Iran J Econ Stud, 2024, 13(1), 109-132

Iranian Journal of Economic Studies

Journal homepage: ijes.shirazu.ac.ir

Risk Estimation for the New Heavy Tail Distribution using Bayesian Approach

Hanieh Panahi[a*], Showkat Ahmad Lone[b], Hooman Aliakbarian[c]

a. Department of Mathematics and Statistics, La.C., Islamic Azad University, Lahijan, Iran.
b. Department of Basic Sciences, Saudi electronic university, Riyadh, Saudi Arabia.
c. Department of financial Management, La.C., Islamic Azad University, Lahijan, Iran.

Highlights

- **Extend the T-X family approach to develop a novel heavy-tailed distribution to estimate the VaR Measure.**
- **Obtain the Bayesian estimates using the MHA under appropriate prior and various novel loss functions.**
- **Demonstrate the applicability of the proposed method using insurance claims data.**

Article History

Received: 18 January 2025
Revised: 09 May 2025
Accepted: 11 May 2025
Published: 01 June 2025

JEL Classification
C11
C13
C32

Keyword
Heavy-tailed Weibull
Precautionary loss function
Metropolis-Hastings's method
Value-at-Risk measure

Abstract

This study addresses the evaluation of Value-at-Risk (*VaR*) using a Bayesian approach, specifically employing the heavy-tailed Weibull (*HTW*) distribution. The *VaR* is a crucial financial metric for business and investment decision-making. While various methods exist for estimating *VaR*, this research focuses on statistical techniques utilizing heavy-tail distributions. The paper extends the heavy-tailed Weibull model, which is particularly relevant for financial applications and provides reliable predictions for heavy-tailed data. The statistical properties of the *HTW* distribution are developed and Bayesian estimates under multiple symmetric and asymmetric loss functions are obtained .The Bayes estimate is evaluated from the posterior distribution that minimizes the corresponding posterior risk. Due to the complexity of the posterior distribution, the Metropolis-Hastings algorithm (MHA) is implemented to draw posterior samples. The Markov Chain Monte Carlo sample convergence is evaluated through diagnostic plots. The insurance loss data is used to display the application of the presented methodology in a real-world situation. The outcomes showed that Bayesian estimates can be used to evaluate the Value-at-Risk measure well. Financial institutions and risk managers can consider implementing Bayesian methods with heavy-tailed distributions, particularly the heavy-tailed Weibull model, for more accurate *VaR* estimation. This approach is especially valuable for portfolios with extreme events or fat-tailed return distributions.

1. Introduction

Heavy-tailed distributions gained wider application in the statistical literature since researchers began to utilize various variants of the Extreme Value Theorem (EVT), namely, the Block Maxima method. The EVT, a specialized

* hpanahi@iau.ac.ir & h.panahi@yahoo.com.
DOI: 10.22099/ijes.2025.52208.1995

domain within statistical theory, provides a rigorous framework for characterizing and forecasting extreme events, defined as observations exceeding previously recorded magnitudes. In the context of risk theory, EVT is of paramount importance due to its capacity to facilitate the analysis of risk measures associated with observations that deviate substantially from the central tendency, particularly those situated within the distributional tail.

A critical risk measure in this regard is Value-at-Risk (VaR). Formally, given a prespecified confidence level $T \in (0,1)$, VaR represents the T -quantile of the aggregate profit and loss distribution. This means that with probability T, the losses will not exceed the VaR value. To put it another way, there's only a (1-T) probability that losses will be greater than the VaR amount. To express this mathematically, if we denote our random loss variable as X, then, $P(X \leq VaR(T)) = T$, or equivalently, $VaR(T) = F^{-1}(T)$, where F^{-1} is the inverse CDF (quantile function) of the loss distribution. The value at risk has become a widely adopted risk measure in financial institutions and regulatory frameworks, primarily due to its ability to summarize market risk exposure into a single, easily interpretable number. VaR estimates the potential loss in value of an asset or portfolio over a defined time horizon for a given confidence level. The first study to estimate potential loss and calculate risk was done in 1888 by Francis Edgeworth. He improved statistical theory and estimated future probabilities using past empirical events. Further, Bali (2007) studied the generalized extreme value approach to financial risk measurement. Trzpiot & Majewska (2010) estimated the VaR using extreme value and robust approaches. Socgnia & Wilcox (2014) performed an empirical investigation of various subclasses of the generalized hyperbolic distribution, namely the hyperbolic, variance gamma, normal inverse Gaussian, and skewed-t distributions, utilizing daily log-returns from the Johannesburg Stock Exchange as their data source. Chinhamu et al. (2015) studied VaR criterion in evaluating the gold market. Peng et al. (2020) improved VaR prediction using model uncertainty. Müller & Brutti Righi (2024) compared different methods in VaR estimation. Chronopoulos (2024) predicted VaR using deep neural network quantile regression. Additionally, the application of heavy-tailed distributions in risk management for estimation tasks is well-established. Martin et al. (2022) considered peaks-over-threshold method for generalized Pareto distribution in risk evaluation. Luger (2012) derived the finite sample bootstrap in GARCH models with heavy innovations. Spierdijk (2016) considered confidence intervals for VaR by using heavy tails and skewness properties. Panahi (2019) evaluated the VaR estimation under the kappa distribution. The Weibull distribution holds a significant position as a parametric probability distribution in reliability and financial analysis, with applications including reliability engineering and decision-making. Research has progressively extended the Weibull distribution's capabilities through modifications and generalizations. For instance, Lai et al. (2003) introduced a three-parameter modified version, while Bebbington et al. (2007) and Sarhan &

Apaloo (2013) further expanded the distribution with two- and four-parameter extensions, respectively, to accommodate various hazard rate shapes, including bathtub and monotone forms. Famoye et al (2005) discussed the beta Weibull distribution with several hazard rate functions. Moradi et al. (2022) estimated the exponentiated Weibull parameters under censored data. Benkhelifa (2021a, 2021b) proposed the Weibull Birnbaum-Saunders model and beta reduced modified Weibull distribution respectively. Ghazal & Radwan (2022) proposed the reduced modified Weibull distribution with its application to medical and engineering data .

Although, several distributions have been introduced for the modeling of applied data, however, there are situations where these distributions are not suitable for modeling different data. For this, we propose to use the heavy-tailed distribution for evaluating the Value-at-Risk (VaR) criterion. The heavy-tailed distributions are essential in econometric modeling and forecasting of financial time series. By incorporating heavy-tailed distributions into portfolio optimization models, investors can construct portfolios that are more robust to market shocks and better aligned with their risk preferences. Alzaatreh et al. (2013) proposed the T-X family approach to obtain heavy-tailed distribution using the following cumulative distribution function (CDF):

$$F^{*}(x) = \int_{0}^{H(F(x;\theta))} u(t)dt \quad (1)$$

Here, $H(F(x;\theta))$ meets certain conditions; see Alzaatreh et al. (2013). So, the probability density function of $F^{*}(x)$ can be written as:

$$f^{*}(x) = \left(\frac{\partial}{\partial x} H(F(x;\theta))\right) u(H(F(x;\theta))); \;\; x \in R. \quad (2)$$

Recently, Zhao et al. (2020) proposed the Type I heavy-tailed Weibull distribution. They obtained the maximum likelihood estimates of parameters and analyzed three real data sets to show the fit of the model to the reliability engineering, bio-medical and financial data. Based on the T-X family pioneered by Alzaatreh et al. (2013), we propose the HTW distribution and provide some mathematical properties. Also, we use this distribution because financial data often doesn't follow normal patterns, especially during extreme events that standard distributions underestimate. Its additional parameter specifically models tail thickness, providing a more realistic representation of financial risks while maintaining mathematical tractability. This makes HTW particularly valuable for VaR estimation, helping financial institutions avoid underestimating potential losses during extreme events. From a statistical perspective, the HTW distribution offers several advantages. It provides a good balance between mathematical tractability and realistic modeling of extreme events. The Bayesian approach further enhances this by incorporating prior knowledge and providing full posterior distributions rather than just point estimates, resulting in more robust risk management strategies that better align with regulatory requirements. Also, extending the application of EVT to a Bayesian framework offers several benefits, mirroring the advantages highlighted earlier with the HTW distribution.

