ARTICLE

Type-I Heavy-Tailed Burr XII Distribution with Applications to Quality Control, Skewed Reliability Engineering Systems and Lifetime Data

Okechukwu J. Obulezi[1,*], Hatem E. Semary[2], Sadia Nadir[3], Chinyere P. Igbokwe[4], Gabriel O. Orji[1], A. S. Al-Moisheer[2] and Mohammed Elgarhy[5]

[1]Department of Statistics, Faculty of Physical Sciences, Nnamdi Azikiwe University, Awka, P.O. Box 5025, Nigeria
[2]Department of Mathematics and Statistics, Faculty of Science, Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh, 11432, Saudi Arabia
[3]Department of Mathematics and Statistics, Faculty of Engineering and Applied Sciences, Riphah International University, Islamabad, 44000, Pakistan
[4]Department of Statistics, School of Computer Science and Engineering, Lovely Professional University, Phagwara, 144411, Punjab, India
[5]Department of Basic Sciences, Higher Institute of Administrative Sciences, Belbeis, AlSharkia, 44621, Egypt
*Corresponding Author: Okechukwu J. Obulezi. Email: oj.obulezi@unizik.edu.ng

ABSTRACT: This study introduces the type-I heavy-tailed Burr XII (TIHTBXII) distribution, a highly flexible and robust statistical model designed to address the limitations of conventional distributions in analyzing data characterized by skewness, heavy tails, and diverse hazard behaviors. We meticulously develop the TIHTBXII's mathematical foundations, including its probability density function (PDF), cumulative distribution function (CDF), and essential statistical properties, crucial for theoretical understanding and practical application. A comprehensive Monte Carlo simulation evaluates four parameter estimation methods: maximum likelihood (MLE), maximum product spacing (MPS), least squares (LS), and weighted least squares (WLS). The simulation results consistently show that as sample sizes increase, the Bias and RMSE of all estimators decrease, with WLS and LS often demonstrating superior and more stable performance. Beyond theoretical development, we present a practical application of the TIHTBXII distribution in constructing a group acceptance sampling plan (GASP) for truncated life tests. This application highlights how the TIHTBXII model can optimize quality control decisions by minimizing the average sample number (ASN) while effectively managing consumer and producer risks. Empirical validation using real-world datasets, including "Active Repair Duration," "Groundwater Contaminant Measurements," and "Dominica COVID-19 Mortality," further demonstrates the TIHTBXII's superior fit compared to existing models. Our findings confirm the TIHTBXII distribution as a powerful and reliable alternative for accurately modeling complex data in fields such as reliability engineering and quality assessment, leading to more informed and robust decision-making.

KEYWORDS: Acceptance sampling; heavy-tailed models; parameter estimation; reliability engineering

1 Introduction

In the analysis of real-world phenomena, rare events and complex hazard behaviors are common, necessitating the use of flexible statistical distributions. Classical models such as the normal, exponential, Weibull, and Rayleigh often prove inadequate in modeling data characterized by significant skewness, heavy tails, or non-monotonic hazard functions. Consequently, the development of more versatile distributions, particularly those with heavy-tailed behavior, has been a central focus of modern applied statistics.

Heavy-tailed distributions, which allocate a higher probability mass to extreme outcomes than Gaussian models, have been observed across a wide range of disciplines. In finance, for example, asset prices and stock returns frequently exhibit fat-tailed tendencies, with the probability of large gains or losses exceeding what would be predicted by a normal distribution [1]. This has led to the adoption of stable laws to better explain high market volatility. Similarly, in actuarial science, catastrophic claims in insurance are a classic example of heavy-tailed behavior, particularly when modeling loss severity [2]. Environmental and natural systems also display these characteristics; hydrological records, rain intensities, and seismic events often follow distributions with slowly decaying tails [3,4]. In medicine and biology, skewed and extreme-value data, such as tumor sizes and survival times, similarly require heavy-tailed models for accurate analysis [5–7].

Formally, a distribution with distribution function $F(x)$ is defined as heavy-tailed if its survival function $\bar{F}(x) = 1 - F(x)$ converges to zero more slowly than any exponential rate, that is,

$$\lim_{x \to \infty} \frac{\bar{F}(x)}{e^{-\lambda x}} = \infty \quad \text{for all } \lambda > 0.$$

This implies that heavy-tailed distributions place significantly more probability mass in the tails than exponential-type distributions. Classic examples include the Cauchy distribution by Cauchy [8], Student's t distribution by Student [9], Fréchet distribution by Fréchet [10], and the Pareto family of distributions by Pareto [11], and Arnold [12], which includes the Pareto II (Lomax) distribution by Lomax [13]. A specific and widely-studied type of heavy-tailed distribution is the power-law distribution, characterized by a survival function that decays as a power of x for large x:

$$\bar{F}(x) \sim x^{-\alpha} \quad \text{as } x \to \infty,$$

where $\alpha > 0$ is the tail index. This slow, polynomial decay of the tail is a defining feature that distinguishes power-law distributions from those with exponential tails. Given that empirical data frequently violate the assumptions of standard models like the exponential, gamma, and Weibull distributions, researchers have increasingly turned to advanced generalization strategies such as transformation techniques, compounding, and mixture models to construct distributions with superior tail behavior and greater structural flexibility, see [14–16] as well as [17–20].

Among these flexible models, the Burr Type XII (BXII) distribution has been particularly influential. Introduced by Burr [21] as part of a system of distributions designed to fit a wide range of empirical data, it is a continuous probability distribution for a non-negative random variable $x \geq 0$. Singh [22] demonstrated the applicability of Burr-type distributions in different data scenarios. The conventional Burr Type XII distribution is characterized by two positive shape parameters, commonly represented as c and k, and its high degree of flexibility is well-known. By altering the shape parameters, the BXII distribution can take on many forms and approximate numerous other common distributions, such as the log-logistic, Lomax (Pareto Type II), and Weibull, and can even approximate the normal distribution under certain conditions. It offers a parsimonious yet powerful framework for analyzing data across diverse fields, including reliability, survival analysis, and econometrics.

The ongoing demand for models capable of capturing complex data characteristics has led to a proliferation of BXII generalizations. These extensions often enhance the original BXII's flexibility by adding new parameters or by embedding it within a larger family of distributions. For example, Gad et al. [23] introduced the Burr XII-Burr XII (BXI-BXII) distribution, a compound model that demonstrated a superior fit to certain datasets compared to the baseline BXII. The literature is rich with such constructions, which can be broadly categorized by their methodological approach.

One common strategy is the use of general families of distributions. The T-X family by Alzaatreh et al. [24], for instance, has been used to derive the Burr XII-moment exponential (BXII-ME) distribution, which offers greater flexibility for modeling lifetime data [25]. Similarly, the Kumaraswamy-G family has led to the Kumaraswamy Burr XII (KBXII) distribution, a four-parameter model with a remarkable ability to accommodate diverse functional shapes and hazard rate behaviors [26]. Another important approach involves the Marshall-Olkin extended family, which produced the Marshall-Olkin extended Burr Type XII (MOEBXII) distribution to provide enhanced flexibility for reliability and survival analysis [27]. For more details see [28–30] as well as [31–33].

Other specialized extensions of the BXII distribution have been developed to address specific challenges. Ocloo [34] introduced a novel extension of the Burr XII distribution, demonstrating its enhanced capability to model data with complex characteristics such as bimodality and heavy tails, and providing a better fit than many existing BXII extensions. Recent innovations also include the Sine-Burr XII distribution, which utilizes a trigonometric transformation to add flexibility without introducing new parameters [35], and the Odd Burr XII Gompertz (OBXIIGo) distribution, which provides a more adaptable model for lifetime data by integrating the BXII family with the Gompertz distribution [36].

The work of Gad et al. [23] and other similar studies underscore a critical motivation in modern statistical modeling: the development of new distributions that can effectively capture complex data structures not adequately described by classical models. The literature reviewed herein indicates a clear trend toward creating higher-order BXII-based models that offer greater flexibility in their distributional shapes and hazard rate functions. These models consistently serve as a better alternative to their classical counterparts, a finding supported by statistical goodness-of-fit metrics.

One such area where the choice of distribution is paramount is in acceptance sampling, particularly within truncated life-testing experiments. In specialized manufacturing or medical diagnosis, using simple, stratified, or cluster sampling may be insufficient. Group acceptance sampling plans (GASPs) provide an efficient alternative by examining several units in batches, conserving resources. The efficacy of a GASP, however, hinges on the choice of the underlying product lifetime distribution. An ill-suited model can lead to misclassifying good or bad lots, resulting in either excessive consumer risk or uneconomical producer rejection. This has prompted researchers to propose GASPs under more generalized distributions to overcome the limitations of standard models. For example, GASPs have been developed for truncated life tests based on the inverse Rayleigh and log-logistic distributions by Aslam & Jun [37], the Marshall-Olkin extended Lomax distribution by Rao [38], and the inverse Weibull distribution with median lifetime as a quality measure by Singh & Tripathi [39]. Other recent studies have employed generalized distributions like the AG-transmuted exponential by Almarashi & Khan [40], transmuted exponential by Owoloko et al. [41], transmuted Rayleigh by Saha et al. [42], and exponential-logarithmic distributions by Ameeq et al. [43] in the GASP framework. This body of research demonstrates that superior modeling of the lifetime distribution directly leads to more efficient and risk-balanced sampling decisions. Most recently, Ekemezie et al. [44] presented the odd Perks-Lomax (OPL) distribution and used it to construct a GASP, showcasing how a novel, versatile distribution can enhance both statistical theory and practical application simultaneously.

Building on this demonstrated need for more intricate models, this research proposes and thoroughly investigates a novel statistical distribution: the Type-I Heavy-Tailed Burr Type XII (TIHTBXII) distribution. The primary motivation is to address the persistent challenges in adequately modeling real-world data with extreme skewness, heavy tails, and complex, non-monotonic hazard functions, for which existing generalized distributions still exhibit limitations. We not only meticulously derive the mathematical properties of the TIHTBXII distribution and examine various parameter estimation procedures but, more significantly, we demonstrate its practical utility by constructing a group acceptance sampling plan (GASP) specifically

tailored to truncated life tests based on this new model. This new tool leverages the TIHTBXII distribution's excellent capability to provide more accurate and reliable data, thereby enabling better decision-making in fields like quality control, reliability engineering, and lifetime data analysis, where current models may fail to capture the true underlying risk mechanisms.

The rest of this paper is organized as follows: In Section 2, we define the basic functions of the proposed TIHTBXII model, including its probability density function (PDF), cumulative distribution function (CDF), survival function, and hazard function. Section 3 is dedicated to the derivation of several key characteristics of the TIHTBXII model, such as its quantile function, moments, incomplete moments, moment-generating function, mean residual life function, entropy, extropy, and order statistics. In Section 4, we discuss the estimation of the model parameters using maximum likelihood, maximum product spacing, least squares, and weighted least squares methods. Section 5 presents a Monte Carlo Markov Chain (MCMC) simulation study conducted with four different parameter settings, with the bias and root mean square error (RMSE) also plotted to support the numerical results. In Section 6, we design a group acceptance sampling plan (GASP) to illustrate how the TIHTBXII model can optimize quality control decisions by minimizing the average sample number. Finally, Section 7 reports the applications of the proposed model to real datasets, including the duration of active repairs of airborne communication transceivers, groundwater contaminant measurements, and Dominica COVID-19 mortality rate data, while Section 8 provides the conclusion of this study.

2 Development of the Type-I Heavy-Tailed Burr XII Distribution

The Burr XII distribution has cumulative distribution function (CDF) and probability density function (PDF), respectively, given as

$$F(x) = 1 - \frac{1}{(1 + x^c)^k}; \quad x > 0, \quad c, k > 0, \tag{1}$$

and

$$f(x) = \frac{ckx^{c-1}}{(1 + x^c)^{k+1}}. \tag{2}$$

Reference [45] proposed the Type I heavy-Tailed (TI-HT) family of distributions with CDF and PDF respectively given as

$$G(x; \theta, \xi) = 1 - \left(\frac{1 - F(x; \xi)}{1 - (1 - \theta)F(x; \xi)} \right)^{\theta}; \quad x \in \mathfrak{R}, \quad \theta > 0, \tag{3}$$

and

$$g(x; \theta, \xi) = \frac{\theta^2 f(x; \xi) \{1 - F(x; \xi)\}^{\theta-1}}{\{1 - (1 - \theta)F(x; \xi)\}^{\theta+1}}, \tag{4}$$

where $F(x; \xi)$ is the baseline distribution which depends on $\xi \in \mathfrak{R}$. Introducing Eqs. (1) and (2) into (3) and (4) produces a special distribution referred to as type-I heavy-tailed Burr XII (TIHTBXII) distribution with CDF and PDF respectively defined as

$$G(x; \theta, c, k) = 1 - \left[\frac{1}{1 - \theta \left\{ 1 - (1 + x^c)^k \right\}} \right]^\theta ; \quad x > 0, \quad \theta, c, k > 0, \tag{5}$$

and

$$g(x; \theta, c, k) = \frac{\theta^2 c k x^{c-1} (1 + x^c)^{k-1}}{\left[1 - \theta \left\{ 1 - (1 + x^c)^k \right\} \right]^{\theta+1}}. \tag{6}$$

The survival function is

$$s(x; \theta, c, k) = \left[\frac{1}{1 - \theta \left\{ 1 - (1 + x^c)^k \right\}} \right]^\theta ; \quad x > 0, \quad \theta, c, k > 0. \tag{7}$$

The hazard function is

$$h(x; \theta, c, k) = \frac{\theta^2 c k x^{c-1} (1 + x^c)^{k-1}}{\left[1 - \theta \left\{ 1 - (1 + x^c)^k \right\} \right]^1}; \quad x > 0, \quad \theta, c, k > 0. \tag{8}$$

The proposed TIHTBXII distribution aims to improve model flexibility, provide the best fit to real-world data, and accommodate heavy-tailed data in fields such as reliability engineering, medical and financial sciences, among others.

Fig. 1a and b represents the graph of the PDF and hazard function, respectively. These plots are at various combination of the parameter values. The plots reflect left-skewed behaviour, both low and high peaks, and bumped-shape. This demonstrate that the TIHTBXII can be utilized in modeling different lifetime events. The hazard rate plots depict bump-shape, bathtub, L-shape, J-shape and strictly non-decreasing shape. This again implies that TIHTBXII model can benefit from modeling different lifetime events.
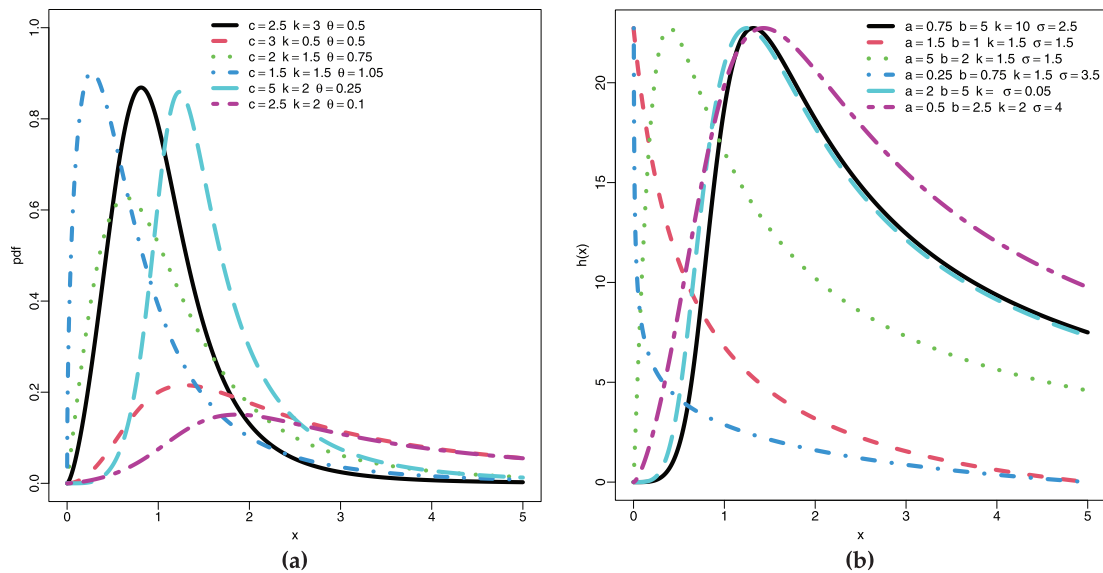

Figure 1: Plots of **(a)** PDF, **(b)** Hazard function of the TIHTBXII distribution

3 Characterization of the TIHTBXII Model

In this section, we study some basic properties of the new distribution which include quantile function, moment and incomplete moment, moment generating function, order statistic, entropy and extropy, and the mean residual life function.

3.1 Quantile Function

The quantile function $Q(u) = G^{-1}(u)$, for $0 < u < 1$, is derived as:

$$Q(u) = \left\{ \left[1 - \left(\frac{1 - (1-u)^{-\frac{1}{\theta}}}{\theta} \right) \right]^{\frac{1}{k}} - 1 \right\}^{\frac{1}{c}} . \tag{9}$$

In inference, the advantage of $Q(u)$ is in generating random samples and overall simulation studies. Further, various quantile measures can be obtained such as the mean, variance, skewness and kurtosis.

