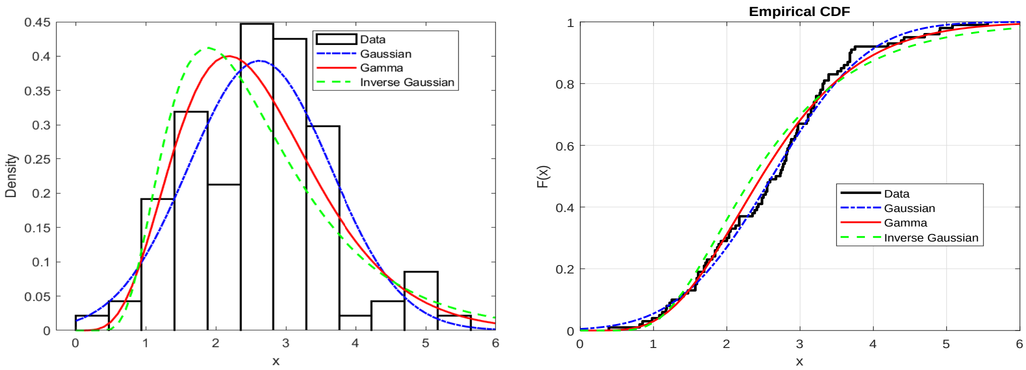

Figure 2. Bayes estimates of pdf and cdf for the breaking stress data.

6. Conclusion and discussion

Our research introduces a groundbreaking Bayesian estimation technique for determining the value of the unknown dispersion parameter, λ. In contrast to conventional informative methods, our approach employs a non-informative prior, enabling more robust estimations of this crucial parameter. By leveraging Bayesian inference and non-informative priors, our method offers a powerful tool for accurately estimating λ across diverse applications and environments.

Furthermore, we have provided different estimators for the dispersion parameter λ, by employing both the Pearson estimator (method of moments) and non-informative Bayesian estimation. We have demonstrated the existence of a non-informative Bayesian estimator for λ using an unbounded uniform and Jeffrey's priors. A comprehensive comparison of these estimators was conducted through a thorough simulation study.

Our analysis revealed that the Bayesian estimator, utilizing the unbounded uniform prior and Jeffrey's prior, outperforms the Pearson method in accurately estimating λ. The implementation of the non-informative Bayesian estimator takes into account both prior knowledge and data information, resulting in more informed predictions. In contrast, the Pearson approach relies solely on the data and disregards prior knowledge, leading to less accurate estimations of λ.

These findings underscore the significance of incorporating prior information in statistical modelling, as demonstrated by the superior prediction accuracy achieved by the Bayesian estimator.

Acknowledgements

We sincerely thank the Associate Editor and two anonymous referees for their valuable comments and constructive suggestions.

References

[1] J.M. Bernardo and A.F.M. Smith, *Bayesian Theory*, Wiley, 1994.

[2] J. Bertrand, P. Bertrand and J.P. Ovarlez, *The Transforms and Applications Handbook*, CRC & IEEE Presses, 2000.

[3] G.E. Box and G.C. Tiao, *Bayesian Inference in Statistical Analysis*, John Wiley & Sons, 2011.

[4] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.

[5] G.S. Datta and M. Ghosh, *Some remarks on noninformative priors*, J. Amer. Statist. Assoc. **90** (432), 1357-1363, 1995.

[6] D.M. Eaves, *On Bayesian nonlinear regression with an enzyme example*, Biometrika **70** (2), 373-379, 1983.

[7] K. Fatima and S.P. Ahmad, *Bayesian approach in estimation of shape parameter of the exponentiated moment exponential distribution*, J. Stat. Theory Appl. **17** (2), 359-374, 2018.

[8] C.W. Garvan and M. Ghosh, *On the property of posteriors for dispersion models*, J. Statist. Plann. Inference **78** (1-2), 229-241, 1999.

[9] A. Gelman and J. Hill, *Data Analysis using Regression and Multilevel/Hierarchical Models*, Cambridge University Press, 2007.

[10] C.B. Guure, N.A. Ibrahim and M.B. Adam, *Bayesian inference of the Weibull model based on interval-censored survival data*, Comput. Math. Methods Med., Doi: 10.1155/2013/849520, 2013.

[11] J.A. Hartigan, *Locally uniform prior distributions*, Ann. Statist. **24** (1), 160-173, 1996.

[12] N.T. Hobbs and M.B. Hooten, *Bayesian Models: A Statistical Primer for Ecologists*, Princeton University Press, 2015.

[13] J.G. Ibrahim and P.W. Laud, *On Bayesian analysis of generalized linear models using Jeffreys's prior*, J. Amer. Statist. Assoc. **86** (416), 981-986, 1991.

[14] H. Jeffreys, *Theory of Probability*, 3rd ed., Oxford University Press, 1961.

[15] B. Jorgensen, *Exponential dispersion models*, J. R. Stat. Soc. Ser. B. Stat. Methodol. **49** (2), 127-145, 1987.

[16] B. Jorgensen, *The Theory of Dispersion Models*, CRC Press, 1997.

[17] B. Jorgensen, J.R. Martinez and M. Tsao, *Asymptotic behaviour of the variance function*, Scand. J. Stat. **21** (3), 223-243, 1994.

[18] R. Kaas, M. Goovaerts, J. Dhaene and M. Denuit, *Modern Actuarial Risk Theory*, Springer, 2008.

[19] A.A. Khan, M. Aslam, Z. Hussain and M. Tahir, *Comparison of loss functions for estimating the scale parameter of log-Normal distribution using non-informative priors*, Hacet. J. Math. Stat. **45** (6), 1831-1845, 2015.