Incorporating prior beliefs about the tail behavior can improve the accuracy and stability of VaR estimates, especially when dealing with limited data. Furthermore, the Bayesian approach provides a natural mechanism for quantifying the uncertainty associated with the parameters and, consequently, with the VaR estimates. This uncertainty quantification is crucial for risk managers, as it allows them to assess the potential range of losses and make more informed decisions about capital allocation and risk mitigation strategies.

The motivations of this study include presented method can be applied across various fields such as engineering, management, insurance and more. The simplicity of the proposed method combined with its remarkable flexibility in modeling real-world data, makes it an attractive alternative to the VaR evaluation.

To the best of our knowledge, there is no study reported in literature that proposed estimation of the HTW parameters and evaluation of the VaR measure. So, we focus on proposing a novel version of Weibull model for gaining the VaR estimation. We illustrate the Bayes methodology with an application to financial data. Bayes estimation under several symmetric and asymmetric loss functions is presented. Since the posterior distribution becomes complex, we propose the use of Metropolis-Hastings (MH) method to draw the posterior samples and summarize the characteristics of the posterior samples. Posterior predictive density is derived for future observations.

In Section 2, we propose the HTW model and gain its mathematical properties. Sections 3 and 4 deal with the parameter estimates using the MLE and Bayesian approaches. Bayesian Section presents the formulation of the Bayesian model for the given problem under various novel loss functions. The insurance loss data is analyzed in Section 5. Finally, section 6 deals with the conclusion of the paper.

2. The HTW Distribution

The field of statistics offers a vast array of continuous distributions. In traditional financial modeling, the assumption of normally distributed returns has been a cornerstone for decades. This assumption simplifies risk assessment and portfolio optimization. However, empirical evidence consistently reveals that financial asset returns often exhibit characteristics that deviate significantly from normality. Specifically, these returns tend to have fatter tails, indicating a higher probability of extreme events than predicted by the normal distribution. This phenomenon necessitates the adoption of heavy-tailed distributions, such as the Student’s t-distribution, generalized hyperbolic distribution, and stable distributions, to more accurately capture the true risk profile of financial assets. These distributions are characterized by their ability to model the increased likelihood of large, infrequent fluctuations, which are crucial for effective risk management and investment strategies. By incorporating these distributions, financial models can better reflect the realities of market dynamics, leading to more robust and reliable predictions. By employing heavy-tailed distributions, Value at risk measure can provide a more conservative and realistic assessment of risk. This is especially critical for institutions that must comply with regulatory capital requirements, as underestimating risk can lead to inadequate capital reserves and potential financial instability. Furthermore, heavy-tailed distributions allow for the incorporation of tail dependencies, which capture the tendency of extreme events to occur simultaneously across different assets, further enhancing the accuracy of risk assessments.

Moreover, the introduction of new heavy-tailed distributions can have a significant impact on better modeling of financial data. One of these distributions is the heavy-tailed Weibull (HTW) distribution, which is used in this article to model financial data. The HTW distribution with parameters α and β, is given by the following probability density function (PDF) and cumulative distribution function (CDF):

$$f(x,\alpha,\beta)=\frac{\alpha\beta^2x^{\alpha-1}e^{-\beta x^{\alpha}}}{\left(\beta+(1-\beta)e^{-x^{\alpha}}\right)^{\beta+1}};\quad x\geq 0,\alpha,\beta\geq 0, \tag{3}$$

and

$$F(x,\alpha,\beta)=1-\left(\frac{e^{-x^{\alpha}}}{\beta+(1-\beta)e^{-x^{\alpha}}}\right)^{\beta};\quad x\geq 0,\alpha,\beta\geq 0. \tag{4}$$

Moreover, the hazard function and the survival function of *HTW* are:

$$H(x,\alpha,\beta)=\frac{\alpha\beta^2x^{\alpha-1}}{\beta+(1-\beta)e^{-x^{\alpha}}}, \tag{5}$$

and

$$S(x,\alpha,\beta)=\left(\frac{e^{-x^{\alpha}}}{\beta+(1-\beta)e^{-x^{\alpha}}}\right)^{\beta}. \tag{6}$$

To gain the heavy-tailed property of the HTW model, we consider the following Theorem.

Theorem: A distribution $F(x;\alpha,\beta)$ for a random variable *X* is considerate to be heavy tail if and only if

$$\lim_{x\to\infty}\frac{1-F(x;\alpha,\beta)}{S(x;\upsilon)}=\infty;\ S(x;\upsilon)\ \textit{is the survival of exponential distribution.}$$
(7)

Proof: Based on Equation (2), we have

$$\lim_{x\to\infty}\frac{e^{-\beta x^{\alpha}}}{e^{-\upsilon x}(\beta+(1-\beta)e^{-x^{\alpha}})^{\beta}}=\frac{e^{-\beta x^{\alpha}}e^{\upsilon x}}{(\beta+(1-\beta)e^{-x^{\alpha}})^{\beta}}. \tag{8}$$

Thus, for $\alpha\in(0,1)$ and any value of β and υ, we can write,

$$\lim_{x\to\infty}e^{-\beta x^{\alpha}+\upsilon x}=\infty \text{ and } \lim_{x\to\infty}(\beta+(1-\beta)e^{-x^{\alpha}})^{\beta}=\beta^{\beta}.$$
(9)

So, for $\alpha\in(0,1)$, we have,

$$\lim_{x\to\infty}\frac{e^{-\beta x^{\alpha}}e^{\upsilon x}}{(\beta+(1-\beta)e^{-x^{\alpha}})^{\beta}}=\infty. \tag{10}$$

Thus, we can conclude that HTW distribution has heavy-tailed property.

3. Properties of HTW Model

In this section, we considered some mathematical properties of *HTW* distribution such as:

3.1. Quantile Function (QuF)

The QuF of the random variabla X which has HTW model is given as,

$$QF(p) = \left[-\ln\left(\frac{\beta}{1-\beta+(1-p)^{-1/\beta}}\right)\right]^{1/\alpha};\ \ p \in (0,1) \quad (11)$$

Proof: For $0 < p < 1$, the *QuF* is the inverse of cumulative distribution of *X*, that is:

$$1-\left(\frac{e^{-x^{\alpha}}}{\beta+(1-\beta^{\beta})e^{-x^{2}}}\right)^{\beta} = p, \quad (12)$$

Therefore,

$$(1-p)^{1/\beta} = \frac{1}{\beta e^{\alpha^{\alpha}}+(1-\beta)}, \quad (13)$$

$$(1-p)^{-1/\beta} = \beta e^{\alpha} + (1-\beta),$$
(14)

$$\frac{(1-p)^{-1/\beta}-(1-\beta)}{\beta} = e^{x^{\alpha}}, \quad (15)$$

Hence,

$$\left[\ln\left(\frac{(1-p)^{-1/\beta}-(1-\beta)}{\beta}\right)\right]^{1/\alpha} = x. \quad (16)$$

which completes the proof.