3.2 Moment

The r-th crude moment of a continuous random variable X which follows any distribution say TIHTBXII (θ, c, k) can be expressed as follows

$$\mu_r' = \mathbb{E}(X^r) = \int_{-\infty}^{\infty} x^r g(x)\, dx = \int_0^{\infty} \frac{\theta^2 c k x^{r+c-1} (1+x^c)^{k-1}}{\left[1 - \theta \left\{ 1 - (1+x^c)^k \right\} \right]^{\theta+1}}\, dx$$

$$= \theta^2 c k \int_0^{\infty} x^{r+c-1} (1+x^c)^{k-1} \left[1 - \theta \left\{ 1 - (1+x^c)^k \right\} \right]^{-(\theta+1)}\, dx.$$

Using series expansion

$$(1+x^c)^{k-1} = \sum_{i=0}^{\infty} \binom{k-1}{i} x^{ic}; \quad \text{for} \quad |x| < 1.$$

For $\theta \in \mathbb{R}$ or $\mathbb{C}$, a generalized binomial theorem provides that

$$\left[1 - \theta \left\{ 1 - (1+x^c)^k \right\} \right]^{-(\theta+1)} = \sum_{j=0}^{\infty} \binom{\theta+j}{j} \theta^j \left[1 - (1+x^c)^k \right]^j.$$

Also, for $j \in \mathbb{R}/\mathbb{N}_{\nmid}$ and $|x| < 1$, standard binomial expansion provides that

$$\left[1 - (1+x^c)^k \right]^j = \sum_{l=0}^{\infty} (-1)^l \binom{j}{l} (1+x^c)^{lk} .$$

Similarly,

$$(1+x^c)^{lk} = \sum_{m=0}^{\infty} \binom{lk}{m} x^{mc}.$$

By back substitution,

$$\mu'_r = \theta^2 ck \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} (-1)^l \binom{k-1}{i}\binom{\theta+j}{j}\binom{j}{l}\binom{lk}{m}\theta^j \int_0^{\infty} x^{r+c+ic+mc-1} \, dx.$$

Notice that

$$\binom{j}{l} = \frac{j!}{l!(j-l)!} = \frac{(j,l)}{l!}; \quad \text{Recall} \quad e^{-x} = \sum_{l=0}^{\infty}(-1)^l \frac{x^l}{l!},$$

so that

$$\mu'_r = \theta^2 ck \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{m=0}^{\infty} \binom{k-1}{i}\binom{\theta+j}{j}\binom{lk}{m}\theta^j (j,l) \int_0^{\infty} \sum_{l=0}^{\infty}(-1)^l \frac{x^l}{l!}\frac{1}{x^l} x^{r+c+ic+mc-1} \, dx$$

$$= \theta^2 ck \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{m=0}^{\infty} \binom{k-1}{i}\binom{\theta+j}{j}\binom{lk}{m}\theta^j (j,l) \int_0^{\infty} x^{r+c+ic+mc-l-1} e^{-x} \, dx.$$

Therefore,

$$\mu'_r = \theta^2 ck \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{m=0}^{\infty} \binom{k-1}{i}\binom{\theta+j}{j}\binom{lk}{m}\theta^j (j,l)\Gamma(r+c+ic+mc-l); \quad r = 1, 2, \cdots \qquad (10)$$

3.3 Incomplete Moment

For a continuous random variable $X \sim \mathrm{TIHTBXII}(\theta, c, k)$, the *r-th incomplete moment* about the origin is given by:

$$\mu_r(x_0) = \mathbb{E}[X^r \mid X \le x_0] = \int_0^{x_0} x^r g(x; \theta, c, k) \, dx.$$

Substitute into the incomplete moment:

$$\mu_r(x_0) = \theta^2 ck \int_0^{x_0} x^{r+c-1}(1+x^c)^{k-1} \left[1 - \theta\left(1-(1+x^c)^k\right)\right]^{-(\theta+1)} \, dx.$$

Using the generalized binomial expansion:

$$(1+x^c)^{k-1} = \sum_{i=0}^{\infty} \binom{k-1}{i} x^{ic}.$$

Expanding the second term yields

$$\left[1 - \theta\left(1-(1+x^c)^k\right)\right]^{-(\theta+1)} = \sum_{j=0}^{\infty} \binom{\theta+j}{j}\theta^j \left[1-(1+x^c)^k\right]^j.$$

and

$$\left[1-(1+x^c)^k\right]^j = \sum_{l=0}^{\infty}(-1)^l \binom{j}{l}(1+x^c)^{lk}.$$

Also,

$$(1 + x^c)^{lk} = \sum_{m=0}^{\infty} \binom{lk}{m} x^{mc}.$$

The integrand becomes a power series in x:

$$x^{r+c-1} \cdot \sum_{i=0}^{\infty} \binom{k-1}{i} x^{ic} \cdot \sum_{j=0}^{\infty} \binom{\theta+j}{j} \theta^j \sum_{l=0}^{\infty} (-1)^l \binom{j}{l} \sum_{m=0}^{\infty} \binom{lk}{m} x^{mc}.$$

Now collect all powers of x, and integrate:

$$\mu_r(x_0) = \theta^2 ck \sum_{i,j,l,m} (-1)^l \binom{k-1}{i} \binom{\theta+j}{j} \theta^j \binom{j}{l} \binom{lk}{m} \int_0^{x_0} x^{r+c-1+ic+mc} \, dx.$$

$$\int_0^{x_0} x^{r+c-1+ic+mc} dx = \frac{x_0^{r+c+ic+mc}}{r + c + ic + mc}.$$

Hence,

$$\mu_r(x_0) = \theta^2 ck \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} (-1)^l \binom{k-1}{i} \binom{\theta+j}{j} \theta^j \binom{j}{l} \binom{lk}{m} \cdot \frac{x_0^{r+c+ic+mc}}{r + c + ic + mc}. \tag{11}$$

3.4 Moment Generating Function

Let $t \in \mathbb{R}^+$, the moment generating function (MGF) of any continuous random variable X given as $M_X(t)$ which assumes the TIHTBXII(θ, c, k) is defined as

$$M_X(t) = \mathbb{E}(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} g(x) \, dx = \int_0^{\infty} \frac{\theta^2 ck x^{c-1} (1 + x^c)^{k-1} e^{tx}}{\left[1 - \theta \left\{1 - (1 + x^c)^k\right\}\right]^{\theta+1}} \, dx$$

$$= \theta^2 ck \int_0^{\infty} x^{c-1} (1 + x^c)^{k-1} e^{tx} \left[1 - \theta \left\{1 - (1 + x^c)^k\right\}\right]^{-(\theta+1)} \, dx.$$

Using series expansion

$$(1 + x^c)^{k-1} = \sum_{i=0}^{\infty} \binom{k-1}{i} x^{ic}; \quad \text{for} \quad |x| < 1.$$

For $\theta \in \mathbb{R}$ or $\mathbb{C}$, a generalized binomial theorem provides that

$$\left[1 - \theta \left\{1 - (1 + x^c)^k\right\}\right]^{-(\theta+1)} = \sum_{j=0}^{\infty} \binom{\theta+j}{j} \theta^j \left[1 - (1 + x^c)^k\right]^j.$$

Also, for $j \in \mathbb{R}/\mathbb{N}_{\nleftarrow}$ and $|x| < 1$, standard binomial expansion provides that

$$\left[1 - (1 + x^c)^k\right]^j = \sum_{l=0}^{\infty} (-1)^l \binom{j}{l} (1 + x^c)^{lk}.$$

Similarly,

$$(1 + x^c)^{lk} = \sum_{m=0}^{\infty} \binom{lk}{m} x^{mc}; \quad e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

By back substitution,

$$M_X(t) = \theta^2 ck \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \frac{(-1)^l}{n!} \binom{k-1}{i} \binom{\theta+j}{j} \binom{j}{l} \binom{lk}{m} \theta^j \int_0^{\infty} x^{n+c+ic+mc-1} \, dx.$$

Notice that

$$\binom{j}{l} = \frac{j!}{l!(j-l)!} = \frac{(j,l)}{l!}; \quad \text{Recall} \quad e^{-x} = \sum_{l=0}^{\infty} (-1)^l \frac{x^l}{l!},$$

so that

$$M_X(t) = \theta^2 ck \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \binom{k-1}{i} \binom{\theta+j}{j} \binom{lk}{m} \frac{\theta^j (j,l)}{n!} \int_0^{\infty} \sum_{l=0}^{\infty} (-1)^l \frac{x^l}{l!} \frac{1}{x^l} x^{n+c+ic+mc-1} \, dx$$

$$= \theta^2 ck \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \binom{k-1}{i} \binom{\theta+j}{j} \binom{lk}{m} \frac{\theta^j (j,l)}{n!} \int_0^{\infty} x^{n+c+ic+mc-l-1} e^{-x} \, dx.$$

Therefore,

$$M_X(t) = \theta^2 ck \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \binom{k-1}{i} \binom{\theta+j}{j} \binom{lk}{m} \frac{\theta^j (j,l) \Gamma(n+c+ic+mc-l)}{n!}. \tag{12}$$

3.5 Mean Residual Life Function

The mean residual life (MRL) function of a non-negative continuous random variable X is defined as

$$m(t) = \mathbb{E}[X - t \mid X > t] = \frac{1}{\bar{G}(t)} \int_t^{\infty} \bar{G}(x) \, dx, \quad t \ge 0, \tag{13}$$

where $\bar{G}(t) = 1 - G(t)$ is the survival function. This function quantifies the expected remaining lifetime given survival up to time t. Suppose $X \sim \text{TIHTBXII}(\theta, c, k)$, the MRL is given as follows

$$m(t) = \left[1 - \theta + \theta (1 + t^c)^k\right]^\theta \int_t^{\infty} \left[1 - \theta + \theta (1 + x^c)^k\right]^{-\theta} \, dx.$$

Recall that for $j \in \mathbb{R}/\mathbb{N}_{\nmid}$ and $|x| < 1$, the standard binomial expansion provides that

$$(1 + x)^{-n} = \sum_{r=0}^{\infty} (-1)^r \binom{n+r-1}{r} x^r,$$

hence

$$m(t) = \left[1 - \theta + \theta (1 + t^c)^k\right]^\theta \int_t^{\infty} \sum_{r=0}^{\infty} (-1)^r \binom{\theta+r-1}{r} \theta^r \left[1 - (1 + x^c)^k\right]^r \, dx.$$

Similarly, by expanding $\left[1 - (1 + x^c)^k\right]^r$, we have

$$m(t) = \left[1 - \theta + \theta (1 + t^c)^k\right]^\theta \int_t^\infty \sum_{r=0}^\infty \sum_{s=0}^\infty (-1)^{r+s} \binom{\theta + r - 1}{r} \binom{r}{s} \theta^r (1 + x^c)^{ks} \, dx.$$

Also, by expanding $(1 + x^c)^{ks}$, we realize

$$m(t) = \left[1 - \theta + \theta (1 + t^c)^k\right]^\theta \int_t^\infty \sum_{r=0}^\infty \sum_{s=0}^\infty (-1)^{r+s} \binom{\theta + r - 1}{r} \binom{r}{s} \theta^r \sum_{y=0}^\infty \binom{ks}{y} x^{yc} \, dx$$

$$= \left[1 - \theta + \theta (1 + t^c)^k\right]^\theta \sum_{r=0}^\infty \sum_{s=0}^\infty \sum_{y=0}^\infty (-1)^{r+s} \binom{\theta + r - 1}{r} \binom{r}{s} \binom{ks}{y} \theta^r \int_t^\infty x^{yc} \, dx.$$

Recall

$$\binom{r}{s} = \frac{r!}{s!(r-s)!} = \frac{(r,s)}{s!}; \quad \text{and} \quad e^{-x} = \sum_{s=0}^\infty (-1)^s \frac{x^s}{s!}$$

$$m(t) = \left[1 - \theta + \theta (1 + t^c)^k\right]^\theta \int_t^\infty \sum_{r=0}^\infty \sum_{y=0}^\infty \sum_{s=0}^\infty \frac{(-1)^s x^s}{s!} \frac{1}{x^s} (-1)^r \binom{\theta + r - 1}{r} \binom{ks}{y} (r,s) \theta^r x^{yc} \, dx$$

$$= \left[1 - \theta + \theta (1 + t^c)^k\right]^\theta \sum_{r=0}^\infty \sum_{y=0}^\infty (-1)^r \binom{\theta + r - 1}{r} \binom{ks}{y} (r,s) \theta^r \int_t^\infty x^{yc-s} e^{-x} \, dx.$$

Therefore, the MRL function is

$$m(t) = \left[1 - \theta + \theta (1 + t^c)^k\right]^\theta \sum_{r=0}^\infty \sum_{y=0}^\infty (-1)^r \binom{\theta + r - 1}{r} \binom{ks}{y} (r,s) \theta^r \Gamma(yc - s + 1, t). \tag{14}$$

3.6 Entropy

Entropy is a fundamental concept in information theory and statistical mechanics, used to quantify the uncertainty or randomness associated with a probability distribution. For a given probability density function $g(x)$, the Rényi entropy of order $q > 0$, $q \neq 1$, is defined as:

$$H_q(x) = \frac{1}{1-q} \log \left\{ \int_{-\infty}^\infty g^q(x) \, dx \right\}.$$

In this section, we derive the Rényi entropy of the proposed distribution. Due to the complexity of the density function, the integral involved is analytically intractable in closed form. However, by applying the generalized binomial expansion and power series techniques, we express the entropy in terms of infinite series and known special functions, particularly the Gamma function.

$$H_q(x) = \frac{1}{1-q} \log \left\{ \int_{-\infty}^\infty g^q(x) \, dx \right\} = \frac{1}{1-q} \log \left\{ (\theta^2 ck)^q \int_0^\infty \frac{x^{q(c-1)} (1 + x^c)^{q(k-1)}}{\left[1 - \theta \left(1 - (1 + x^c)^k\right)\right]^{q(\theta+1)}} \, dx \right\}$$

$$= \frac{1}{1-q} \log \left\{ (\theta^2 ck)^q \int_0^\infty x^{q(c-1)} (1 + x^c)^{q(k-1)} \left[1 - \theta \left(1 - (1 + x^c)^k\right)\right]^{-q(\theta+1)} \, dx \right\}.$$

Using generalized binomial expansion when $q(\theta + 1) \in \mathbb{R}$ and $|x| < 1$, it can be shown that

$$\left[1 - \theta\left(1 - (1 + x^c)^k\right)\right]^{-q(\theta+1)} = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{l=0}^{\infty}(-1)^j\binom{q(\theta+1)+i-1}{i}\binom{i}{j}\binom{ik}{l}\theta^i x^{lc}.$$

Similarly,

$$(1 + x^c)^{q(k-1)} = \sum_{h=0}^{\infty}\binom{q(k-1)}{h}x^{hc}.$$

Therefore,

$$H_q(x) = \frac{1}{1-q}\log\left\{(\theta^2 ck)^q \int_0^{\infty} x^{q(c-1)}\sum_{h=0}^{\infty}\binom{q(k-1)}{h}x^{hc}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{l=0}^{\infty}(-1)^j\binom{q(\theta+1)+i-1}{i}\binom{i}{j}\binom{ik}{l}\theta^i x^{lc}\right\}$$

$$= \frac{1}{1-q}\log\left\{(\theta^2 ck)^q \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{l=0}^{\infty}\sum_{h=0}^{\infty}(-1)^j\binom{q(\theta+1)+i-1}{i}\binom{i}{j}\binom{ik}{l}\binom{q(k-1)}{h}\theta^i\int_0^{\infty}x^{qc-q+hc+lc}\,dx\right\}.$$

Recall that

$$\binom{i}{j} = \frac{i!}{j!(i-j)!} = \frac{P(i,j)}{j!}; \quad \text{and} \quad e^{-x} = \sum_{j=0}^{\infty}(-1)^j\frac{x^j}{j!},$$

$$H_q(x) = \frac{1}{1-q}\log\left\{(\theta^2 ck)^q \sum_{i=0}^{\infty}\sum_{l=0}^{\infty}\sum_{h=0}^{\infty}\sum_{j=0}^{\infty}(-1)^j\int_0^{\infty}\frac{x^j}{j!}\frac{1}{x^j}(i,j)\theta^i\binom{q(\theta+1)+i-1}{i}\binom{ik}{l}\binom{q(k-1)}{h}x^{qc-q+hc+lc}\,dx\right\}$$

$$= \frac{1}{1-q}\log\left\{(\theta^2 ck)^q \sum_{i=0}^{\infty}\sum_{l=0}^{\infty}\sum_{h=0}^{\infty}P(i,j)\theta^i\binom{q(\theta+1)+i-1}{i}\binom{ik}{l}\binom{q(k-1)}{h}\int_0^{\infty}x^{qc-q+hc+lc-j}e^{-x}\,dx\right\}$$

$$= \frac{1}{1-q}\log\left\{(\theta^2 ck)^q \sum_{i=0}^{\infty}\sum_{l=0}^{\infty}\sum_{h=0}^{\infty}P(i,j)\theta^i\binom{q(\theta+1)+i-1}{i}\binom{ik}{l}\binom{q(k-1)}{h}\Gamma(qc-q+hc+lc-j+1)\right\}.$$