[20] N.P. Lemoine, *Moving beyond noninformative priors: Why and how to choose weakly informative priors in Bayesian analyses*, Oikos **128** (7), 912-928, 2019.

[21] P. McCullagh and J.A. Nelder, *Generalized Linear Models*, 2nd ed., Chapman & Hall, 1989.

[22] M.D. Nichols and W.J. Padgett, *A bootstrap control chart for Weibull percentiles*, Qual. Reliab. Eng. Int. **22** (2), 141-151, 2006.

[23] I. Sadok and A. Masmoudi, *New parametrization of stochastic volatility models*, Comm. Statist. Theory Methods **51** (7), 1936-1953, 2022.

[24] I. Sadok, A. Masmoudi and M. Zribi, *Integrating the EM algorithm with particle filter for image restoration with exponential dispersion noise*, Comm. Statist. Theory Methods **52** (2), 446-462, 2023.

[25] I. Sadok and M. Zribi, *Image restoration using Weibull particle filters*, 4th International Conference on Pattern Analysis and Intelligent Systems (PAIS), 12-13 October 2022, Algeria, 2022.
[26] S.K. Sinha, *Bayes estimation of the reliability function and hazard rate of a Weibull failure time distribution*, Trabajos de Estadistica **1** (2), 47-56, 1986.

Appendix

Lemma 6.1. *Let ν be a probability measure concentrated on $\mathbb{R}^+$, then*

(1) $\exists t_0 \in \Theta(\nu)$ *such that,* $(-\infty, t_0] \subset \Theta(\nu)$,
(2) $\lim_{t\to-\infty} K'_\nu(t) = 0$,
(3) $\lim_{t\to-\infty} \dfrac{K_\nu(t)}{t} = 0$.

Proof of Lemma 6.1

Proof. (1) Indeed, let $t_0 \in \Theta(\nu)$, if $t \leqslant t_0$, then $ty \leqslant t_0 y$ for all $y \geqslant 0$

$$L_\nu(t) = \int_0^{+\infty} e^{ty}\nu(dy) \leqslant \int_0^{+\infty} e^{t_0 y}\nu(dy) = L_\nu(t_0) < +\infty.$$

This implies that for all $t \in \Theta(\nu)$, $(-\infty, t_0] \subset \Theta(\nu)$.

(2) In fact,

$$K'_\nu(t) = \frac{L'_\nu(t)}{L_\nu(t)} = \frac{\int_0^{+\infty} xe^{tx}\nu(dx)}{\int_0^{+\infty} e^{tx}\nu(dx)},$$

upon setting $t = t_0 - s$, for all $s \geqslant 0$, then

$$\begin{aligned} K'_\nu(t) &= \frac{\int_0^{+\infty} xe^{(t_0-s)x}\nu(dx)}{\int_0^{+\infty} e^{(t_0-s)x}\nu(dx)} \\ &\leqslant \frac{\int_0^{\varepsilon} xe^{(t_0-s)x}\nu(dx) + \int_\varepsilon^{+\infty} xe^{(t_0-s)x}\nu(dx)}{\int_0^{\varepsilon} e^{(t_0-s)x}\nu(dx)} \\ &\leqslant \varepsilon + \frac{\int_\varepsilon^{+\infty} xe^{(t_0-s)x}\nu(dx)}{\int_0^{\varepsilon} e^{(t_0-s)x}\nu(dx)}. \end{aligned}$$

As $x \mapsto e^{(t_0-s)x}$ is a convex function, then

$$e^{(t_0-s)x} \geqslant (t_0-s)e^{(t_0-s)\varepsilon}(x-\varepsilon) + e^{(t_0-s)\varepsilon} \geqslant (t_0-s)e^{(t_0-s)x}(x-\varepsilon), \forall x \in [0,\varepsilon].$$

Implying that,

$$\begin{aligned} \int_0^{\varepsilon} e^{(t_0-s)x}\nu(dx) &\geqslant \int_0^{\varepsilon} (t_0-s)e^{(t_0-s)\varepsilon}(x-\varepsilon)\nu(dx) \\ &\geqslant (t_0-s)e^{-s\varepsilon}\int_0^{\varepsilon}(x-\varepsilon)e^{t_0 x}\nu(dx). \end{aligned}$$

Hence,

$$\begin{aligned} \frac{\int_\varepsilon^{+\infty} xe^{(t_0-s)x}\nu(dx)}{\int_0^{\varepsilon} e^{(t_0-s)x}\nu(dx)} &\leqslant \frac{\int_\varepsilon^{+\infty} xe^{(t_0-s)x}\nu(dx)}{(t_0-s)e^{-s\varepsilon}\int_0^{\varepsilon}(x-\varepsilon)e^{t_0x}\nu(dx)} \\ &\leqslant \frac{\int_\varepsilon^{+\infty} xe^{t_0x}\nu(dx)}{(t_0-s)\int_0^{\varepsilon}(x-\varepsilon)e^{t_0x}\nu(dx)} \underset{s\to+\infty}{\longrightarrow} 0. \end{aligned}$$

Therefore,

$$0 \leqslant \lim_{s\to+\infty} K'_\nu(t_0-s) \leqslant \varepsilon;\ \forall\varepsilon > 0.$$

Consequently,

$$\lim_{s\to+\infty} K'_\nu(t_0-s) = 0.$$

(3) Note that
$$\lim_{t\to-\infty}\frac{K_\nu(t)}{t}=\int_0^1\lim_{t\to-\infty}K'_\nu(tu)du=0.$$
□