3.2. Measures for HTW model

Suppose the continuous random variable X has HTW model $X : HTW(\alpha,\beta)$. The rth-moment of HTW model is defined by:

$$u_r^{'} = \sum_{i,j,k=0}^{\infty} Z(i,j,k)N(\alpha,\beta,r);\quad r = 1,2,\ldots,$$
(17)

where, $Z(i,j,k) = \binom{i}{j}\binom{i}{k}\binom{\alpha+1}{i}(-1)^{i+j+k}$ and $N(\alpha,\beta,r) = \Gamma((r/\alpha)+1)\beta^{j+2}(1+\beta)^{-((r/\alpha)+1)}$.

The mean and variance for the HTW model are given by:

$$u_1^{'} = \sum_{i,j,k=0}^{\infty}\binom{i}{j}\binom{i}{k}\binom{\alpha+1}{i}(-1)^{i+j+k}\Gamma(1/\alpha+1)\beta^{j+2}(1+\beta)^{-(1/\alpha+1)},$$
(18)

$$u_2^{'} = \sum_{i,j,k=0}^{\infty}\binom{i}{j}\binom{i}{k}\binom{\alpha+1}{i}(-1)^{i+j+k}\Gamma(2/\alpha+1)\beta^{j+2}(1+\beta)^{-(2/\alpha+1)},$$
(19)

$$Var(X) = u_2^{'} - (u_1^{'})^2. \quad (20)$$

The index of dispersion for X is:

$$CV = \frac{Var(X)}{u_1^{'}}. \quad (21)$$

3.3. Moment generating function (MoGF)

The MoGF of the HTW model can be given as:

$$M_X(t) = \int_0^{\infty} e^{tx} f(x,\alpha,\beta)dx = \int_0^{\infty} e^{tx} \frac{\alpha\beta^2 x^{\alpha-1} e^{-\beta\alpha^{\alpha}}}{(\beta+(1-\beta)e^{-x^{\alpha}})^{\beta+1}} dx$$

$$= \sum_{i,j,k,h=0}^{\infty} \frac{t^h}{h!} Z(i,j,k)\Gamma((h/\alpha)+1)(k+\beta)^{-((1/\alpha)+1)}. \quad (22)$$

3.4. Stress- Strength Reliability

Studies have indicated that certain parts or equipment endure because of their durability. These gadgets can withstand a given amount of stress, but when more force is placed on them, they malfunction because they are unable to handle it. The stress strength model defines a component's life as follows: if stress surpasses strength, the component will fail. Let the model consists of strength variable such that $X : HTW(\alpha,\beta)$ and independent stress variable $Y : HTW(\lambda,\beta)$. Then, the stress-strength reliability model can be written as:

$$Reliability = P(X>Y) = \int_0^{\infty} P(X>Y|Y=y)\, f_Y(y)dy$$

$$= \int_0^{\infty} f_Y(y)F_Y(x)\, dy = \frac{\lambda}{\alpha+\lambda}, \quad (23)$$

which is a continuous function (Figure 1).

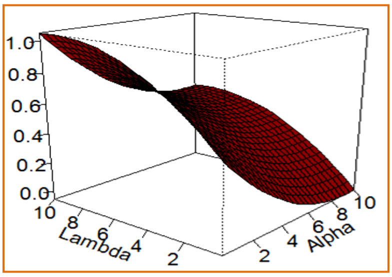


Figure 1: The 3D plot of reliability model.
Source: *Research Findings*

Figure 1 demonstrates how reliability decreases as these parameters change. The highest reliability (values close to 1.0) occurs when λ is high and α is low (back corner of the plot). The reliability then decreases following a non-linear pattern as these parameters shift. This visualization likely represents a reliability model where:

- Higher values on the z-axis (closer to 1.0) indicate better reliability.
- The saddle-like shape shows there are optimal combinations of λ and α . The key insight is to understand which combinations of λ and α

provide acceptable reliability thresholds for whatever system or component is being modeled.

3.5. Value-at-Risk Evaluation

Value at Risk (VaR) has emerged as a prevalent risk metric within financial institutions and regulatory structures, primarily attributable to its capacity to condense market risk exposure into a singular, easily understandable figure. The calculation of VaR typically involves employing statistical techniques, such as historical simulation, Monte Carlo simulation, or parametric methods. Each of these approaches possesses its own strengths and weaknesses. Historical simulation utilizes past market data to forecast potential future losses, assuming that historical patterns will persist. Monte Carlo simulation, on the other hand, involves generating numerous random scenarios to simulate potential market movements. Parametric methods rely on distributional assumptions about the underlying risk factors, such as normality, which may not always hold true in practice, particularly during periods of market stress. Therefore, choosing the appropriate method for estimating value at risk is important.

Let X represents random variable that were detected from $HTW(\alpha, \beta)$, the Value-at-Risk (VaR) criterion can be evaluated by finding the inverse of CDF for our proposed HTW model. Thus, the VaR can be written as:

$$VaR_T(X) = F^{-1}(T, \alpha, \beta) = \left[- ln\left(\frac{\hat{\beta}}{1-\hat{\beta}+(1-T)^{-1/\hat{\beta}}}\right)\right]^{1/\hat{\alpha}}. \tag{24}$$

4. Likelihood Function

A well-liked technique for parameter estimation in statistical models is called maximum likelihood estimation (MLE). It is a method for finding a statistical model's unknown parameters using sample data. The MLE stands as a cornerstone of statistical inference, offering a powerful and versatile framework for estimating parameters of probability distributions based on observed data. Unlike simpler methods like the method of moments, MLE leverages the full distributional assumptions to derive estimators with desirable properties, making it a preferred choice in many statistical applications. In this section, we discuss estimation of parameters of the HTW parameters using the likelihood approach. The likelihood function (LikF) for the set of parameter $\theta = (\alpha, \beta)$ is given by:

$$L\left(\alpha,\beta\middle|data\right)=\prod_{i=1}^{n}\frac{\alpha\beta^{2}x_i^{\alpha-1}e^{-\beta x_i^{\alpha}}}{\left(\beta+\left(1-\beta\right)e^{-x_i^{\alpha}}\right)^{\beta+1}}$$

$$=\alpha^{n}\beta^{2n}\prod_{i=1}^{n}x_i^{\alpha-1}e^{-\beta\sum_{i=1}^{n}x_i^{\alpha}}\times\prod_{i=1}^{n}\left(\left(1-\beta\right)e^{-x_i^{\alpha}}+\beta\right)^{-\beta-1}, \quad (25)$$

In practice, it is often more convenient to work with the log-likelihood function, which is the natural logarithm of the likelihood function. The logarithm transformation preserves the location of the maximum but converts products into sums, simplifying calculations. Moreover, the log-likelihood often exhibits more desirable properties for optimization, such as convexity, which aids in finding the global maximum. So, the log-*LikF* is:

$$ln\,L\,(\alpha,\beta|data)=n\,ln\alpha+2nln\beta+(\alpha-1)\sum_{i=1}^{n}lnx_i-\beta\sum_{i=1}^{n}x_i^{\alpha}-(1-\beta)\sum_{i=1}^{n}\ln((1-\beta)e^{-x_i^{\alpha}}+\beta).$$

(26)