3.7 Extropy

This subsection defines extropy, a measure of order or information content within a system, serving as a dual to the more commonly known concept of entropy. We consider a specific mathematical formulation of extropy, denoted by the functional $J(g)$. The initial representation of $J(g)$ is an infinite series where each term encompasses an integral over a complex function $g(x)$. To analytically evaluate this integral, we systematically apply the generalized binomial expansion to several nested terms, including $\left[1 - \theta\left\{1 - (1 + x^c)^k\right\}\right]^{-n(\theta+1)}$, $\left[1 - (1 + x^c)^k\right]^i$, and $(1 + x^c)^{jk}$. This decomposition transforms the original integral into a convergent multifold summation. The final integration step is then performed, yielding an expression for $J(g)$ in terms of the Gamma function, $\Gamma(z)$, thus providing a complete analytical series expansion.

$$J(g) = \sum_{n=2}^{\infty}\frac{1}{n(n-2)}\int_0^{\infty}g^n(x)\,dx = \sum_{n=2}^{\infty}\frac{1}{n(n-2)}\int_0^{\infty}\frac{(\theta^2 ck)^n x^{n(c-1)}(1+x^c)^{n(k-1)}}{\left[1 - \theta\left\{1 - (1 + x^c)^k\right\}\right]^{n(\theta+1)}}\,dx$$

$$= \sum_{n=2}^{\infty}\frac{(\theta^2 ck)^n}{n(n-2)}\int_0^{\infty}x^{n(c-1)}(1+x^c)^{n(k-1)}\left[1 - \theta\left\{1 - (1 + x^c)^k\right\}\right]^{-n(\theta+1)}\,dx.$$

Using generalized binomial expansion; for $n(\theta + 1) \in \mathbb{R}$ and $|x| < 1$,

$$\left[1 - \theta \left\{1 - (1 + x^c)^k\right\}\right]^{-n(\theta+1)} = \sum_{i=0}^{\infty} \binom{n(\theta+1) + i - 1}{i} \theta^i \left[1 - (1 + x^c)^k\right]^i.$$

Also,

$$\left[1 - (1 + x^c)^k\right]^i = \sum_{j=0}^{\infty} (-1)^j \binom{i}{j} (1 + x^c)^{jk}.$$

Similarly,

$$(1 + x^c)^{jk} = \sum_{h=0}^{\infty} \binom{jk}{h} x^{hc}; \quad \text{and} \quad (1 + x^c)^{n(k-1)} = \sum_{l=0}^{\infty} \binom{n(k-1)}{l} x^{lc}.$$

So that,

$$J(g) = \sum_{n=2}^{\infty} \frac{(\theta^2 ck)^2}{n(n-1)} \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{h=0}^{\infty} \sum_{l=0}^{\infty} (-1)^j \binom{i}{j} \binom{n(k-1)}{l} \binom{jk}{h} \binom{n(\theta) + i - 1}{i} \theta^i \int_0^{\infty} x^{nc-n+lc+hc} \, dx.$$

Since,

$$\binom{i}{j} = \frac{i!}{j!(i-j)!} = \frac{P(i, j)}{j!}; \quad e^{-x} = \sum_{j=0}^{\infty} (-1)^j \frac{x^j}{j!},$$

$$J(g) = \sum_{n=2}^{\infty} \frac{(\theta^2 ck)^n}{n(n-1)} \sum_{i=0}^{\infty} \sum_{l=0}^{\infty} \sum_{h=0}^{\infty} \binom{n(k-1)}{l} \binom{jk}{h} \binom{n(\theta+1) + i - 1}{i} \int_0^{\infty} \sum_{j=0}^{\infty} (-1)^j \frac{x^j}{j!} \frac{1}{x^j} P(i, j) \theta^i x^{nc-n+lc+hc} \, dx$$

$$= \sum_{n=2}^{\infty} \frac{(\theta^2 ck)^n}{n(n-1)} \sum_{i=0}^{\infty} \sum_{l=0}^{\infty} \sum_{h=0}^{\infty} \binom{n(k-1)}{l} \binom{jk}{h} \binom{n(\theta+1) + i - 1}{i} \theta^i P(i, j) \int_0^{\infty} x^{nc-n+lc+hc-j} e^{-x} \, dx$$

$$= \sum_{n=2}^{\infty} \frac{(\theta^2 ck)^n}{n(n-1)} \sum_{i=0}^{\infty} \sum_{l=0}^{\infty} \sum_{h=0}^{\infty} \binom{n(k-1)}{l} \binom{jk}{h} \binom{n(\theta+1) + i - 1}{i} \theta^i P(i, j) \Gamma(nc - n + lc + hc - j + 1).$$

3.8 Order Statistic

In the realm of probability theory and statistics, an order statistic is a fundamental concept representing the r-th smallest value among a set of n random variables. Given a set of n independent and identically distributed (i.i.d.) random variables, say $X_1, X_2, \ldots, X_n$, arranged in ascending order as $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$, $X_{(r)}$ is the r-th order statistic. The PDF of the r-th order statistic, $f_{X_{r:n}}(x)$, is derived from the PDF of the individual random variables, $g(x)$, and their CDF, $G(x)$. For the proposed TIHTBXII model, it is given as

$$f_{X_{r:n}}(x) = \frac{n!}{(r-1)!(n-r)!} \left[G(x)\right]^{r-1} \left[1 - G(x)\right]^{n-r} g(x)$$

$$= \frac{n!}{(r-1)!(n-r)!} \left\{1 - \left[1 - \theta \left(1 - (1 + x^c)^k\right)\right]^{-\theta}\right\}^{r-1} \left\{1 - \theta \left(1 - (1 + x^c)^k\right)\right\}^{-\theta(n-r)}$$

$$\times \frac{\theta^2 ck x^{c-1} (1 + x^c)^{k-1}}{\left[1 - \theta \left\{1 - (1 + x^c)^k\right\}\right]^{\theta+1}}$$

$$= \frac{n!\theta^2 ckx^{c-1}}{(r-1)!(n-r)!} \sum_{i,j=0}^{\infty} \sum_{l,h=0}^{\infty} \sum_{m,p=0}^{\infty} \sum_{q,s=0}^{\infty} \sum_{t,w=0}^{\infty} (-1)^{i+p+m+s} \theta^{h+j+l} \binom{r-1}{i} \binom{i\theta+h-1}{h} \binom{h}{p} \binom{pk}{q}$$

$$\times \binom{\theta(n-r)+j-1}{j} \binom{j}{m} \binom{km}{w} \binom{\theta+l}{l} \binom{l}{s} \binom{ks}{t} x^{qc+wc+tc}.$$

4 Estimation

We proceed with the estimation of the parameters of the TIHTBXII distribution using maximum likelihood, maximum product spacing, least squares and weighted least squares procedures.

4.1 Maximum Likelihood Estimation

Maximum Likelihood Estimation (MLE) involves finding the values of the parameters that maximize the likelihood of observing the given data. First, we formulate the likelihood function. Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from the distribution with the given PDF $g(x; \theta, c, k)$. Assuming the observations are independent and identically distributed (i.i.d.), the likelihood function $L(\theta, c, k|\mathbf{x})$ is the product of the individual PDFs:

$$L(\theta, c, k|\mathbf{x}) = \prod_{i=1}^{n} g(x_i; \theta, c, k) = \prod_{i=1}^{n} \frac{\theta^2 ck x_i^{c-1} (1+x_i^c)^{k-1}}{\left[1 - \theta\left\{1 - (1+x_i^c)^k\right\}\right]^{\theta+1}}.$$

Formulate the log-likelihood function since it is almost always easier to work with the natural logarithm of the likelihood function, called the log-likelihood function, denoted as $l(\theta, c, k|\mathbf{x})$ or $\ln L$. This is because the logarithm is a monotonically increasing function, so maximizing $\ln L$ is equivalent to maximizing L. Also, products become sums, which are much simpler to differentiate. So, the log-likelihood function is:

$$l(\theta, c, k|\mathbf{x}) = n(2\ln\theta + \ln c + \ln k) + (c-1)\sum_{i=1}^{n}\ln x_i + (k-1)\sum_{i=1}^{n}\ln(1+x_i^c)$$

$$- (\theta+1)\sum_{i=1}^{n}\ln\left[1 - \theta\left\{1 - (1+x_i^c)^k\right\}\right],$$

$$\frac{2n}{\theta} - \sum_{i=1}^{n}\ln\left[1 - \theta\left\{1 - (1+x_i^c)^k\right\}\right] + (\theta+1)\sum_{i=1}^{n}\frac{1-(1+x_i^c)^k}{\left[1 - \theta\left\{1 - (1+x_i^c)^k\right\}\right]} = 0 \qquad (15)$$

$$\frac{\partial l}{\partial c} = \frac{n}{c} + \sum_{i=1}^{n}\ln x_i + (k-1)\sum_{i=1}^{n}\frac{x_i^c \ln x_i}{1+x_i^c} - (\theta+1)\sum_{i=1}^{n}\frac{\theta k(1+x_i^c)^{k-1} x_i^c \ln x_i}{\left[1 - \theta\left\{1 - (1+x_i^c)^k\right\}\right]} = 0,$$

and

$$\frac{\partial l}{\partial k} = \frac{n}{k} + \sum_{i=1}^{n}\ln(1+x_i^c) - (\theta+1)\sum_{i=1}^{n}\frac{\theta(1+x_i^c)^k \ln(1+x_i^c)}{\left[1 - \theta\left\{1 - (1+x_i^c)^k\right\}\right]} = 0. \qquad (16)$$

Next, solve the system of score equations. The MLEs for θ, c, and k (denoted as $\hat{\theta}$, $\hat{c}$, and $\hat{k}$) are the solutions to the system of these three non-linear equations:

$$\frac{2n}{\hat{\theta}} - \sum_{i=1}^{n} \ln\left[1 - \hat{\theta}\left\{1 - \left(1 + x_i^{\hat{c}}\right)^{\hat{k}}\right\}\right] + (\hat{\theta} + 1)\sum_{i=1}^{n} \frac{1 - \left(1 + x_i^{\hat{c}}\right)^{\hat{k}}}{\left[1 - \hat{\theta}\left\{1 - \left(1 + x_i^{\hat{c}}\right)^{\hat{k}}\right\}\right]} = 0,$$

$$\frac{n}{\hat{c}} + \sum_{i=1}^{n} \ln x_i + (\hat{k} - 1)\sum_{i=1}^{n} \frac{x_i^{\hat{c}} \ln x_i}{1 + x_i^{\hat{c}}} - (\hat{\theta} + 1)\sum_{i=1}^{n} \frac{\hat{\theta}\hat{k}(1 + x_i^{\hat{c}})^{\hat{k}-1} x_i^{\hat{c}} \ln x_i}{\left[1 - \hat{\theta}\left\{1 - \left(1 + x_i^{\hat{c}}\right)^{\hat{k}}\right\}\right]} = 0,$$

$$\frac{n}{\hat{k}} + \sum_{i=1}^{n} \ln(1 + x_i^{\hat{c}}) - (\hat{\theta} + 1)\sum_{i=1}^{n} \frac{\hat{\theta}(1 + x_i^{\hat{c}})^{\hat{k}} \ln(1 + x_i^{\hat{c}})}{\left[1 - \hat{\theta}\left\{1 - \left(1 + x_i^{\hat{c}}\right)^{\hat{k}}\right\}\right]} = 0.$$

The solutions can be obtained numerically.

4.2 Maximum Product Spacing Estimation

The Maximum Product Spacing (MPS) method estimates parameters by maximizing the product of the differences between consecutive values of the CDF, evaluated at ordered observations. Let $X_1, X_2, \ldots, X_n$ be a random sample from the distribution, and let $X_{(1)} \le X_{(2)} \le \ldots \le X_{(n)}$ be the ordered statistics. Define the uniform spacings $D_i(\theta, c, k)$ as:

$$D_i(\theta, c, k) = G(X_{(i)}; \theta, c, k) - G(X_{(i-1)}; \theta, c, k),$$

for $i = 1, \ldots, n + 1$, where we set $G(X_{(0)}; \theta, c, k) = 0$ and $G(X_{(n+1)}; \theta, c, k) = 1$. The product spacing function to be maximized is:

$$P(\theta, c, k) = \prod_{i=1}^{n+1} D_i(\theta, c, k)$$

Alternatively, it is often more convenient to maximize the logarithm of the product spacing function:

$$S(\theta, c, k) = \sum_{i=1}^{n+1} \ln D_i(\theta, c, k).$$

The MPS estimates $(\hat{\theta}_{MPS}, \hat{c}_{MPS}, \hat{k}_{MPS})$ are obtained by solving the following system of equations: $\frac{\partial S}{\partial \theta} = 0, \frac{\partial S}{\partial c} = 0, \frac{\partial S}{\partial k} = 0$. These equations are solved using numerical optimization techniques.

4.3 Least Squares Estimation

The Least Squares (LS) method estimates parameters by minimizing the sum of squared differences between the empirical CDF and the theoretical CDF. Let $X_{(1)} \le X_{(2)} \le \ldots \le X_{(n)}$ be the ordered statistics of the observed sample. The empirical CDF at $X_{(i)}$ is approximated by $F_n(X_{(i)}) = \frac{i}{n+1}$. The LS estimates $(\hat{\theta}_{LS}, \hat{c}_{LS}, \hat{k}_{LS})$ are found by minimizing the sum of squares function:

$$L(\theta, c, k) = \sum_{i=1}^{n} \left[G(X_{(i)}; \theta, c, k) - \frac{i}{n+1}\right]^2.$$

To find the estimates, we set the partial derivatives of $L(\theta, c, k)$ with respect to each parameter to zero:

$$\frac{\partial L}{\partial \theta} = 2\sum_{i=1}^{n}\left[G(X_{(i)}; \theta, c, k) - \frac{i}{n+1}\right]\frac{\partial G(X_{(i)}; \theta, c, k)}{\partial \theta} = 0,$$

$$\frac{\partial L}{\partial c} = 2 \sum_{i=1}^{n} \left[G(X_{(i)}; \theta, c, k) - \frac{i}{n+1} \right] \frac{\partial G(X_{(i)}; \theta, c, k)}{\partial c} = 0,$$

$$\frac{\partial L}{\partial k} = 2 \sum_{i=1}^{n} \left[G(X_{(i)}; \theta, c, k) - \frac{i}{n+1} \right] \frac{\partial G(X_{(i)}; \theta, c, k)}{\partial k} = 0. \tag{17}$$

These equations are typically solved numerically.

4.4 Weighted Least Squares Estimation

Weighted Least Squares (WLS) is an extension of LS that incorporates weights to account for varying precision across observations. In the context of empirical CDF fitting, weights are often chosen to reflect the variance of the ordered statistics. The WLS estimates $(\hat{\theta}_{WLS}, \hat{c}_{WLS}, \hat{k}_{WLS})$ are obtained by minimizing the weighted sum of squares function:

$$W(\theta, c, k) = \sum_{i=1}^{n} w_i \left[G(X_{(i)}; \theta, c, k) - \frac{i}{n+1} \right]^2.$$

A common choice for weights w_i based on the variance of uniform order statistics is:

$$w_i = \frac{(n+1)^2 (n+2)}{i(n-i+1)}.$$

The system of equations to solve for the WLS estimates is:

$$\frac{\partial W}{\partial \theta} = 2 \sum_{i=1}^{n} w_i \left[G(X_{(i)}; \theta, c, k) - \frac{i}{n+1} \right] \frac{\partial G(X_{(i)}; \theta, c, k)}{\partial \theta} = 0,$$

$$\frac{\partial W}{\partial c} = 2 \sum_{i=1}^{n} w_i \left[G(X_{(i)}; \theta, c, k) - \frac{i}{n+1} \right] \frac{\partial G(X_{(i)}; \theta, c, k)}{\partial c} = 0,$$

$$\frac{\partial W}{\partial k} = 2 \sum_{i=1}^{n} w_i \left[G(X_{(i)}; \theta, c, k) - \frac{i}{n+1} \right] \frac{\partial G(X_{(i)}; \theta, c, k)}{\partial k} = 0. \tag{18}$$

Numerical optimization is required to solve this system. To implement the LS and WLS methods, the partial derivatives of $G(x; \theta, c, k)$ with respect to θ, c, and k are required, which are given as follows

$$\frac{\partial G}{\partial \theta} = -(1 - \theta\{1 - (1 + x^c)^k\})^{-\theta} \left[-\ln(1 - \theta\{1 - (1 + x^c)^k\}) + \frac{\theta\{1 - (1 + x^c)^k\}}{1 - \theta\{1 - (1 + x^c)^k\}} \right],$$

$$\frac{\partial G}{\partial c} = \theta^2 k (1 + x^c)^{k-1} x^c \ln x [1 - \theta\{1 - (1 + x^c)^k\}]^{-\theta-1},$$

$$\frac{\partial G}{\partial k} = \theta^2 (1 + x^c)^k \ln(1 + x^c)[1 - \theta\{1 - (1 + x^c)^k\}]^{-\theta-1}. \tag{19}$$

5 Simulation

To contrast the performance of the estimation methods (MLE, MPS, LS, and WLS), we conduct a Monte Carlo Markov Chain (MCMC) simulation study. In the simulation, we test the bias and root mean square error (RMSE) of the parameter estimates for various sample sizes and true parameter values. MCMC simulation is a powerful and widely used approach for evaluating the efficiency of estimators, particularly for advanced statistical models in which analytical solutions are impossible, based on early work by Metropolis et al. [46] and then subsequently broadened by Hastings [47].