In order to obtain the MLEs of the parameters, we maximize the log-LikF (6) with respect to the proposed parameters. MLEs are the solution of the nonlinear equation system:

$$\frac{\partial\ln L\left(\alpha,\beta\middle|data\right)}{\partial\alpha}=\frac{n}{\alpha}+\sum_{i=1}^{n}lnx_i-\beta\sum_{i=1}^{n}x_i^{\alpha}\ln x_i-\left(\beta+1\right)\sum_{i=1}^{n}\frac{\left(1+\beta\right)e^{-x_i^{\alpha}}x_i^{\alpha}\ln x_i}{\beta+\left(1+\beta\right)e^{-x_i^{\alpha}}}=0, \quad (27)$$

and

$$\frac{\partial\ln L\left(\alpha,\beta\middle|data\right)}{\partial\beta}=\frac{n}{\beta}-\sum_{i=1}^{n}x_i^{\alpha}-\sum_{i=1}^{n}\ln\left(\left(1-\beta\right)e^{-x_i^{\alpha}}+\beta\right)-(\beta+1)\sum_{i=1}^{n}\frac{1-e^{-x_i^{\alpha}}}{\beta+\left(1-\beta\right)e^{-x_i^{\alpha}}}=0 \quad (28)$$

These equations do not provide analytical expressions of the MLEs. Therefore, we maximize the log-likelihood function numerically using a suitable iterative method (e.g. Newton-Raphson method). We suggest the readers to use optim() function of R-software to compute MLEs numerically. This function offers various optimization methods, including Newton-Raphson and Nelder-Mead methods. The Nelder-Mead method is a popular optimization algorithm introduced by Nelder (1965), which is a direct search method and does not require to derive the objective function. The R function is as follows.

- *optim*(set of initial values, $-\,ln\,L\,(\alpha,\beta|data)$, data)

There is no definitive rule for selecting initial values in numerical optimization, as noted in the literature. But the right choice must be made.

Because choosing poor initial values can lead to convergence or local optimality problems.

5. Bayesian Estimation

In contrast to classical theory, the parameter in the Bayesian approach is regarded as random variable whose distribution is known to the investigator. This is a reasonable assumption because no population's parameters remain constant over the course of a study (Asadi et al. (2018), Mahdavi & Ehsani (2022), Vardani et al. (2024) and Panahi (2025)). Bayesian methods, a cornerstone of modern statistical inference, offer a fundamentally different approach compared to classical, frequentist statistics. Instead of treating parameters as fixed but unknown quantities, Bayesian methods treat them as random variables, possessing a probability distribution that reflects our uncertainty or prior beliefs about their values. This probabilistic perspective allows for a more nuanced and flexible framework for data analysis, inference, and decision-making. At the heart of Bayesian methods lies Bayes' theorem, a mathematical equation that updates our prior beliefs in light of new evidence. One of the key advantages of Bayesian methods is their ability to incorporate prior knowledge into the analysis. This is particularly valuable when dealing with limited data or when there is substantial prior information about the parameters of interest. The prior distribution can reflect expert opinions, previous studies, or even vague notions about the plausible range of parameter values. By combining prior knowledge with the data, Bayesian methods can provide more robust and reliable inferences, especially in situations where the data alone is insufficient.

One of the main challenges in Bayesian methods is the specification of the prior distribution. The choice of prior can have a significant impact on the posterior distribution, especially when the data is limited. It is important to carefully consider the prior distribution and to conduct sensitivity analysis to assess how the results change under different prior assumptions. Non-informative priors, which aim to minimize the influence of the prior, can be used, but they may not always be appropriate or well-defined. It is important to mention here that no conjugate prior is related to the random variables. So, we use the gamma prior distribution as an informative prior leveraging its flexibility. On the basis of joint prior density ($\upsilon(\alpha,\beta) \propto \alpha^{\varsigma_1-1}\beta^{\varsigma_2-1}e^{-\alpha\tau_1-\beta\tau_2}$) along with the likelihood function (6), the joint posterior function under can be written as:

$$\upsilon(\alpha,\beta|data) \propto L\left(\alpha,\beta|data\right)\upsilon(\alpha,\beta) = \alpha^{n+\varsigma_1-1}\beta^{2n+\varsigma_2-1}\prod_{i=1}^{n}x_i^{\alpha-1}e^{-\alpha\tau_1-\beta(\sum_{i=1}^{n}x_i^{\alpha}+\tau_2)} \times\prod_{i=1}^{n}\left(\left(1-\beta\right)e^{-x_i^{\alpha}}+\beta\right)^{-\beta-1} \tag{29}$$

Hence, the conditional posterior (ConP) densities of α and β can be obtained, up to proportionality, as:

$$\upsilon_1(\alpha|\beta, data) = \alpha^{n+\varsigma_1-1} e^{-\alpha\tau_1-\beta\sum_{i=1}^{n} x_i^{\alpha}} \prod_{i=1}^{n} x_i^{\alpha-1}\left(\left(1-\beta\right)e^{-x_i^{\alpha}}+\beta\right)^{-\beta-1} \quad (30)$$

and

$$\upsilon_2(\beta|\alpha, data) = \beta^{2n+\varsigma_2-1} e^{-\beta(\sum_{i=1}^{n} x_i^{\alpha}+\tau_2)} \times \prod_{i=1}^{n}\left(\left(1-\beta\right)e^{-x_i^{\alpha}}+\beta\right)^{-\beta-1} \quad (31)$$

We considered different asymmetric and symmetric loss functions. These loss functions and the associated Bayesian estimators of them are presented in Table 1. Loss function measures the difference between the actual value and the predicted value in statistics. These functions can be symmetric or asymmetric. A loss function is symmetric if the penalty for overestimating and underestimating the true value is the same. In other words, whether the predicted value is higher or lower than the actual value, the loss remains equal, for example, squared error loss. This function squares the difference, meaning larger errors are penalized more heavily. In some cases, the cost of overestimation and underestimation is not equal. Asymmetric loss functions are used when one type of error is costlier than the other. Other loss functions are in the case of asymmetric loss functions. In financial models, like stock market predictions, overestimating losses or gains may have different financial consequences. Choosing the right loss function depends on the problem and the relative importance of different types of errors.

Table 1. loss functions and the associated Bayesian estimators for $\theta = (\alpha, \beta)$.

Loss Functions	Bayesian estimators
Squared error loss(SELo): $L_1(\theta) = (\theta - \hat{\theta})^2$	$E(\theta \| data)$
Weighted squared error:(*WSEr*) $L_2(\theta) = \frac{(\theta - \hat{\theta})^2}{\theta}$	$\left(E(\theta^{-1} \| data)\right)^{-1}$
K-loss function (*KLFu*) $L_3(\theta) = \left(\sqrt{\frac{\theta}{\hat{\theta}}} - \sqrt{\frac{\hat{\theta}}{\theta}}\right)^2$	$\sqrt{\frac{E(\theta \| data)}{E(\theta^{-1} \| data)}}$
Precautionary loss function(PLFu): $L_4(\theta) = \left(1 - \frac{\theta}{\hat{\theta}}\right)^2$	$\frac{E(\theta^{-1} \| data)}{E(\theta^{-2} \| data)}$
Linex loss function (LLFu): $L_5(\theta) = e^{\varphi(\Delta)} - \varphi(\Delta) - 1; \Delta = (\theta - \hat{\theta})$ *See, Figure 2.*	$-\frac{1}{\varphi}\log\left[E(e^{-\varphi\theta} \| data)\right]$

Source: *Research Findings*

Since the Bayesian estimators do not possess analytical solutions, a Markov Chain Monte Carlo (MCMC) method, namely the Metropolis-Hastings (MH) algorithm, is implemented to derive an approximate solution.

Figure 2 shows plots of the Linex (Linear-Exponential) loss function for different values of the shape parameter φ. This loss function is asymmetric, which is its primary feature. Unlike symmetric loss functions (like squared error), Linex penalizes errors in one direction more heavily than errors in the opposite direction. For negative φ values, the function penalizes negative errors ($\Delta < 0$) more heavily than positive errors. For positive φ values, the function penalizes positive errors ($\Delta > 0$) more heavily than negative errors. As |φ| increases (comparing the black, orange, and green lines), the asymmetry becomes more pronounced, showing greater penalization in the respective directions.

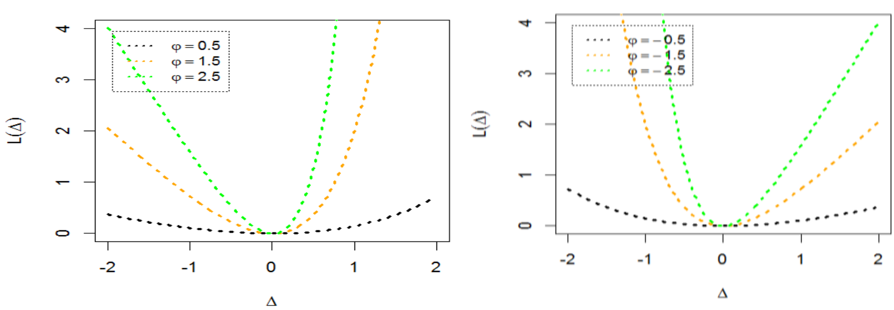


Figure 2. Plot for Linex loss function for: positive values (left) and negative values (right) of φ .

Source: *Research Findings*

5.1. The MH Algorithm

In the area of statistical computing, MCMC is a vital approach. A number of parameters can be estimated using this effective technique, which can be used to sample from a particular probability distribution. One well-known MCMC methodology is MH technique. The following is a description of the MH technique.

Step 1: Set the starting value to the MLEs, represent by $\hat{\alpha}$ and $\hat{\beta}$.

Step 2: Set $k = 1$.

Step 3: At the k^{th} iteration, generate a proposal point $\hat{\alpha}^*$ from $N(\alpha^{i-1}, Variance(\alpha))$.

Step 4: Obtain the acceptance probability as:

$$Acceptance\ Probability(\hat{\alpha}^{k-1}, \hat{\alpha}^{k}) = \min\left(1, \upsilon_1(\hat{\alpha}^* | data) / \upsilon(\hat{\alpha}^{k-1} | data)\right).$$

Step 5: Generate a random variable U_1 from standard uniform distribution.

$Step\,6:$ If $U_1 \le Acceptance\ Probability(\hat{\alpha}^{k\text{-}1}, \hat{\alpha}^{k})$, accept the proposal, else set $\hat{\alpha}^{k\text{-}1} = \hat{\alpha}^{k}$.

$Step\,7:$ Generate $\hat{\beta}^{*}$ from the Normal distribution $N(\beta^{i-1}, Variance(\beta))$.

$Step\,8:$ Evaluate the acceptance probability as:

$$Acceptance\ Probability(\hat{\beta}^{k\text{-}1}, \hat{\beta}^{k}) = \min\left(1, \upsilon_2(\hat{\beta}^{*} \mid data) / \upsilon_2(\hat{\beta}^{k\text{-}1} \mid data)\right).$$

$Step\,9:$ Generate $U_2 : Uniform(0,1)$

$Step\,10:$ If $U_2 \le Acceptance\ Probability(\hat{\beta}^{k\text{-}1}, \hat{\beta}^{k})$, set $\hat{\beta}^{*} = \hat{\beta}^{k}$, otherwise $\hat{\beta}^{k\text{-}1} = \hat{\beta}^{k}$

$Step\,11:$ Repeat Steps 2-10; for N time.

$Step\,12:$ The Bayes estimates of the parameters under $L_1(\theta),\ L_2(\theta),\ L_3(\theta)$, $L_4(\theta)$ and $L_5(\theta)$ can be written as:

$$\hat{\theta}_{SELo} = \frac{1}{N-M} \sum_{K=M+1}^{N} \theta^{(K)}; \quad \theta = (\alpha, \beta)\ , \tag{32}$$

$$\hat{\theta}_{WSEr} = \left[\frac{1}{N-M} \sum_{K=M+1}^{N} 1/\theta^{(K)} \right]^{-1}; \quad \theta = (\alpha, \beta)\,, \tag{33}$$

$$\hat{\theta}_{KLFu} = \sqrt{\frac{\sum_{K=M+1}^{N} \theta^{(K)}}{\sum_{K=M+1}^{N} \frac{1}{\theta^{(K)}}}}; \quad \theta = (\alpha, \beta)\,, \tag{34}$$

$$\hat{\theta}_{PLFu} = \sqrt{\frac{1}{N-M} \sum_{K=M+1}^{N} \theta^{2(K)}}; \quad \theta = (\alpha, \beta)\,, \tag{35}$$

and

$$\hat{\theta}_{LLFu} = -\frac{1}{\varphi} \log\left\{ \frac{1}{N-M} \sum_{K=M+1}^{N} e^{-\varphi \theta^{(K)}} \right\}; \quad \theta = (\alpha, \beta)\,. \tag{36}$$

Where, M is the burn-in-period of Markov chain.

6. Insurance Loss Data Analysis

This section provides an example of the suggested *VaR* and estimation approaches utilizing real-insurance loss data set. The data set is taken from Riad et al. (2022). To gain the fit of the *HTW* distribution to data set, we apply the log-likelihood (*LL*) criterion, Kolmogorov-Smirnov (*KS*) criterion, Anderson-Darling

(*AD*) criterion and Cramer-von Mises (*CVM*) criterion. The proposed distribution has been compared with five other distributions.

- *HTW* distribution:
- $LL_{HTW} = 27.1960,\ KS_{HTW} = 0.10856,\ AD_{HTW} = 0.88899,\ CVM_{HTW} = 0.13269$.
- *Weibull* distribution:
- $LL_W = 25.5290,\ KS_W = 0.12484,\ AD_W = 1.19055,\ CVM_W = 0.18564.$
- *Burr* distribution:
- $LL_B = 27.5486,\ KS_B = 0.11061,\ AD_B = 0.865653,\ CVM_B = 0.137021$.
- *Pareto* distribution:
- $LL_P = \text{-}309.4259,\ KS_P = 0.608971,\ AD_P = 28.009168,\ CVM_P = 6.245090$.
- *Exponentiated Pareto*:
- $LL_{EP} = \text{-}297.0731,\ KS_{EP} = 0.382,\ AD_{EP} = 13.582,\ CVM_{EP} = 2.6939$.
- *Exponential* distribution:
- $LL_E = \text{-}305.6592,\ KS_E = 0.490299,\ AD_E = 17.18987,\ CVM_E = 3.64775$.

It is observed that the *HTW* distribution better fits than any of these distributions. The maximum likelihood estimates of the proposed model parameters are plotted in Figure 3. This visualization is showing how the likelihood function value changes as the parameters α and β vary. In maximum likelihood estimation, we are typically looking for the parameter combination that minimizes the negative of this function. We also present the goodness-of-fit graphically by presented the empirical CDF and the fitted CDF plot, PP plot and also density plot, which clearly demonstrate that the *HTW* distribution adequately fits the data set (see, Figures 4, 5 and 6). Based on the following considerations, the performance of the VaR has been estimated.