For our computation, we performed 1000 Monte Carlo repetitions for each of the scenarios and sample sizes. In each of these, we generated data from the TIHTBXII distribution, via its quantile function, from uniformly distributed random numbers. Numerical optimization (with in particular `optim` in R) was used to estimate MLE, LS, and WLS parameters, while the MPS estimates were calculated using the function `mpsedist` from the package `BMT`.

Our simulation experiment is done on the basis of four distinct scenarios, which are equivalent to different combinations of the distribution parameters θ, c, and k:

Scenario I $\theta = 0.5$, $c = 1.75$ and $k = 2.75$
Scenario II $\theta = 0.75$, $c = 2.75$ and $k = 3.5$
Scenario III $\theta = 1.5$, $c = 1.70$ and $k = 2.5$
Scenario IV $\theta = 0.75$, $c = 2.5$ and $k = 2.75$

For each situation, we generate random samples of varying sizes ($n = 25, 100, 200, 500$). We then go ahead and apply each estimation method on these test datasets, compute estimates of the parameters, and compute their Bias and RMSE.

The simulation results, which are reported in Tables 1 and 2, provide information about the performance of the MLE, MPS, LS, and WLS estimators of the parameters of the TIHTBXII distribution for various sample sizes and settings. The overall pattern in all the methods and settings is that Bias and Root Mean Square Error (RMSE) decrease as the sample size increases. This demonstrates that all estimators are more precise and accurate with larger data. Estimation of $\hat{\theta}$ and $\hat{c}$ tends to have lower bias and RMSE than $\hat{k}$, demonstrating that $\hat{k}$ is usually more difficult to estimate. Throughout the methods, WLS tends to perform better with relatively lower bias and RMSE for $\hat{\theta}$ and $\hat{c}$ in most scenarios, where the benefits of WLS are more pronounced at larger sample sizes. LS also performs well, with a good balance between accuracy and precision. While MLE improves significantly with large sample sizes, with small sample sizes its performance can sometimes be less stable compared to WLS and LS. MPS performs similarly but sometimes has larger RMSE, particularly for $\hat{k}$ with small sample sizes.

Table 1: MCMC simulation results for scenarios I & II

Method	Estimator	n = 25		n = 100		n = 200		n = 500	
		Bias	**RMSE**	**Bias**	**RMSE**	**Bias**	**RMSE**	**Bias**	**RMSE**
					Scenario I				
MLE	$\hat{\theta}$	0.17995	0.58668	0.09285	0.40835	0.07448	0.30844	0.02802	0.20063
	$\hat{c}$	0.20444	0.51330	0.06712	0.28662	0.05960	0.24033	0.01602	0.17618
	$\hat{k}$	1.85327	9.89314	0.71210	3.71272	0.32691	1.97224	0.29046	1.46998
MPS	$\hat{\theta}$	0.05554	0.69423	0.01180	0.48287	0.00620	0.38081	0.02208	0.19508
	$\hat{c}$	0.13268	0.41376	0.08127	0.27046	0.04396	0.23190	0.04449	0.17507
	$\hat{k}$	2.54496	6.31906	1.38572	3.04250	0.97100	2.31894	0.67102	1.61581
LS	$\hat{\theta}$	0.01926	0.48189	0.03204	0.25960	0.03399	0.21379	0.03358	0.16623
	$\hat{c}$	0.02307	0.41589	0.04482	0.23189	0.03236	0.18905	0.03852	0.15853
	$\hat{k}$	2.05443	4.60830	1.36641	3.24356	1.04080	2.54635	0.70745	1.71386

(Continued)

Table 1 (continued)

Method	Estimator	n = 25		n = 100		n = 200		n = 500	
		Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
WLS	$\hat{\theta}$	0.05814	0.56244	0.02143	0.34217	0.01430	0.25489	0.00387	0.18565
	$\hat{c}$	0.00552	0.43013	0.00941	0.25968	0.00539	0.20767	0.00746	0.16196
	$\hat{k}$	1.94456	4.69110	1.01071	2.90617	0.67189	2.19200	0.42847	1.51416
				Scenario II					
MLE	$\hat{\theta}$	0.31921	0.80005	0.24869	0.77164	0.17024	0.66364	0.06082	0.36400
	$\hat{c}$	0.24985	0.64586	0.08335	0.38315	0.03061	0.31631	0.00513	0.23990
	$\hat{k}$	0.49244	4.14513	0.34461	2.78228	0.40032	2.53688	0.36263	2.00259
MPS	$\hat{\theta}$	0.20099	0.93397	0.21607	0.97536	0.14415	0.88775	0.02810	0.50396
	$\hat{c}$	0.18479	0.60298	0.11748	0.41160	0.11485	0.36003	0.07419	0.27140
	$\hat{k}$	2.13792	6.74145	1.48787	3.87291	1.36216	3.38651	0.88632	2.45378
LS	$\hat{\theta}$	0.38204	0.95043	0.34648	0.77473	0.24021	0.60351	0.14089	0.42083
	$\hat{c}$	0.01194	0.62021	0.05696	0.38574	0.04270	0.31201	0.02837	0.25000
	$\hat{k}$	0.44317	3.70687	0.20983	2.28670	0.15856	2.00869	0.07759	1.70025
WLS	$\hat{\theta}$	0.41882	1.03326	0.32804	0.80379	0.20356	0.59471	0.09772	0.36325
	$\hat{c}$	0.03128	0.61731	0.05434	0.38371	0.03046	0.31320	0.01816	0.23707
	$\hat{k}$	0.51868	3.57274	0.00275	2.40681	0.06440	2.15360	0.06715	1.68018

Table 2: MCMC simulation results for scenarios III & IV

Method	Estimator	n = 25		n = 100		n = 200		n = 500	
		Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
				Scenario III					
MLE	$\hat{\theta}$	0.03225	1.01405	0.14890	1.10252	0.20828	1.11438	0.19798	0.98548
	$\hat{c}$	0.01821	0.31046	0.04359	0.18766	0.03864	0.15207	0.02860	0.11743
	$\hat{k}$	2.43943	5.18841	1.60911	4.08507	1.07862	3.24293	0.62665	2.51134
MPS	$\hat{\theta}$	0.27658	1.38419	0.57008	1.65001	0.64230	1.67052	0.50956	1.41165
	$\hat{c}$	0.17766	0.34573	0.11401	0.22601	0.08193	0.18301	0.04807	0.13857
	$\hat{k}$	2.34683	5.35857	1.63571	4.78528	1.11429	4.13791	0.57606	3.12439
LS	$\hat{\theta}$	0.25807	1.10924	0.16082	0.94809	0.16117	0.86683	0.14925	0.74437
	$\hat{c}$	0.08878	0.36169	0.08173	0.21259	0.06665	0.17392	0.04435	0.13095
	$\hat{k}$	1.45022	3.92998	1.30302	3.51828	1.01555	3.06586	0.66314	2.52461

(Continued)

Table 2 (continued)

Method	Estimator	n = 25		n = 100		n = 200		n = 500	
		Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
WLS	$\hat{\theta}$	0.30659	1.17572	0.21222	1.02355	0.18744	0.91604	0.16649	0.75752
	$\hat{c}$	0.07909	0.34304	0.07364	0.20768	0.05792	0.16782	0.03605	0.12387
	$\hat{k}$	1.57857	4.14790	1.37480	3.72453	1.03745	3.17618	0.58840	2.48032
Scenario IV									
MLE	$\hat{\theta}$	0.31846	0.79170	0.23122	0.76457	0.16800	0.62207	0.05269	0.31189
	$\hat{c}$	0.23073	0.57615	0.06602	0.34054	0.04722	0.29100	0.00863	0.21561
	$\hat{k}$	0.18949	2.41963	0.36323	2.33271	0.25611	1.97845	0.25212	1.53135
MPS	$\hat{\theta}$	0.28752	1.01289	0.27740	1.08300	0.18879	0.87512	0.03631	0.49300
	$\hat{c}$	0.15038	0.53142	0.09967	0.37901	0.06121	0.32498	0.05609	0.24667
	$\hat{k}$	1.33882	4.50712	1.12505	3.03613	0.85302	2.63671	0.63895	1.95149
LS	$\hat{\theta}$	0.43714	0.98028	0.30746	0.73448	0.25368	0.62240	0.09422	0.37176
	$\hat{c}$	0.01739	0.55000	0.04261	0.35575	0.04718	0.29523	0.00488	0.22805
	$\hat{k}$	0.19195	2.70779	0.06645	1.79729	0.09700	1.67350	0.09500	1.35778
WLS	$\hat{\theta}$	0.44203	1.04554	0.28139	0.74862	0.20635	0.56581	0.06411	0.31850
	$\hat{c}$	0.02642	0.54946	0.03950	0.35342	0.04639	0.28810	0.00275	0.21745
	$\hat{k}$	0.36631	2.73767	0.11744	1.94959	0.01094	1.70641	0.16137	1.35045

Scenario-dependent results show that the parameter values chosen influence estimation challenge. For instance, Scenario I is "easier" to estimate with lower overall RMSE values, whereas Scenarios II and IV are challenging, particularly for $\hat{\theta}$. Basically, the simulation confirms that the quality of estimation improves with larger data. WLS and LS prove to be robust and stable methods for estimating the parameters of the TIHTBXII distribution under the conditions considered.

Fig. 2a and b is the plots of the Bias and RMSE of the four non-Bayesian methods for Scenarios I and II while Fig. 3a and b is the plots of the Bias and RMSE of the four non-Bayesian methods for Scenarios III and IV. These support the numerical results in Tables 1 and 2 since as the sample size increases, the Bias remain positive and the RMSE decrease. These are indicators of good convergence behavior which indicate that the TIHTBXII model is fit for modeling lifetime datasets.
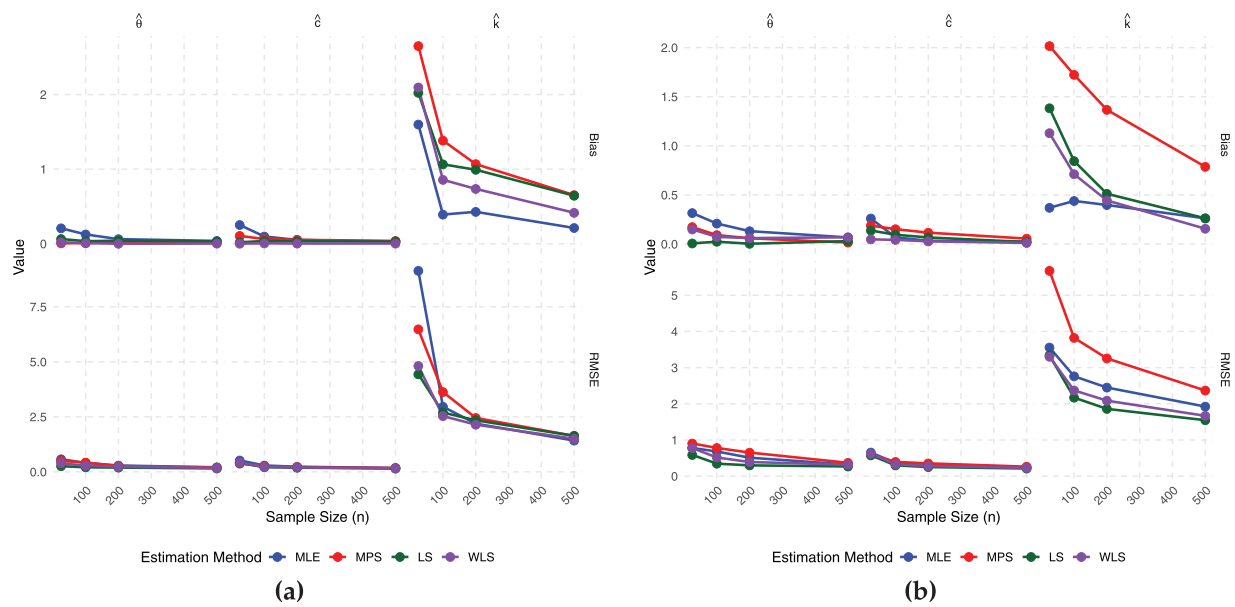
Figure 2: Estimator performance **(a)** scenario I **(b)** scenario II
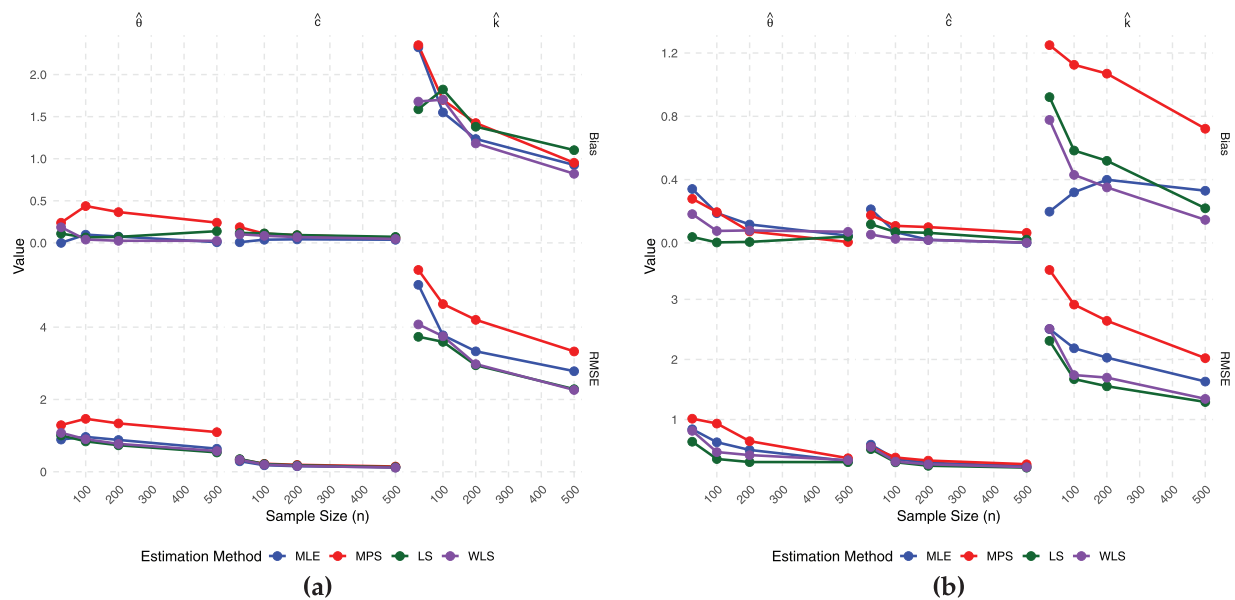

Figure 3: Estimator performance **(a)** scenario III **(b)** scenario IV

6 GASP Based on Truncated Life Tests for TIHTBXII Distribution

This section presents the use of Group Acceptance Sampling Plan (GASP) in truncated life tests for items having lifetimes that are TIHTBXII-distributed. Truncated life tests are very important for effective quality control, particularly in the case of testing long-lifetime components. The median lifetime of an item with TIHTBXII-distributed lifetime is obtained using its quantile function, given in Eq. (9). For the median ($u = 0.5$), the quantile function provides the median lifetime, μ:

$$\mu = \left\{ \left[1 - \left(\frac{1 - (1 - u)^{-\frac{1}{\theta}}}{\theta} \right) \right]^{\frac{1}{k}} - 1 \right\}^{\frac{1}{c}}.$$

(20)

For ease of calculations under the GASP platform, we split μ with the introduction of an intermediate term, $\bar{\omega}$. At the median ($u = 0.5$), $\bar{\omega}$ is:

$$\bar{\omega} = \left[1 - \left(\frac{1 - (1 - u)^{-\frac{1}{\theta}}}{\theta} \right) \right]^{\frac{1}{k}} - 1, \quad \text{at } u = 0.5 \quad \bar{\omega} = \left[1 - \left(\frac{1 - 2^{\frac{1}{\theta}}}{\theta} \right) \right]^{\frac{1}{k}} - 1.$$

This split is such that $\mu = \bar{\omega}^{\frac{1}{c}}$.

For the truncated life tests, the test duration is provided by $t_0 = a_1 \times \mu_0$, where a_1 is a specified constant and μ_0 is the specified median lifetime. The ratio of the actual median lifetime to the specified median lifetime is $r_2 = \frac{\mu}{\mu_0}$.

The failure probability, p, of a single item before time t_0 is derived by inserting these relationships into the CDF of the TIHTBXII distribution in Eq. (3). This yields:

$$p = 1 - \left[\frac{1}{1 - \theta \left\{ 1 - \left(1 + \left(\left(\frac{a_1}{r_2} \right)^c \times \bar{\omega} \right) \right)^k \right\}} \right]^{\theta}.$$

(21)

This expression for p plays a crucial role in calculating the acceptance probabilities in the GASP scheme.

The GASP design parameters, i.e., the number of groups (G), the acceptance number (n), the group size (r), and the test duration (t_0), are determined such that both consumer and producer risks are balanced. For determining the optimal GASP design parameters, we formulate an optimization problem. We want to minimize the Average Sample Number (ASN), which is $n = r \times G$. This minimization is subject to the following, which balance consumer risk (β) and producer risk (α):

$$P_{\text{accept}}\left(p_1 \mid \frac{\mu}{\mu_0} = r_1 \right) = \left[\sum_{i=0}^{n} \binom{r}{i} p_1^i (1 - p_1)^{r-i} \right]^g \le \beta,$$

(22)

$$P_{\text{accept}}\left(p_2 \mid \frac{\mu}{\mu_0} = r_2 \right) = \left[\sum_{i=0}^{n} \binom{r}{i} p_2^i (1 - p_2)^{r-i} \right]^g \ge 1 - \alpha.$$

(23)

Here, r_1 and r_2 are the mean ratios relating to the consumer and producer risks, respectively. The failure probabilities p_1 and p_2 are calculated in these P_{accept} functions, as provided in Eqs. (22) and (23).

Table 3 presents an illustrative set of GASP design parameters for the TIHTBXII distribution, based on specific parameter values ($\theta = 0.75$, $c = 2.25$, $k = 1.75$). The table shows the minimum required number of groups (G) and the acceptance number (n) for varying consumer risks (β), mean ratios ($\frac{\mu}{\mu_0}$), and test duration constants (a_1), for fixed group sizes ($r = 5$ and $r = 10$). The $P_{\text{accept}}(p)$ values indicate the probability of accepting a lot under these conditions.

Table 3: GASP under TIHTBXII at $\theta = 0.75$, $c = 2.25$, $k = 1.75$ and with minimum G and n

β	$\frac{\mu}{\mu_0}$	$r = 5$									$r = 10$								
		$a_1 = 0.5$				$a_1 = 1$					$a_1 = 0.5$				$a_1 = 1$		
		G	n	$P_{\text{accept}}(p)$		G	n	$P_{\text{accept}}(p)$			G	n	$P_{\text{accept}}(p)$		G	n	$P_{\text{accept}}(p)$
0.25	2	44	2	0.978689		7	3	0.9805			7	2	0.964823		2	4	0.974626
	4	7	1	0.995574		1	1	0.987112			3	1	0.991697		1	2	0.994897
	6	2	0	0.968123		1	1	0.997748			1	0	0.968123		1	1	0.990365
	8	2	0	0.983177		1	0	0.960522			1	0	0.983177		1	1	0.997225
0.10	2	73	2	0.964892		12	3	0.966805			37	3	0.987557		3	4	0.962181
	4	12	1	0.992425		2	1	0.974389			4	1	0.988944		1	2	0.994897
	6	3	0	0.952567		2	1	0.9955			4	1	0.99815		1	1	0.990365
	8	3	0	0.974872		1	0	0.960522			2	0	0.966638		1	1	0.997225
0.05	2	94	2	0.955023		15	3	0.95868			48	3	0.983888		7	5	0.986683
	4	15	1	0.99054		2	1	0.974389			5	1	0.9862		2	2	0.98982
	6	15	1	0.998442		2	1	0.9955			5	1	0.997689		1	1	0.990365
	8	4	0	0.966638		1	0	0.960522			2	0	0.966638		1	1	0.997225
0.01	2	0	0	–		146	4	0.985135			74	3	0.97527		10	5	0.98103
	4	23	1	0.985531		3	1	0.961831			7	1	0.980733		2	2	0.98982
	6	23	1	0.997612		3	1	0.993258			7	1	0.996766		2	1	0.980823
	8	6	0	0.950376		3	1	0.998101			3	0	0.950376		2	1	0.994457

The factors influencing the optimal plan are:

(a) β: Risk of consumer, measuring the largest acceptable probability of accepting a lot with actual defect level p.

(b) $r_2 = \mu/\mu_0$: True mean lifetime to specified mean lifetime ratio (μ to μ_0), a measure of lot quality.

(c) r: Number of inspection repetition or rounds.

(d) a_1: A scale factor for the truncation time, such that the test length $t_0 = a_1 \times \mu_0$. This parameter impacts the probability of failure p in cases of truncated testing.

The triplet given $(G, n, P_{\text{accept}}(p))$ is the minimal number of groups to be inspected, the group size, and the respective probability of accepting a lot.

(i) Mean Ratio Influence (r_2): With β, r, a_1 held constant, increasing r_2 (i.e., improved quality lots) causes a monotonic decrease in G and n. For example, when $\beta = 0.25$, $r = 5$, $a_1 = 0.5$, increasing r_2 from 2 to 8 decreases G from 44 to 2 and n from 2 to 0, demonstrating fewer inspection resources are needed for better lot quality. Acceptance probability, P_{accept}, is always high, occasionally increasing slightly, since it captures the ease of satisfying the acceptance condition for good lots. Influence of Consumer's Risk (β): Reduction in β (i.e., tighter consumer risk condition) produces a clear trend of increasing G and/or n. For example, for $r_2 = 2$, $r = 5$, $a_1 = 0.5$, a decrease in β from 0.25 to 0.05 raises G from 44 to 94 while n remains 2. This shows that decreasing the risk of the consumer requires an increased sampling effort so as to maintain the required $P_{\text{accept}} \geq 1 - \beta$. Notably, under $\beta = 0.01$ and $r_2 = 2$, $r = 5$, $a_1 = 0.5$, $G = 0$, $n = 0$, P_{accept} is evaluated as '–' when it is not feasible to construct a valid GASP for such stringent conditions for the specified test parameters. Repetition Factor (r) Impact: With a rise in the repetition

factor r (from 5 to 10), one often finds a reduction in G and/or n. For example, if $\beta = 0.25, r_2 = 2$, $a_1 = 0.5$, raising r from 5 to 10 lowers G from 44 to 7 (with n still 2).

This means that more inspection repetitions enable a more effective sampling plan with fewer groups or items in each run needed to get the required acceptance probability.

(ii)   Impact of the truncation time scaling factor (a_1): Increasing a_1 from 0.5 to 1 has a tendency to reduce G and/or n. To illustrate, at $\beta = 0.25, r_2 = 2, r = 5$, increasing a_1 from 0.5 to 1 reduces G from 44 to 7 (increasing n from 2 to 3). This means that a slightly longer test duration ($t_0 = a_1\mu_0$) can provide more information about product quality and therefore potentially allow a less costly sampling plan.

The trend, while generally generating more frugal plans, will be at the expense of G or n.

(iii) Non-Feasible Plans ($G = 0$ or $n = 0$): Cells of $G = 0$ and/or $n = 0$ generally reflect cases where no feasible sampling plan was possible to achieve the acceptance criteria. The '−' for P_{accept} also signals non-feasibility, particularly under highly strict β values, suggesting such lots would automatically be rejected or the conditions of the test are insufficient to distinguish lot quality adequately. In some cases where $n = 0$ but P_{accept} is available (e.g. $\beta = 0.25, r_2 = 6, r = 5, a_1 = 0.5$), acceptance implies that acceptance is solely based on the number of groups G.

The $P_{accept}(p)$ values always satisfy the available constraint, being large (usually ≥0.95), which is a testament to the efficacy of the optimized GASP designs in guaranteeing lot quality under the specified conditions.

Fig. 4a–d is the operating characteristic (OC) curve for

Case I: $(a_1 = 0.5, r = 5)$
Case II: $(a_1 = 1, r = 5)$
Case III: $(a_1 = 0.5, r = 10)$
Case IV: $(a_1 = 1, r = 10)$,

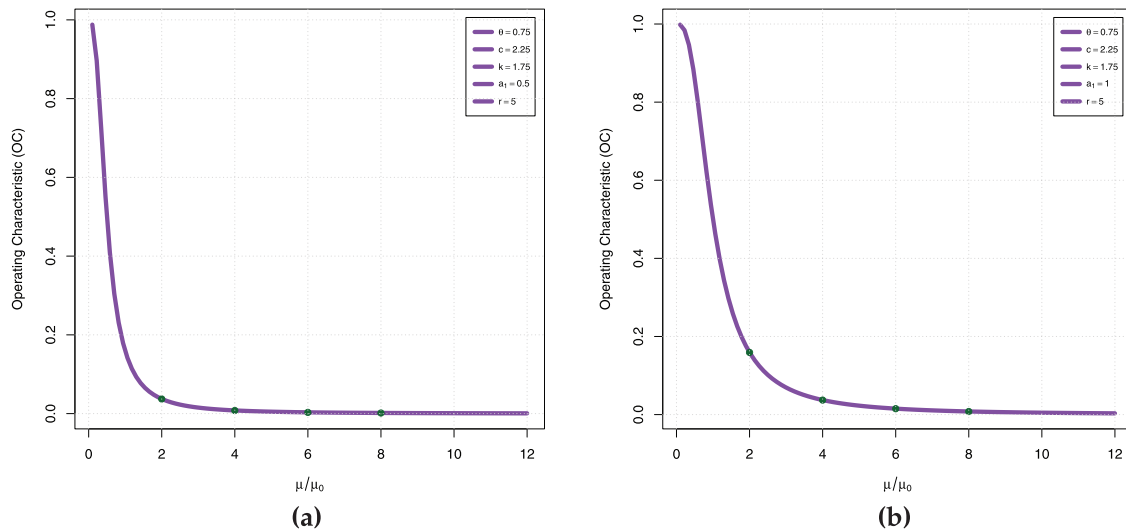respectively, based on $\theta = 0.75, c = 2.25, k = 1.75$. These support the results in Table 3.

Figure 4: (Continued)
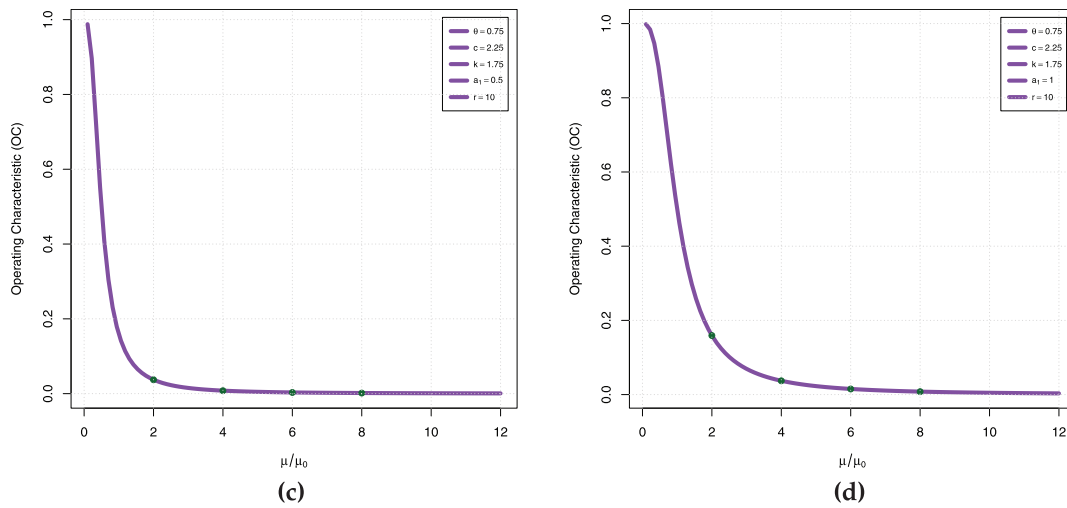
Figure 4: OC curve **(a)** case I **(b)** case II **(c)** case III **(d)** case IV

Table 4 provides the optimum parameters $(G, n, P_{\text{accept}}(p))$ of a Group Acceptance Sampling Plan (GASP) specially designed under truncated life tests for the TIHTBXII distribution. The distribution's own parameters in such designs are assigned as $-\theta = 0.5$, $c = 1.25$, and $k = 2.25$. The primary objective of any design is to determine the smallest number of groups (G) and group size per item (n) that satisfies the consumer's risk requirement, $P_{\text{accept}}(p) \geq 1 - \beta$, where p is the failure probability.

Table 4: GASP under TIHTBXII at $\theta = 0.5$, $c = 1.25$, $k = 2.25$ and with minimum G and n

β	$\frac{\mu}{\mu_0}$	$r = 5$											
		$a_1 = 0.5$			$a_1 = 1$			$r = 10$					
								$a_1 = 0.5$			$a_1 = 1$		
		G	n	$P_{\text{accept}}(p)$	G	n	$P_{\text{accept}}(p)$	G	n	$P_{\text{accept}}(p)$	G	n	$P_{\text{accept}}(p)$
0.25	2	903	4	0.969222	0	0	–	45	5	0.973959	8	6	0.952273
	4	10	2	0.983934	2	2	0.96585	2	2	0.968834	1	3	0.97012
	6	3	1	0.967734	2	2	0.991163	1	1	0.956315	1	2	0.960693
	8	3	1	0.98372	1	1	0.971935	1	1	0.977325	1	2	0.984293
0.10	2	0	0	–	0	0	–	75	5	0.956976	41	7	0.967275
	4	17	2	0.972842	12	3	0.985555	3	2	0.953617	3	4	0.985135
	6	17	2	0.993603	4	2	0.982404	3	2	0.9881	1	2	0.960693
	8	5	1	0.973015	4	2	0.993542	2	1	0.955164	1	2	0.984293
0.05	2	0	0	–	0	0	–	491	6	0.976796	54	7	0.957124
	4	22	2	0.964995	15	3	0.981976	10	3	0.984323	4	4	0.980229
	6	22	2	0.991729	5	2	0.978054	4	2	0.984165	2	3	0.98874
	8	6	1	0.967706	5	2	0.991935	2	1	0.955164	2	2	0.968834
0.01	2	0	0	–	0	0	–	754	6	0.964589	427	8	0.972694
	4	209	3	0.990179	23	3	0.972497	14	3	0.978122	5	4	0.975348
	6	34	2	0.987246	7	2	0.969411	6	2	0.976342	3	3	0.983158
	8	9	1	0.951952	7	2	0.988727	6	2	0.991338	2	2	0.968834

Each triplet $(G, n, P_{\text{accept}}(p))$ is the best configuration: G is the number of groups to sample, n is the number of elements in each group, and $P_{\text{accept}}(p)$ is the resulting accept probability when the true defect rate is p.

 (i) Influence of Mean Ratio (r_2): In all β, r, and a_1 regimes, an increase in r_2 (expressing a better quality lot) always brings about a substantial decrease in G and/or n. For example, for $\beta = 0.25$, $r = 5$, $a_1 = 0.5$, increasing r_2 from 2 to 8 reduces G from 903 to 3 and n from 4 to 1. This demonstrates a steep inverse relationship: improved lot quality needs far less severe sampling to reach the specified $P_{\text{accept}}(p)$. The $P_{\text{accept}}(p)$ values increase with r_2, indicating smoother acceptance of improved quality lots.
 (ii) Consumer's Risk (β) Influence: As β decreases (denoting a more risk-averse consumer's risk tolerance), there is the overall trend of increasing G and/or n. To illustrate, at $r_2 = 4$, $r = 10$, $a_1 = 0.5$, reducing β from 0.25 ($G = 2$, $n = 2$) to 0.01 ($G = 14$, $n = 3$) necessitates more extensive inspection. This confirms stricter risk control takes more extensive inspection in order to obtain the desired $P_{\text{accept}}(p) \geq 1 - \beta$. Few instances of $G = 0$, $n = 0$ (for example, for $\beta = 0.10$, $r_2 = 2$, $r = 5$, $a_1 = 0.5$) appear, particularly for lower levels of β, i.e., for which an optimal plan cannot be formulated in these extreme cases.
 (iii) Impact of Repetition Factor (r): An increase in the repetition factor r from 5 to 10 tends to decrease G and/or n significantly, thus making the sampling plan more effective.
     For example, for $\beta = 0.25$, $r_2 = 2$, $a_1 = 0.5$, increasing r from 5 to 10 decreases G from 903 to 45 (with n increasing from 4 to 5). This shows that increasing the number of rounds of inspection can drastically decrease the number of groups or items per round, essentially optimalizing resource distribution for a certain $P_{\text{accept}}(p)$.
 (iv) Effect of Truncation Time Scaling Factor (a_1): It tends to produce more cost-effective sampling plans (lower G and/or n) to raise the truncation time scaling factor a_1 from 0.5 to 1. Illustratively, when $\beta = 0.25$, $r_2 = 2$, $r = 10$, to lower a_1 from 0.5 ($G = 45$, $n = 5$) to 1 ($G = 8$, $n = 6$) drops G significantly.
     This indicates that having a relatively longer test duration ($t_0 = a_1\mu_0$) can result in more conclusive information on lot quality and may allow a less resource-intensive sampling plan. But a noteworthy observation to make here is that in case $r = 5$, increasing a_1 for $r_2 = 2$ and lower β values results in giving $G = 0$, $n = 0$ outcomes (for instance when $\beta = 0.25$, $r_2 = 2$, $r = 5$, $a_1 = 0.5$ yields $G = 903$, $n = 4$ while $a_1 = 1$ yields $G = 0$, $n = 0$). This shows a complex interaction where some combinations of parameters (with the current fixed distribution parameters $\theta = 0.5$, $c = 1.25$, $k = 2.25$) can render a plan infeasible even with increased test duration, and show specific challenges in designing optimum plans for such situations.
 (v) Non-Feasible Plans ($G = 0$, $n = 0$ with '–'): The high occurrence of $G = 0$, $n = 0$ with a '–' for $P_{\text{accept}}(p)$ (especially for lower levels of β and $r_2 = 2$) suggests cases where there was no possible GASP design to meet the requirements of acceptance. This implies that for such relatively high levels of consumer risk or low quality of lots (small r_2), along with specific values of r and a_1, the lot would be rejected outright or the current testing program is not adequate to offer the desired level of confidence. The observation that their prevalence is greater in this table than in previous analyses suggests that the current fixed parameters of distribution ($\theta = 0.5$, $c = 1.25$, $k = 2.25$) might be leading to more challenging scenarios for designing GASP plans with feasibility.