- **Step 1:** Estimation of model parameters by maximum likelihood method (Figure 5) as initial values in the MH algorithm.
- **Step 2:** Obtain the Bayesian estimates of α and β using different loss function, notably, *SELo*, *WSEr* , *KLFu* and *PLFu*.
- **Step 3:** Determine the convergence property of the MCMC using the trace plots.
- **Step 4:** Evaluate the VaR measure using

$$VaR_T(X) = \left[-\ln\left(\frac{\hat{\beta}}{1-\hat{\beta}+(1-T)^{-1/\hat{\beta}}} \right) \right]^{1/\hat{\alpha}} \tag{37}$$

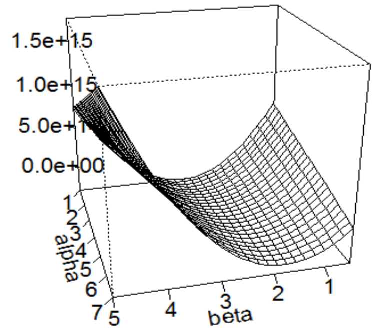


Figure 3. Plot for ML estimates of parameters

Source: *Research Findings*

The libraries that we used for the modeling process are:

- Function "*fitdist*" for "*fitdistrplus*" package: it is used to fit a probability distribution to a given dataset.
- Function "*gofstat*" for "*fitdistrplus*" package: we use this library in application part for calculating the goodness-of-fit tests with its associated p-value.
- Function "*MCMC*" for "*coda*" package: In the context of Bayesian analysis with Markov Chain Monte Carlo (MCMC) methods, we use this function to convert the sequence of samples from the posterior distribution of a parameter into an MCMC object.

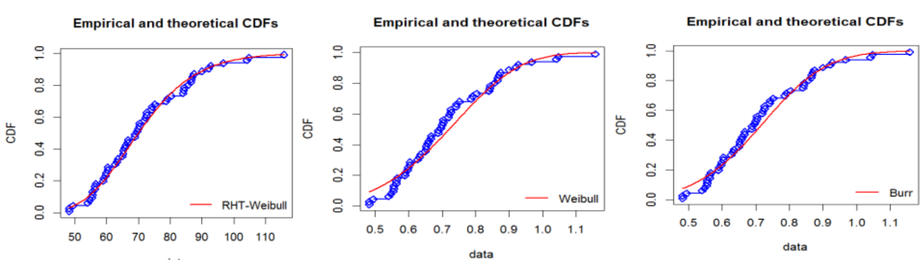


Figure 4: Empirical CDF and the fitted CDF plots .

Source: *Research Findings*

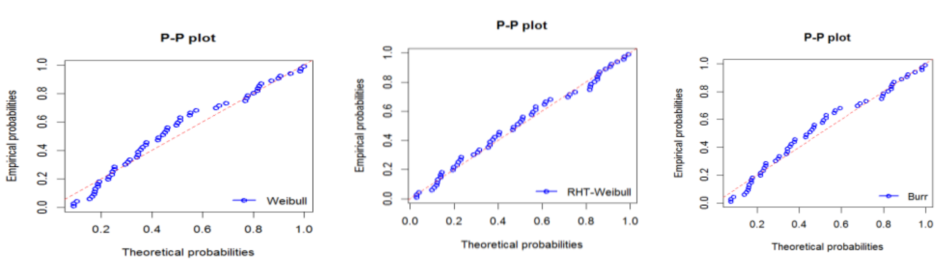


Figure 5. The PP plots .

Source: *Research Findings*

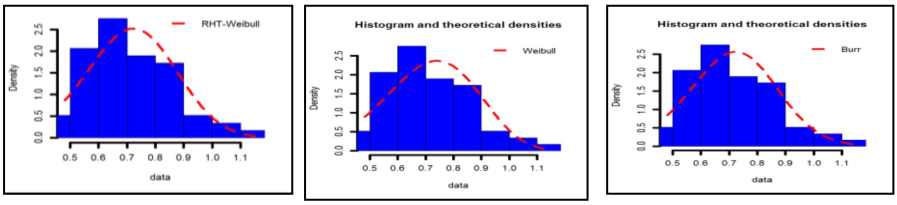


Figure 6. The density and fitted plot for different distributions.
Source: *Research Findings*

The Bayesian estimates are obtained under the non-informative prior assumption. Since we do not know the true value of the parameters and we use iterative estimation as the initial value, we used a large number of iterations to obtain a stable chain, and to eliminate the influence of the initial values, we used the first 5000 iterations as the burn-in period. For the MCMC method, we sample from the posterior distributions of the parameters using the MH algorithm with $N = 25000$. The results of the parameters estimation and risk evaluation are reported in Tables 2 and 3 respectively. To check the convergence of the MCMC samples in Bayes estimation of parameters, the trace plots of the MCMC samples are presented in Figures 7 and 8. It is worth noting that the plots are shown for two loss functions. The rest of the cases are similar. The blue line shows how the parameter value fluctuates across iterations. The values fluctuate around the center line, suggesting the chain has likely converged. The histogram plots are also presented. It is clear that the MCMC method converges extremely effectively.

Therefore, we can conclude that the VaR measure can be obtained satisfactorily based on the HTW model and using the proposed methods.

Table 2. Different estimates using different LFu.

Parameters	$\hat{\theta}_{SELo}$	$\hat{\theta}_{WSEr}$	$\hat{\theta}_{KLFu}$	$\hat{\theta}_{PLFu}$	$\hat{\theta}_{LLFu}$
$\hat{\alpha}$	5.706	5.811	6.331	5.309	5.694
$\hat{\beta}$	1.259	1.112	1.078	1.020	1.785

Source: *Research Findings*

VaR	$\hat{\theta}_{SELo}$	$\hat{\theta}_{WSEr}$	$\hat{\theta}_{KLFu}$	$\hat{\theta}_{PLFu}$	$\hat{\theta}_{LLFu}$

$T = 0.85$	1.0350	1.0818	1.0844	1.1213	0.7303
$T = 0.95$	1.0805	1.1218	1.1202	1.1636	0.9892
$T = 0.99$	1.1412	1.1771	1.1697	1.2234	1.1233

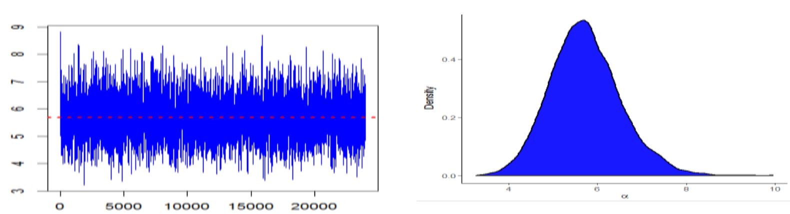
Table 3. The VaR estimations using different LFu.