The accept values of $P_{\text{accept}}(p)$ normally adhere to the required constraint ($P_{\text{accept}}(p) \geq 1 - \beta$), normally maintaining high (often ≥ 0.95), justifying the optimisation process in ensuring the offered lot acceptance probabilities in case there exists an implementable plan.

Fig. 5a–d is the operating characteristic (OC) curve for Case I, Case II, Case III and Case IV, respectively, based on $\theta = 0.5$, $c = 1.25$, $k = 2.75$. These support the results in Table 4.
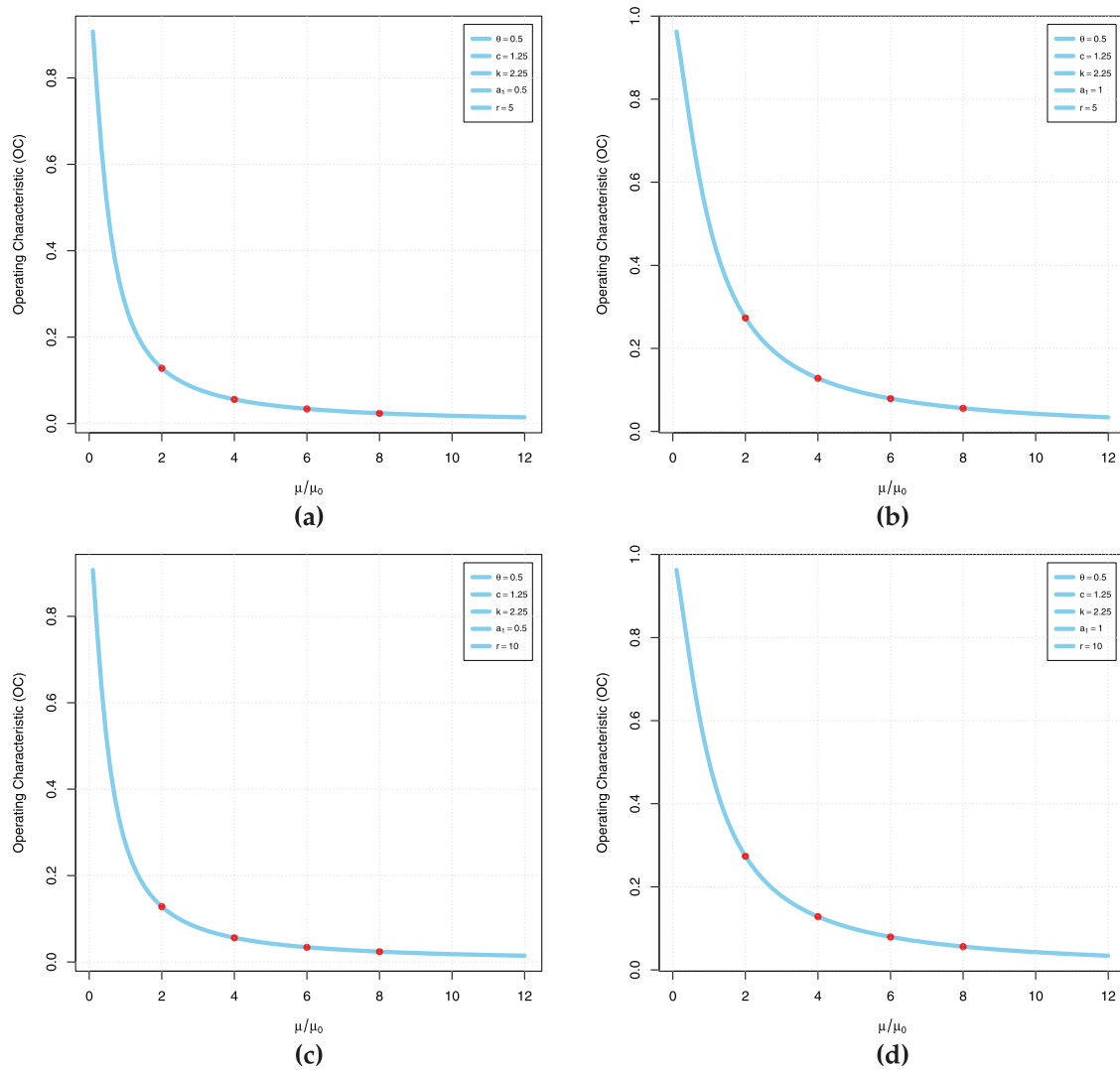
Figure 5: OC curve for $\theta = 0.5$, $c = 1.25$, $k = 2.25$ **(a)** case I **(b)** case II **(c)** case III **(d)** case IV

7 Applications

The first data to be used for illustrating the importance of the TIHTBXII model is the duration of active repairs for airborne communication transceivers. This data was studied by El-Saeed et al. [48] and Mead et al. [49] and presented in Table 5.

Table 5: Duration of active repairs for airborne communication transceivers

0.50	0.60	0.60	0.70	0.70	0.70	0.80	0.80	1.00	1.00
1.00	1.00	1.10	1.30	1.50	1.50	1.50	1.50	2.00	2.00
2.20	2.50	2.70	3.00	3.00	3.30	4.00	4.00	4.50	4.70
5.00	5.40	5.40	7.00	7.50	8.80	9.00	10.20	22.00	24.50

The second application data is related to groundwater contaminants measurements taken to monitor and assess the effectiveness of environmental cleanup efforts. It has been studied by Bhaumik et al. [50] and contained in Table 6, clean-up-gradient monitoring wells, expressed in milligrams per liter (mg/L).

Table 6: Groundwater contaminants measurements

5.1	1.2	1.3	0.6	0.5	2.4	0.5	1.1	8.0	0.8	0.4	0.6	0.9	0.4	2.0	0.5	5.3
3.2	2.7	2.9	2.5	2.3	1.0	0.2	0.1	0.1	1.8	0.9	2.0	4.0	6.8	1.2	0.4	0.2

Weekly Death rate due to COVID-19 from 22/3/2020 to 20/12/2020 in Dominica retrieved from https://data.who.int/dashboards/covid19/data?n=c (accessed on 20 August 2025) and reported in Table 7 below.

Table 7: Weekly death rate due to COVID-19 in Dominica (Data-III)

0.029850746	0.035363458	0.05292172	0.051236749	0.053061224	0.035197989
0.032082324	0.025607639	0.01929982	0.012975779	0.016090105	0.016615654
0.012140954	0.024329382	0.01301384	0.012219227	0.013564214	0.010826889
0.008958089	0.010932598	0.016161891	0.02201862	0.023125997	0.037598736
0.033069307	0.02618165	0.019748264	0.015044519	0.011788481	0.009540117
0.009618688	0.008082768	0.008333333	0.006851922	0.005307263	0.005720572
0.004095843	0.003398641	0.002529511	0.004015331		

We present a table of summary statistics with important measures such as the Q_1, Q_3, variance, standard deviation, skewness and kurtosis. This results are contained in Table 8.

Table 8: Summary statistics for the different datasets

Statistics	Duration of active repair	Groundwater contaminant measurements	Dominica COVID-19 mortality
n	40	34	40
Q_1	1	0.5	0.0094
Q_3	4.775	2.4750	0.0258
IQR	3.775	1.9750	0.0164
Outlier	22, 24.5	6, 6.8	0.05292172, 0.05123675, 0.05306122
Mean	4.0125	1.8794	0.0190
Median	2.1000	1.1500	0.0143
Variance	26.6801	3.8126	0.0002
Standard deviation	5.1653	1.9526	0.0136
Range	24	7.9	0.050531709
Skewness	2.7171	1.6037	1.0777
Kurtosis	10.5433	5.0054	3.4169

Table 8 brings together three different data sets, showing a similar pattern of positive skewness in each. This means each data set holds more small values and a tail that extends to high values due to occasional

gigantic observations or outliers. The Active Repair Duration data is the most skewed and dispersed, and it accounts for many short repairs but also for every now and then a very long one, as indicated by its very high skewness (2.7171) and kurtosis (10.5433). Groundwater Contaminant Measurements also show clear right-skewness (1.6037), indicating that most measurements are low but with the periodic high contaminant levels pulling the average quite high. Lastly, Dominica COVID-19 Mortality data, though still positively skewed (1.0777), is the most compact and least spread out with only outliers at higher mortality rates.

Fig. 6 contains boxplots superimposed on the violin plots for the three datasets, respectively. The image provide evidence that the three datasets contains outliers.

Duration of Active Repairs   Groundwater Contaminants Measurements   Dominica COVID−19 Mortality
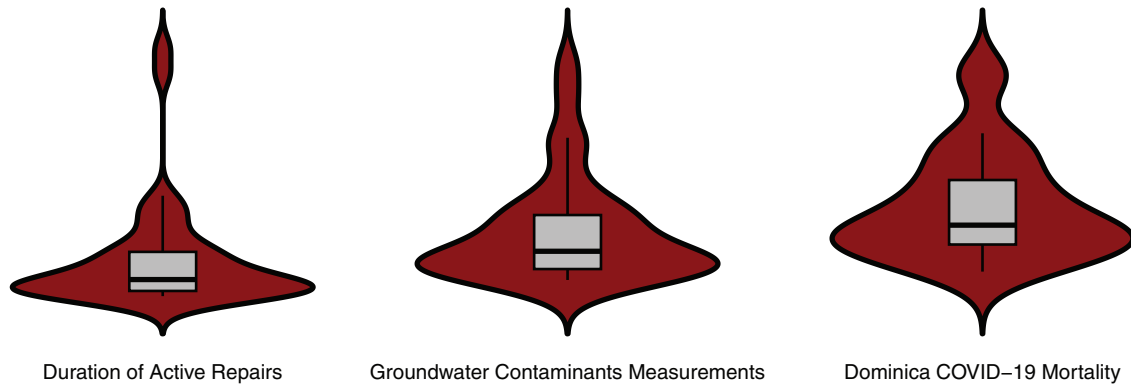
Figure 6: Boxplot superimposed on violin plots

Fig. 7 contains the density plots sumperimposed on the histogram of the respective datasets. All three show positive skewness.

Duration of Active Repairs   Groundwater Contaminants Measurements   Dominica COVID−19 Mortality
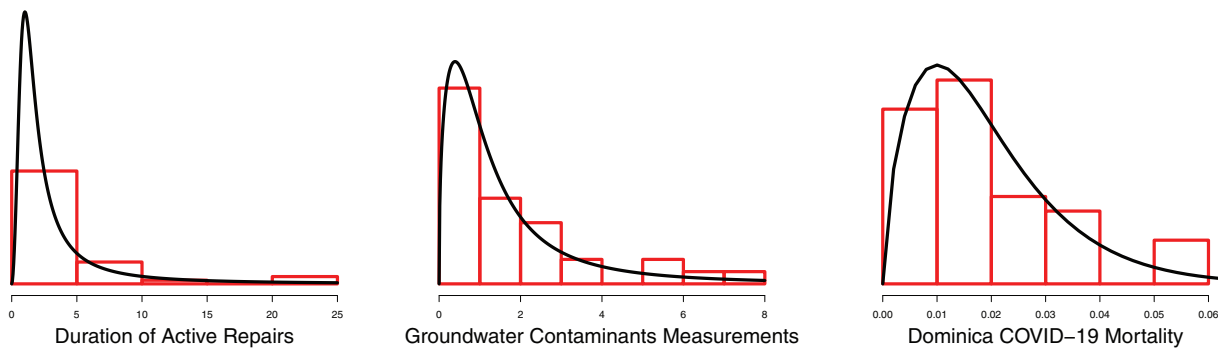
Figure 7: Kernel density superimposed on histogram

Figs. 8 and 9 are the TIHTBXII CDF and survival functions with those of the empirical from the datasets. These reflect the degree of fit of the proposed TIHTBXII to the various datasets.

Fig. 10 is the total time on test plot. The duration of active repairs and the groundwater contaminants measurements data show convex upwards TTT plots. These indicate a decreasing hazard rate (DHR). That is, as time goes on, the probability of failure for a surviving item decreases. This is often seen in early life failures where "weak" items fail quickly, leaving stronger ones. However, the Dominica COVID-19 mortality rate shows concave upward TTT plot. This indicates an increasing hazard rate (IHR). That is, as time goes on, the probability of failure for a surviving item increases.
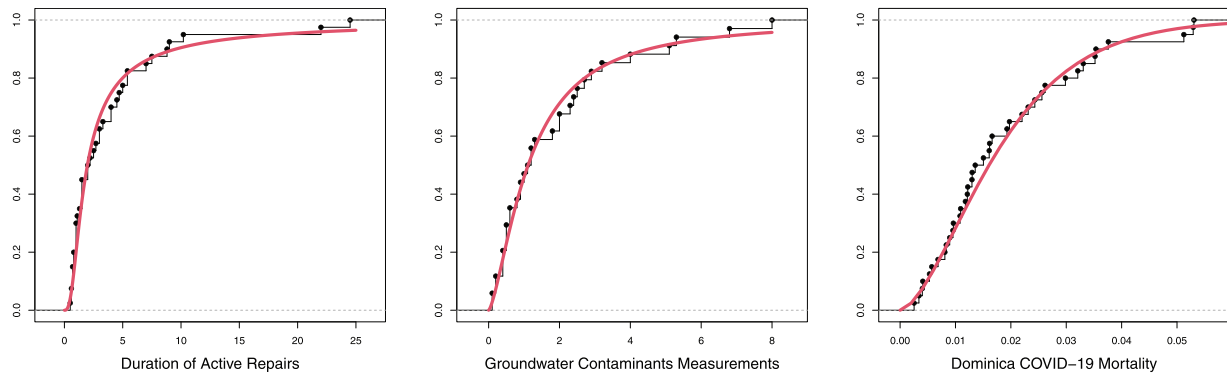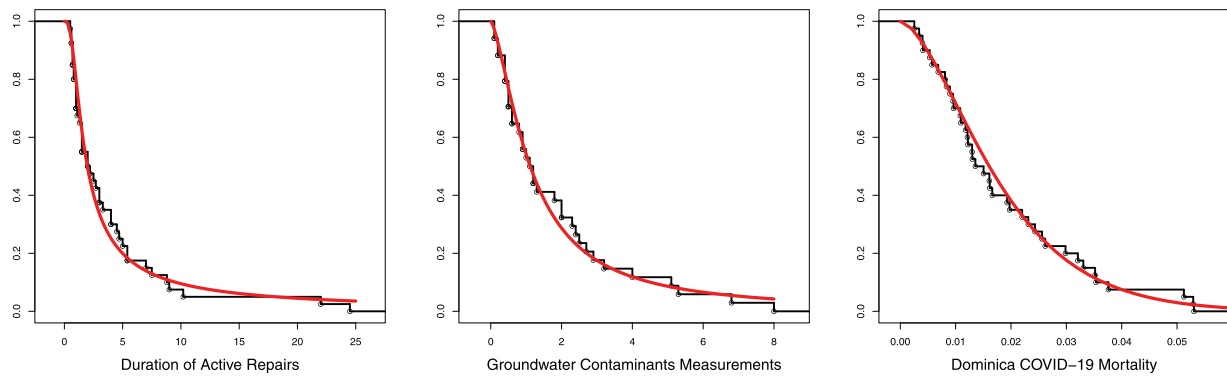
Figure 8: Empirical with TIHTBXII CDF

Figure 9: Empirical with TIHTBXII survival functions
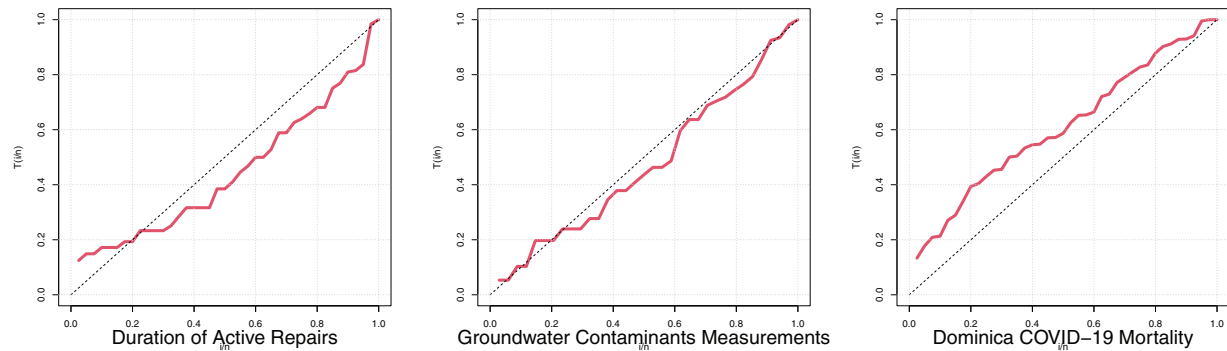

Figure 10: TTT plots

Fig. 11 is the P-P plots. For the three datasets, the P-P plots showed S-shapes or curved patterns that which indicate differences in skewness or kurtosis.

Fig. 12 is the Q-Q plots. For the groundwater contaminants measurements and the Dominica COVID-19 mortality rate, the Q-Q plots curved upwards at both ends (S-shape). These indicate heavier tails (more extreme values) in data than the TIHTBXII distribution. Hence, the datasets have higher kurtosis.
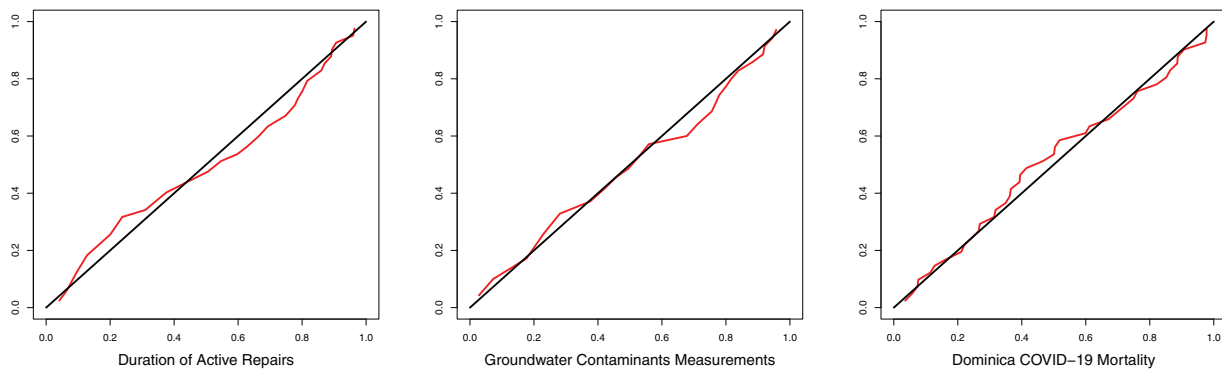
Figure 11: P-P plots
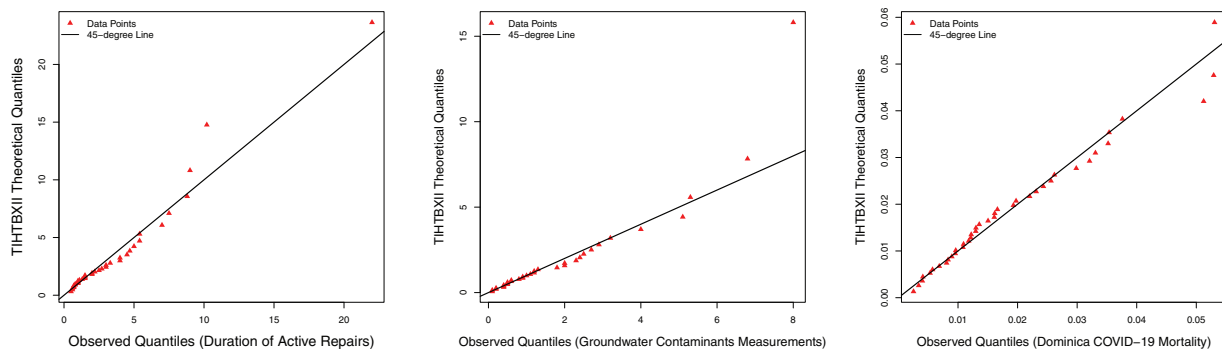
Figure 12: Q-Q plots

Figs. 13–15 are the contour plots for the three datasets. Each is a combination of any two of the parameters. The plots suggest the optimal points of the likelihood function which are contained in the MLEs table in Table 9.

The choice of competing distributions are the baseline distribution being modified which is Burr type XII (BXII) distribution by Burr [21], Kumaraswamy bell-Rayleigh (KwBR) distribution by Nadir et al. [51], Gamma distribution by Johnson et al. [52], Weibull distribution by [53], Gumbel distribution by Weibull [54] and Pareto I and III by Arnold [12].
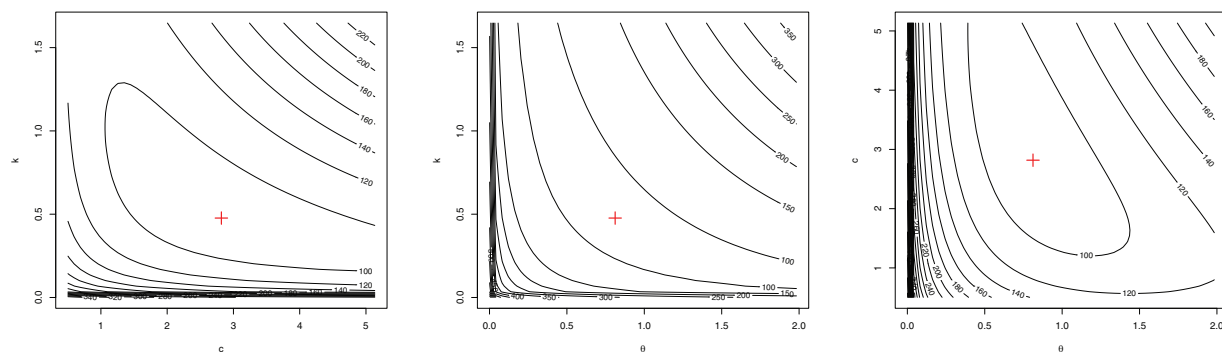
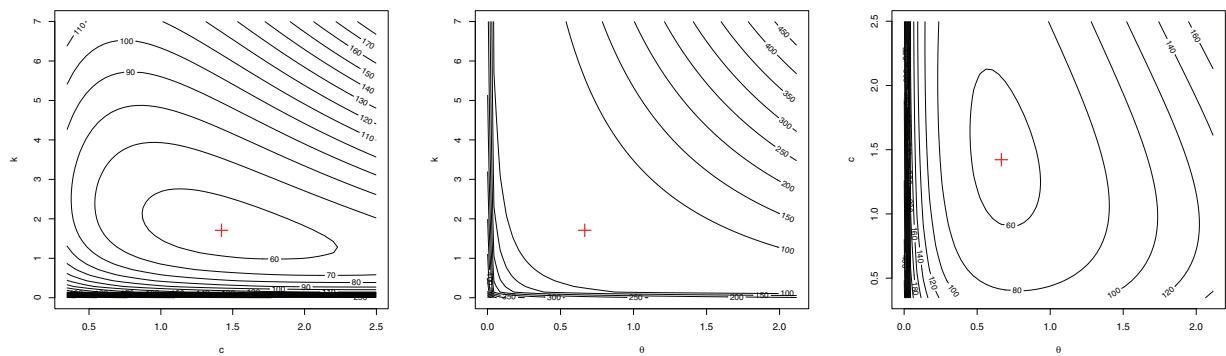**Figure 13:** Contour plots for duration of active repairs

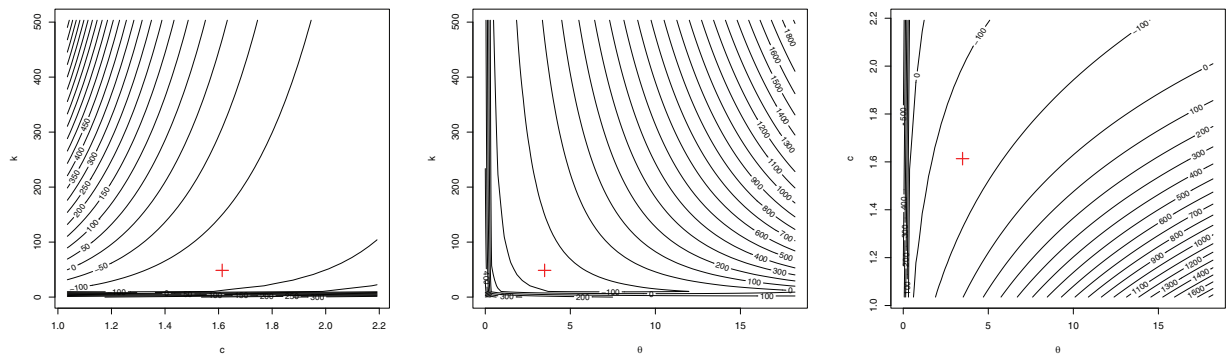**Figure 14:** Contour plots for groundwater contaminants measurements

Figure 15: Contour plots for Dominica COVID-19 mortality

Table 9: MLEs (Standard errors) for the parameters

Data	Distribution	MLEs			
	TIHTBXII	$\hat{\theta}$ 0.8117 (0.3903)	c 2.8181 (0.7715)	k 0.4768 (0.3901)	–
	BXII	– 	c 2.9725 (0.6776)	k 0.3383 (0.0871)	–
	KwBR	a 0.2766 (0.0108)	b 1100.345 (156.4983)	t 1.2590×10^{-12} (NAN)	s 49.3533 (0.0258)
Duration of active repairs	Gamma	a 1.0617 (0.2101)	b 0.2646 (0.0663)	–	–
	Weibull	l 3.9273 (0.6872)	k 0.9604 (0.1089)	–	–
	Gumbel	m 2.2238 (0.4033)	b 2.4764 (0.3475)	–	–
	Pareto I	$\hat{\alpha}$ 0.6482 (0.1025)	–	–	–
	Pareto III	$\hat{\alpha}$ 0.1031 (0.1157)	$\hat{\gamma}$ 0.1346 (0.1307)	$\hat{\sigma}$ 0.6467 (0.1566)	–

(Continued)

Table 9 (continued)

Data	Distribution	MLEs			
Groundwater contaminants measurements	TIHTBXII	$\hat{\theta}$ 0.6659 (0.4829)	$\hat{\sigma}$ 1.4221 (0.3580)	k 1.7090 (1.7620)	–
	BXII	– 	c 1.5623 (0.2480)	k 0.9304 (0.1791)	–
	KwBR	a 0.4808 (0.1894)	b 1.0997×10^3 (1.9067×10^2)	t 5.7×10^{-11} 1.2886 $\times 10^{-4}$	s 41.3964 (2.3539×10^{-3})
	Gamma	a 1.0624 (0.2281)	b 0.5653 (0.1535)	–	–
	Weibull	l 1.8879 (0.3390)	k 1.0103 (0.1327)	–	–
	Gumbel	m 1.0916 (0.2085)	b 1.1697 (0.1741)	–	–
	Pareto I	$\hat{\alpha}$ 0.4176 (0.0716)	–	–	–
	Pareto III	$\hat{\alpha}$ 4.2821 (9.5842)	$\hat{\gamma}$ 0.8735 (0.2503)	$\hat{\sigma}$ 5.7208 (14.8994)	–
Dominica COVID-19 mortality	TIHTBXII	$\hat{\theta}$ 1.8810 (1.7911)	c 1.6384 (0.2228)	k 185.1424 (293.9936)	–
	BXII	– 	c 1.4853 (0.1782)	k 309.4328 (201.5188)	–
	KwBR	a 0.7393 (0.0901)	b 329.5917 (1.2467×10^3)	t 4.6103 (3.4478)	s 16.2273 (NAN)
	Gamma	a 2.0449 (0.4252)	b 107.8358 (25.3974)	–	–
	Weibull	l 0.0211 (0.0024)	k 1.4830 (0.1797)	–	–
	Gumbel	m 0.0130 (0.0016)	b 0.0095 (0.0012)	–	–
	Pareto I	$\hat{\alpha}$ 0.5713 (0.0903)	–	–	–
	Pareto III	$\hat{\alpha}$ 3.9606 (6.8800)	$\hat{\gamma}$ 0.5899 (0.1425)	$\hat{\sigma}$ 0.0422 (0.0581)	–

Table 10 compares a number of statistical distributions for best fit for three datasets: "Duration of Active Repairs," "Groundwater Contaminant Measurements," and "Dominica COVID-19 Mortality." Comparison is by information criteria (Log-Lik, AIC, CAIC, BIC, HQIC) where lower values are preferable, and goodness-of-fit tests (W, A, D, and their p-values) where higher p-values (ideally > 0.05) are desirable.

Table 10: Model performance and goodness of fit measures

Data	Distribution	Log-Lik	AIC	CAIC	BIC	HQIC	W	A	D	p-value
Duration of active repairs	TIHTBXII	−90.89	187.7884	188.4551	192.8550	189.6203	0.0850	0.5369	0.0992	0.8256
	BXII	−90.99	185.9732	186.2975	189.3510	187.1945	0.0898	0.5624	0.1033	0.7867
	KwBR	85.03	−279.7573	−278.6145	−273.0018	−277.3148	–	–	1.0000	2.2×10^{-16}
	Gamma	−95.53	195.0636	195.3880	198.4414	196.2849	0.1506	1.0824	0.1507	0.3239
	Weibull	−95.51	195.0227	195.3470	198.4005	196.2440	0.1455	1.0483	0.1290	0.5184
	Gumbel	−105.16	214.3195	214.6439	217.6939	215.5408	0.3148	2.0151	0.1880	0.1182
	Pareto I	−91.31	184.6242	184.7295	186.3131	185.2349	–	–	0.1619	0.2449
	Pareto III	−90.59	187.1752	187.8419	192.2418	189.0071	0.1091	0.6654	0.1289	0.5193

(Continued)

Table 10 (continued)

Data	Distribution	Log-Lik	AIC	CAIC	BIC	HQIC	W	A	D	*p*-value
Groundwater contaminants measurements	TIHTBXII	−55.89	117.7857	118.5857	122.3648	119.3473	0.0392	0.2667	0.0950	0.9188
	BXII	−56.08	116.1526	116.5396	119.2053	117.1936	0.0418	0.2927	0.1040	0.8555
	KwBR	85.81	−199.9203	−198.5409	−193.8148	−197.8381	0.2814	1.9492	0.9319	2.2×10^{-16}
	Gamma	−55.41	114.8263	115.2134	117.8790	115.8674	0.0459	0.2975	0.0973	0.9044
	Weibull	−55.45	114.8992	115.2862	117.9520	115.9403	0.0463	0.3001	0.0918	0.9366
	Gumbel	−62.23	128.4557	128.8428	131.5085	129.4968	0.1633	1.0407	0.1569	0.3724
	Pareto I	−66.8	135.5943	135.7193	137.1207	136.1149	–	–	0.3219	0.0017
	Pareto III	−55.44	114.8789	115.2660	117.9316	115.9200	0.0433	0.2800	0.0814	0.9779
Dominica COVID-19 mortality	TIHTBXII	123.17	−240.5629	−239.8962	−235.4963	−238.7310	0.0421	0.2767	0.0856	0.9068
	BXII	122.98	−241.9562	−241.6319	−238.5785	−240.7349	0.0540	0.3482	0.0851	0.8286
	KwBR	122.92	−237.9403	−236.7974	−231.1847	−235.4977	0.0543	0.3507	0.0958	0.8225
	Gamma	123.51	−243.0156	−242.6913	−239.6378	−241.7943	0.0346	0.2328	0.0843	0.9157
	Weibull	122.97	−241.9478	−241.6235	−238.5701	−240.7266	0.0543	0.3501	0.0955	0.8257
	Gumbel	121.26	−238.5247	−238.2004	−235.147	−237.3034	0.0752	0.4787	0.1110	0.6664
	Pareto I	106.78	−211.5517	−211.4464	−209.8628	−210.9410	–	–	0.3100	0.0007
	Pareto III	123.12	−240.247	−239.5803	−235.1804	−238.4151	0.0381	0.2525	0.0827	0.9266

For the Active Repairs Duration data, the KwBR distribution, despite seemingly good information criteria, is definitely a poor fit due to its extremely low *p*-value (2.2×10^{-16}). Among the remainder, the BXII and TIHTBXII distributions are the most promising with the lowest information criteria and high *p*-values, which is indicative of a good fit to the data. The Pareto I distribution provides a decent fit, with a competitive AIC and a *p*-value of 0.2449. The Pareto III distribution also performs well, with a slightly higher AIC but a very good *p*-value of 0.5193, suggesting it is a suitable model.