Source: *Research Findings*

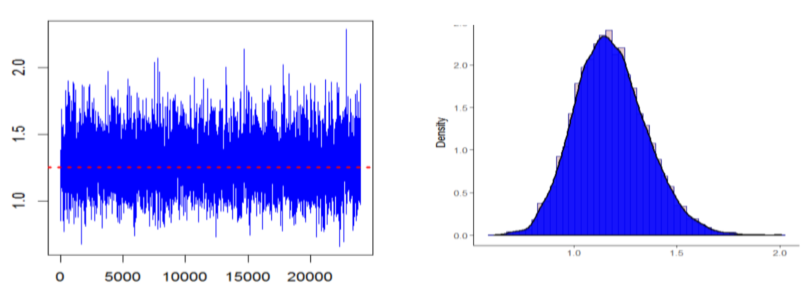
7. Conclusion

This paper proposes a heavy-tailed distribution, known as *HTW* model. In the context of analysis of financial data, it is a significant and novel contribution to the field of risk modeling. This new distribution has several distributional and mathematical properties. It was focused on obtaining the Bayes estimates of the parameters of the *HTW* distribution based on different loss functions. In the Bayesian framework, the Bayes estimates of the unknown parameters were computed using the MH algorithm. Bayes estimates are obtained with a variety of symmetric and asymmetric loss functions. A thorough exploration is presented on the implementation of the MH algorithm for generating samples from the posterior distribution. The Bayesian estimates are applied to VaR evaluation. Our analysis using insurance loss data confirms that the *HTW* distribution can effectively model real-world financial phenomena characterized by extreme events. Also, we have compared the goodness of fit of this data with other distributions used in financial issues, and it is found that the *HTW* model provides better inferences for this data compared to all suggested models. Also, Bayesian estimates based on different loss functions have led to satisfactory risk estimation. The improved accuracy in tail risk estimation enables more informed capital allocation decisions and potentially reduces the likelihood of insolvency during market stress periods.

There are two interesting works for the future work. The first is to apply the *HTW* model to obtain tail value at risk ($TVaR$) and tail variance premium (TVP). Secondly, other estimators such as least square procedure and moment method can be adopted to estimate the model parameters and then to evaluate the VaR criteria.



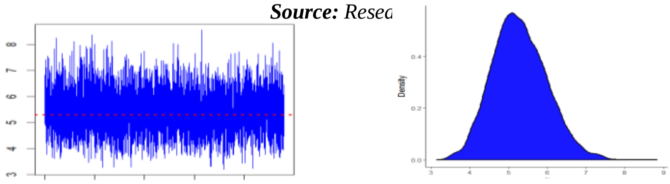
(a) Plots for α



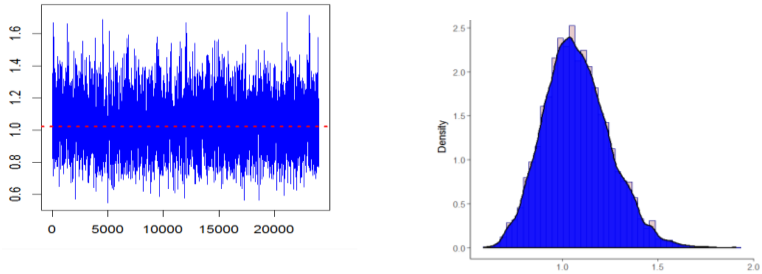
(b) Plots for β

Figure 7.The histogram and trace plots based on SELo.

Source: *Resea*



(a) Plots for α



(b) Plots for β

Figure 8. The histogram and trace plots based on PLFu.

Source: *Research Findings*

Author Contributions

Conceptualization, all authors; methodology all authors; formal analysis, all authors; resources, all authors; writing—original draft preparation, all authors; writing—review and editing, all authors. All authors have read and agreed to the published version of the manuscript.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data used is available in the paper.

Acknowledgements

Not applicable

References

Alzaatreh, A., Lee, C., & Famoye F. (2013). A new method for generating families of continuous distributions. Metron, 71(1), 63–79.

Asadi, A., Zare, H., Ebrahimi, M. & Piraiee, K. (2018). Sentiment Shock and Stock Price Bubbles in a Dynamic Stochastic General Equilibrium Model Framework: The Case of Iran. Iranian Journal of Economic Studies, 7(2), 115-150.

Bali, T.G., (2007). A Generalized Extreme Value Approach to Financial Risk Measurement, Journal of Money, Credit and Banking, Blackwell Publishing, 39(7), 1613-1649.

Benkhelifa, L., (2021a). The Weibull Birnbaum-Saunders Distribution And Its Applications. Statistics, Optimization & Information Computing, 9(1), 61-81.

Benkhelifa, L. (2021b). The Beta Reduced Modified Weibull Distribution with Applications to Reliability Data. Journal of Reliability and Statistical Studies, 14, 323–352.

Bebbington, M., Lai, C.D. & Zitikis, A. (2007). A flexible Weibull extension. Reliability Engineering and System Safety, 92:719-726.

Chronopoulos, I., Raftapostolos , A., & Kapetanios, G. (2024). Forecasting Value-at-Risk Using Deep Neural Network Quantile Regression, Journal of Financial Econometrics, 22(3), 636–669.

Chinhamu, K., Huang, C.K., Huang, C.S., & Chikobvu, D., (2015). Extreme risk, value-at-risk and expected shortfall in the gold market, International Business & Economics Research Journal, 14 (1),107–122.

Riad, F, Hussam, E, Gemeay, A, Aldallal,R, &Afify, A. (2022). Classical and Bayesian inference of the weighted-exponential distribution with an application to insurance data, Mathematical Biosciences and Engineering, 19(7), 6551-6581.

Lai, C.D., Xie, M., & Murthy, D.N.P., (2003). A modified Weibull distribution. IEEE Transactions on Reliability, 52(1):33-37.

Luger, R., (2012). Finite-sample bootstrap inference in GARCH models with heavy-tailed innovations, Computational Statistics & Data Analysis, 56(11), 3198-3211.

Famoye, F., Lee, C. & Olumolade, O. (2005). The beta-Weibull distribution. Journal of Statistical Theory and Applications, 4(2):121-136.

Ghazal, M.G.M., & Radwan, H.M.M. (2022). A reduced distribution of the modified Weibull distribution and its applications to medical and engineering data, 19, 13193-13213.

Mahdavi, P. Ehsani, M.A. (2022). Dynamic Causal Effects in Econometrics with a Focus on the Nonparametric Method: A Review Paper, Iranian Journal of Economic Studies, 11(2), 427-449.

Martın, J., Parra, M.I., Pizarro, M.M., & Sanjuan, E.L. (2022). Baseline methods for the parameter estimation of the generalized Pareto distribution, Entropy, 24 (2), 178.

Müller, F.M., & Brutti Righi, M. (2024). Comparison of Value at Risk (VaR), Multivariate Forecast Models, 63, 75–110.

Moradi, N., Panahi, H., & Habibirad, A. (2022). Estimation for the Three-Parameter Exponentiated Weibull Distribution under Progressive Censored Data, Journal of the Iranian Statistical Society, 21 (1), 153-177.

Nelder, J. (1965). A simplex algorithm for function minimization. Computational Journal, 7, 308–313.

Panahi, H. (2019). Value at Risk Estimation using the Kappa Distribution with Application to Insurance Data, International Journal of Finance & Managerial Accounting, 4 (14), 91-100.

Panahi, H. (2025). Statistical inferences of reliability and order-restricted nano-droplet rebound by comparative Kumaraswamy populations based on balanced joint progressive censoring. Physica Scripta, 100(2), 025024.

Peng, S., Yang, S., & Yao, J. (2020). Improving Value-at-Risk prediction under model uncertainty, Journal of Financial Econometrics, 7, 34-42.

Socgnia, V.K., & Wilcox, D. (2014). A Comparison of Generalized Hyperbolic Distribution Models for Equity Returns, Journal of applied Mathematics, 4, 55-72.

Sarhan A.M., & Apaloo, J. (2013). Exponentiated modified Weibull extension distribution. Reliability Engineering and System Safety, 112:137-144.

Spierdijk, S. (2016). Confidence intervals for ARMA–GARCH Value-at-Risk: The case of heavy tails and skewness, Computational Statistics & Data Analysis, 100, 545-559.

Trzpiot, G., & Majewska, J. (2010). Estimation of value at risk: Extreme value and robust approaches. Operations Research and Decisions, 20 (1), 131–143.

Vardani, M. H., Panahi, H., & Behzadi, M. H. (2024). Statistical inference for marshall-olkin bivariate Kumaraswamy distribution under adaptive progressive hybrid censored dependent competing risks data. Physica Scripta, 99(8), 085272.

Zhao, W., Khosa, S.K., Ahmad, Z., Aslam, M., & Afify, A.Z. (2020). Type-I heavy tailed family with applications in medicine, engineering and insurance. PLoS ONE, 15(8): e0237462.

Appendix

The R programming code for the Linex loss function plot and MCMC steps have been provided below:

```
linex_loss <- function(delta, phi) {return((exp(phi * delta) - phi * delta - 1))}
delta <- seq(-2, 2, length.out = 200)
phi_pos <- c(0.5, 1.5, 2.5)
```

```
phi_neg <- -phi_pos
par(mfrow = c(1,2))
f1 <- function(alpha, beta, x) {n <- length(x)}
if(alpha <= 0 || beta <= 0) return(0)
alpha_part1 <- alpha^(n+r1-1)
    alpha_part2 <- prod(x^(alpha-1))
    alpha_part3 <- exp(-alpha * tu1)
    alpha_part4 <- exp(- alpha * (tu1-sum(log(x))))
    beta_part2 <- prod((beta + (1-beta)*exp(-x^alpha))^(-beta-1))
    J <- alpha_part1 * alpha_part2 * alpha_part3 * beta_part2
 if(is.nan(J) || is.na(J) || is.infinite(J)) return(0)
    return(J)
  }, silent=TRUE)
  return(0)
f2 <- function(alpha, beta, x) {
  n <- length(x)
  if(alpha = 0 || beta <= 0) return(0)
  try({
    beta_part1 <- beta^(2*n+r2-1)
    beta_part2 <- exp(-beta*tu2)
    beta_part3 <- prod((beta + (1-beta)*exp(-x^alpha))^(-beta-1))
   J <- beta_part1 * beta_part2 * beta_part3
    if(is.nan(J) || is.na(J) || is.infinite(J)) return(0)
    return(J)
  }, silent=TRUE)
 return(0)
}
Metropolis <- function(sigma) {
  M1 <- 5000
  beta1 <- alpha1 <- rep(NA, M1)
  alpha1[1] <- 1
  beta1[1] <- 1
  u <- runif(M1)
  k_alpha <- k_beta <- 0

  for(l in 2:M1) {
    alphat <- alpha1[l-1]
    betat <- beta1[l-1]
    y_alpha <- rnorm(1, alphat, sigma)
    ratio_alpha - 0
    if(y_alpha > 0) {
      num <- f1(y_alpha, betat, x) * dnorm(alphat, y_alpha, sigma)
      denom <- f1(alphat, betat, x) * dnorm(y_alpha, alphat, sigma)
      if(denom > 0) ratio_alpha <- num/denom
```

```
    if(is.nan(ratio_alpha) || is.na(ratio_alpha) || is.infinite(ratio_alpha))
ratio_alpha <- 0
   }
    if(y_alpha > 0 && !is.na(ratio_alpha) && u[l] <= ratio_alpha)
     alpha1[l] <- y_alpha
   else {
     alpha1[l] <- alphat
     k_alpha <- k_alpha + 1
   }
   y_beta <- rnorm(1, betat, sigma)
   ratio_beta <- 0
   if(y_beta > 0) {
     num <- f2(alpha1[l], y_beta, x) * dnorm(betat, y_beta, sigma)
     denom <- f2(alpha1[l], betat, x) * dnorm(y_beta, betat, sigma)
       if(denom > 0) ratio_beta <- num/denom
     if(is.nan(ratio_beta) || is.na(ratio_beta) || is.infinite(ratio_beta)) ratio_beta
<- 0
   }
   if(y_beta > 0 && !is.na(ratio_beta) && u[l] <= ratio_beta)
     beta1[l] <- y_beta
   else {
     beta1[l] <- betat
     k_beta <- k_beta + 1
  k_alpha <- k_alpha / M1
 k_beta <- k_beta / M1
 return(list(alpha1=alpha1, beta1=beta1, k_alpha=k_alpha, k_beta=k_beta))
}
# Run with different sigma values to find best sigma
MH <- function(s) {
 M1=2000; burn=500; sigma=s
 beta1 <- alpha1 <- rep(NA, (M1+burn))
 alpha1[1] <- 1
 beta1[1] <- 5
 u <- runif(M1+burn)
 k_alpha <- k_beta <- 0

 for(l in 2:(M1+burn)) {
   alphat <- alpha1[l-1]
   betat <- beta1[l-1]
   y_alpha <- rnorm(1, alphat, sigma)
   if(y_alpha > 0) {
     num <- f1(y_alpha, betat, x) * dnorm(alphat, y_alpha, sigma)
     denom <- f1(alphat, betat, x) * dnorm(y_alpha, alphat, sigma)
     if(denom > 0) ratio_alpha <- num/denom
```

```
    if(is.nan(ratio_alpha) || is.na(ratio_alpha) || is.infinite(ratio_alpha))
ratio_alpha <- 0
   }
   if(y_alpha > 0 && !is.na(ratio_alpha) && u[l] <= ratio_alpha)
     alpha1[l] <- y_alpha
   else {
     alpha1[l] <- alphat
     k_alpha <- k_alpha + 1
   }
   y_beta <- rnorm(1, btat, sigma)
    ratio_beta <- 0
   if(y_beta > 0) {
     num <- f2(alpha1[l], y_beta, x) * dnorm(betat, y_beta, sigma)
     denom <- f2(alpha1[l], betat, x) * dnorm(y_beta, betat, sigma)
      if(denom > 0) ratio_beta <- num/denom
     if(is.nan(ratio_beta) || is.na(ratio_beta) || is.infinite(ratio_beta)) ratio_beta
<- 0
   }
    if(y_beta > 0 && !is.na(ratio_beta) && u[l] <= ratio_beta)
     beta1[l] <- y_beta
   else {
     beta1[l] <- betat
     k_beta <- k_beta + 1
  alpha1 <- alpha1[(burn+1):(burn+M1)]
 beta1 <- beta1[(burn+1):(burn+M1)]
   alpha_geometric <- sqrt(mean(alpha1)/mean(1/alpha1))
 beta_geometric <- sqrt(mean(beta1)/mean(1/beta1))
 alpha_ratio <- mean(1/alpha1)/mean(1/(alpha1^2))
 beta_ratio <- mean(1/beta1)/mean(1/(beta1^2))
  alpha_linex <- (-1/h) * log(mean(exp(-h*alpha1)))
 beta_linex <- (-1/h) * log(mean(exp(-h*beta1)))
   names_estimators <- c( "l3", "l4", "l5")
  results <- data.frame(Estimator = names_estimators,
Alpha = alpha_estimators,
   Beta = beta_estimators
 )
 rejection_rates <- c(k_alpha/(M1+burn), k_beta/(M1+burn))
 return(list(estimators = results))
```