For Groundwater Contaminants Measurements, the KwBR distribution again has a very poor fit with a null *p*-value. The Gamma and Weibull distributions and TIHTBXII provide the best fits, which are reinforced by their competitive information criteria and high *p*-values, suggesting that these models adequately describe the features of the data. The Pareto I distribution is a very poor fit for this data, as indicated by its high AIC and an extremely low *p*-value of 0.0017. In contrast, the Pareto III distribution is an excellent fit, with a low AIC and the highest *p*-value (0.9779) among all tested distributions, making it a very strong candidate for this dataset.

While comparable standard errors might make the additional parameter of TIHTBXII seem less useful, its real value is that it can provide a better fit for datasets that have more complex and highly skewed distributions. The "Duration of Active Repairs" data provides a particular case where, even with the penalty for an extra parameter, the TIHTBXII outperformed the BurrXII on the key goodness-of-fit measures. This is an indication that the extra parameter is not superfluous; it is capturing valuable information on the data structure that the simpler BurrXII model cannot capture effectively. The benefit of TIHTBXII is not always the best model, but its additional flexibility, so that it is a more viable choice for a larger array of highly skewed data from the real world, such as a hypothetical time series of volatile stock returns.

Finally, for the Dominica COVID-19 Mortality Data, the Gamma Distribution is by far the most appropriate model. It has the lowest information criteria and a high *p*-value (0.9157), highly indicating its suitability. Other distributions like TIHTBXII and BXII also yield good fits. The Pareto I distribution performs very poorly, with a high AIC and a *p*-value of 0.0007. The Pareto III distribution, however, provides a very good fit, with an excellent *p*-value of 0.9266 and competitive information criteria, making it a strong contender for this dataset. Again, the TIHTBXII has the best fit in the duration of active repairs data.

Table 10 provides the Maximum Likelihood Estimates (MLEs) of the parameters of various statistical distributions fitted to three data sets: "Duration of Active Repairs," "Groundwater Contaminant Measurements," and "Dominica COVID-19 Mortality." Along with each MLE, in parentheses, is its standard error, a measure of the precision of the estimate; smaller standard errors indicate more precise estimates. For the Duration of Active Repairs data, the parameter estimates vary considerably across distributions. The BXII distribution, for instance, estimates the parameters c as 2.9725 and k as 0.3383, both of which have relatively small standard errors, suggesting stable estimates. Similar results were obtained for the proposed TIHTBXII distribution. For the Pareto distributions, Pareto I yields an $\hat{\alpha}$ of 0.6482 with a standard error of 0.1025, while the three-parameter Pareto III distribution gives an $\hat{\alpha}$ of 0.1031 with a larger standard error of 0.1157, indicating less precision for that parameter. The Pareto III's other parameters, $\hat{\gamma}$ and $\hat{\sigma}$, also have relatively large standard errors, suggesting potential instability. In contrast, the KwBR distribution shows a parameter t with an extremely small MLE (1.2590×10^{-12}) and a "NAN" (Not a Number) standard error, which often signifies estimation problems or model instability, supporting the poor fit of the KwBR from a previous analysis.In the Groundwater Contaminant Measurements data, parameter estimates differ by distribution. For the Gamma distribution, a is estimated as 1.0624 and b as 0.5653, both with reasonable standard errors. Again, the KwBR model for this dataset has unusual estimates, notably for t (5.7×10^{-11}) with a very small standard error, which can still be problematic given its overall lack of fit. For the Pareto distributions, Pareto I estimates $\hat{\alpha}$ at 0.4176 with a small standard error of 0.0716, pointing to a stable estimate. However, for Pareto III, the $\hat{\alpha}$ estimate is 4.2821 with a very large standard error (9.5842), and the $\hat{\sigma}$ estimate is 5.7208 with an even larger standard error (14.8994), suggesting these parameters are not well-determined by the data. For Dominica COVID-19 Mortality, the Gamma distribution, which was a strong candidate in previous goodness-of-fit analyses, has MLEs for a of 2.0449 and b of 107.8358, both of which are stable with relatively small standard errors. By contrast, the BXII distribution for these data has a very large estimate of k (309.4328) with a very large standard error (201.5188), indicating less precision in that estimate. The Pareto I distribution gives an $\hat{\alpha}$ of 0.5713 with a standard error of 0.0903. The Pareto III distribution, on the other hand, shows an $\hat{\alpha}$ of 3.9606 with a very large standard error (6.8800), indicating high uncertainty in this estimate. The KwBR distribution once more has an estimate of s with a "NAN" standard error, reconfirming its unreliability for this data too. In general, the MLEs provide the true values of the parameters that best describe each data set according to its respective distribution. The standard errors provide a sense of how reliable these estimates are. The systematic presence of "NAN" or extremely small/enormous estimates with large standard errors, particularly for the KwBR and in some cases for the Pareto III distribution across all data sets, is a clear sign of issues with its parameter estimation and general applicability, which is in line with its poor goodness-of-fit behavior.

8 Conclusion

In this study, we introduced the Type-I Heavy-Tailed Burr Type XII (TIHTBXII) distribution as a highly flexible and robust statistical model. Our primary motivation was to address the limitations of conventional distributions in accurately capturing the complexities of real-world data, particularly those exhibiting skewness, heavy tails, and diverse hazard behaviors, common in fields such as finance, insurance, environmental science, and reliability engineering. We meticulously defined the TIHTBXII distribution, providing its PDF and CDF, and thoroughly investigated its fundamental statistical properties, including its quantile function, moments, moment-generating function, order statistics, and entropy. These foundational elements are crucial for both theoretical understanding and practical application. A significant part of our research focused on parameter estimation, where we rigorously compared four standard methods: MLE, MPS, LS, and WLS. Our extensive Monte Carlo simulation studies consistently demonstrated the consistency

of these estimators; as sample size increased, the bias and RMSE of the parameter estimates decreased for all methods. Notably, the WLS and LS procedures generally outperformed MLE and MPS, exhibiting lower bias and RMSE, which underscores their stability and robustness across various scenarios and sample sizes. While MLE and MPS improved with larger sample sizes, LS and WLS consistently proved more stable and accurate in estimation for all sample sizes and parameter values considered.

Beyond theoretical exploration and estimation, we showcased the practical utility of the TIHTBXII distribution by developing a Group Acceptance Sampling Plan (GASP) using truncated life tests. This methodology is particularly valuable for quality control and life-testing procedures involving long-lifespan products and specific failure modes. We provided a comprehensive guide for optimizing GASP design parameters (number of groups, acceptance number, group size, and test duration) by minimizing the Average Sample Number (ASN) while effectively balancing consumer and producer risks. Our research highlighted how factors such as the mean ratio, consumer's risk, repetition factor, and truncation time scaling factor influence the stringency and efficiency of the sampling plan. The TIHTBXII distribution's ability to realistically model lifetime data directly translates into more efficient and risk-balanced decision-making in industrial settings. Empirical validation further solidified the TIHTBXII distribution's applicability and versatility. We successfully applied the distribution to real-world datasets, including Active Repair Duration, Groundwater Contaminant Measurements, and Dominica COVID-19 Mortality. In cases like the "Active Repair Duration" data, the TIHTBXII distribution demonstrated a superior fit compared to other established models, as supported by favorable information criteria and goodness-of-fit test statistics. This empirical evidence underscores the TIHTBXII's potential as a valuable tool for analysts and engineers working with heavy-tailed, skewed, and complex data in reliability engineering and quality measurement.

While this study effectively demonstrates the significant potential of the TIHTBXII distribution, it's important to acknowledge certain limitations. Our Monte Carlo simulation studies, though extensive, were conducted under specific assumptions regarding parameter values and sample sizes. The performance of estimation methods, particularly MLE and MPS, might exhibit different convergence rates or biases in scenarios with extremely small sample sizes or highly unusual parameter combinations not explored in our simulations. Second, the GASP design was developed specifically for truncated life tests. The applicability and optimality of the proposed GASP framework might vary for other life-testing scenarios, such as complete life tests or different truncation mechanisms. Finally, while the empirical applications showcased the TIHTBXII's superior fit for the selected datasets, its generalizability and superiority across all possible heavy-tailed and skewed datasets need further validation through broader comparative studies. Future research could explore its performance with even more diverse and complex real-world data from various domains to fully ascertain its range of applicability.

In conclusion, the Type-I Heavy-Tailed Burr Type XII distribution is a robust and valuable addition to the family of statistical models. It offers enhanced flexibility and strength to accurately identify realistic data patterns, particularly for data characterized by skewness and heavy tails. Its improved parameter estimation capabilities and proven effectiveness in real-world problems like acceptance sampling make it an effective alternative to traditional distributions, ultimately leading to more reliable and accurate decisions across various scientific and engineering disciplines.

References

1. Nolan JP. Financial modeling with heavy-tailed stable distributions. Wiley Interdiscip Rev Comput Stat. 2014;6(1):45–55. doi:10.1002/wics.1286.

2. Alsubie A. On modeling the insurance claims data using a new heavy-tailed distribution. In: Intelligent Decision Technologies: Proceedings of the 14th KES-IDT 2022 Conference. Cham, Switzerland: Springer. 2022, pp. 149–58. doi:10.1007/978-981-19-3444-5.

3. Macdonald E, Merz B, Nguyen VD, Vorogushyn S. Heavy-tailed flood peak distributions: what is the effect of the spatial variability of rainfall and runoff generation? Hydrol Earth Syst Sci. 2025;29(2):447–63. doi:10.5194/hess-29-447-2025.

4. Navas-Portella V, González Á, Serra I, Vives E, Corral Á. Universality of power-law exponents by means of maximum-likelihood estimation. Phys Rev E. 2019;100(6):062106. doi:10.1103/PhysRevE.100.062106.

5. Endo A, Murayama H, Abbott S, Ratnayake R, Pearson CB, Edmunds WJ, et al. Heavy-tailed sexual contact networks and monkeypox epidemiology in the global outbreak, 2022. Science. 2022;378(6615):90–4. doi:10.1126/science.add4507.

6. Karlsson M, Wang Y, Ziebarth NR. Getting the right tail right: modeling tails of health expenditure distributions. J Health Econ. 2024;97(3):102912. doi:10.1016/j.jhealeco.2024.102912.

7. Tokutomi N, Nakai K, Sugano S. Extreme value theory as a framework for understanding mutation frequency distribution in cancer genomes. PLoS One. 2021;16(8):e0243595. doi:10.1371/journal.pone.0243595.

8. Cauchy AL. Sur les résultats moyens d'observations de même nature, et sur les résultats les plus probables. CR Acad Sci Paris. 1853;37:198–206.

9. Student. The probable error of a mean. Biometrika. 1908;6(1):1–25. doi:10.2307/2331554.

10. Fréchet M. Sur la loi de probabilité de l'écart maximum. Ann De La Soc Polonaise De Math. 1927;6(2):123–49.

11. Pareto V. Cours d'économie politique. Lausanne, Switzerland: F. Rouge; 1897.

12. Arnold BC. Pareto distributions. Fairland, MD, USA: International Cooperative Publishing House; 1983. doi:10.1007/978-1-4615-6805-4.

13. Lomax KS. Business failures: another example of the analysis of failure data. J Am Stat Assoc. 1954;49(268):847–52. doi:10.1080/01621459.1954.10501239.

14. Beirlant J, Matthys G, Dierckx G. Heavy-tailed distributions and rating. ASTIN Bull J IAA. 2001;31(1):37–58. doi:10.2143/AST.31.1.993.

15. Orji GO, Etaga HO, Almetwally EM, Igbokwe CP, Aguwa OC, Obulezi OJ. A new odd reparameterized exponential transformed-X family of distributions with applications to public health data. Innov Stat and Prob. 2025;1(1):88–118. doi:10.64389/isp.2025.01107.

16. Husain QN, Qaddoori AS, Noori NA, Abdullah KN, Suleiman AA, Balogun OS. New expansion of Chen distribution according to the nitrosophic logic using the Gompertz family. Innov Stat Prob. 2025;1(1):60–75. doi:10.64389/isp.2025.01105.

17. Gemeay AM, Moakofi T, Balogun OS, Ozkan E, Hossain MM. Analyzing real data by a new heavy-tailed statistical model. Modern J Stat. 2025;1(1):1–24. doi:10.64389/mjs.2025.01108.

18. Mazza A, Punzo A. Modeling household income with contaminated unimodal distributions. In: New statistical developments in data science (SIS 2017). Cham, Switzerland: Springer; 2017. p. 373–91. doi:10.1007/978-3-030-21158-5.

19. Punzo A, Mazza A, Maruotti A. Fitting insurance and economic data with outliers: a flexible approach based on finite mixtures of contaminated gamma distributions. J Appl Stat. 2018;45(14):2563–84. doi:10.1080/02664763.2018.1428288.

20. Punzo A, Bagnato L, Maruotti A. Compound unimodal distributions for insurance losses. Insur Math Econ. 2018;81(13–14):95–107. doi:10.1016/j.insmatheco.2017.10.007.

21. Burr IW. Cumulative frequency functions. Ann Math Stat. 1942;13(2):215–32.

22. Singh SK, Maddala GS. A function for size distribution of incomes. In: Modeling income distributions and Lorenz curves. Cham, Switzerland: Springer; 2008. p. 27–35. doi:10.1007/978-0-387-72796-7_2.

23. Gad AM, Hamedani GG, Salehabadi SM, Yousof HM. The Burr XII-Burr XII distribution: mathematical properties and characterizations. Pak J Stat. 2019;35(3):229–48.

24. Alzaatreh A, Lee C, Famoye F. A new method for generating families of continuous distributions. J Appl Stat. 2013;40(9):1606–21. doi:10.1007/s40300-013-0007-y.

25. Bhatti FA, Hamedani GG, Korkmaz M, Sheng W, Ali A. On the Burr XII-moment exponential distribution. PLoS One. 2021;16(2):e0246935. doi:10.1371/journal.pone.0246935.

26. Alizadeh M, Cordeiro GM, de Castro A, Santos PA. The Kumaraswamy-Burr XII distribution. J Stat Comput Simul. 2015;85(13):2697–2715. doi:10.1080/00949655.2012.683003.

27. Al-Saiari AY, Baharith LA, Mousa SA. Marshall-Olkin extended Burr type XII distribution. Int J Stat Probab. 2014;3(1):78–84.

28. Alsadat N, Nagarjuna VBV, Hassan AS, Elgarhy M, Ahmad H, Almetwally EM. Marshall–Olkin Weibull–Burr XII distribution with application to physics data. AIP Adv. 2023;13(9):095325. doi:10.1063/5.0172143 2023.

29. Hassan MHO, Elbatal I, Al-Nefaie AH, Elgarhy M. On the Kavya–Manoharan–Burr X model: estimations under ranked set sampling and applications. J Risk Financ Manag. 2023;16(1):19.

30. Algarni A, MAlmarashi A, Elbatal I, Hassan SA, Almetwally EM, Daghistani MA, et al. Type I half logistic Burr XG family: properties, Bayesian, and non-Bayesian estimation under censored samples and applications to COVID-19 data. Math Probl Eng. 2021;2021(1):5461130.

31. Bantan RA, Chesneau C, Jamal F, Elbatal I, Elgarhy M. The truncated burr X-G family of distributions: properties and applications to actuarial and financial data. Entropy. 2021;23(8):1088.

32. Ahsanullah M, Shakil M, Elgarhy M, Kibria BM. On a generalized Burr life-testing model: characterization, reliability, simulation, and Akaike information criterion. J Stat Theory Appl. 2019;18(3):259–69.

33. Haq M, Elgarhy M, Hashmi S. The generalized odd Burr III family of distributions: properties, and applications. J Taibah Univ Sci. 2019;13(1):961–71.

34. Ocloo SK, Brew L, Nasiru S, Odoi B. On the extension of the Burr XII distribution: applications and regression. Comput J Math Stat Sci. 2023;2(1):1–30. doi:10.21608/cjmss.2023.181739.1000.

35. Isa AM, Ali BA, Zannah U. Sine Burr XII distribution: properties and application to real data sets. Arid Zone J Basic Appl Res. 2022;1:48–58.

36. Noori NA. Exploring the properties, simulation, and applications of the odd Burr XII Gompertz distribution. Adv Theory Nonlinear Anal Appl. 2023;7(4):60–75. doi:10.17762/atnaa.v7.i4.283.

37. Aslam M, Jun C-H. A group acceptance sampling plans for truncated life tests based on the inverse Rayleigh and log-logistics distributions. Pak J Stat. 2009;25(2):107–19.

38. Rao GS. A group acceptance sampling plans based on truncated life tests for Marshall-Olkin extended Lomax distribution. Electronic J Appl Stat Anal. 2009;3(1):18–27. doi:10.1285/i20705948v3n1p18.

39. Singh S, Tripathi YM. Acceptance sampling plans for inverse Weibull distribution based on truncated life test. Life Cycle Reliab Saf Eng. 2017;6(3):169–78. doi:10.1007/s41872-017-0022-8.

40. Almarashi AM, Khan K. Optimizing group size using percentile based group acceptance sampling plans with application. Contemp Math. 2024;5(4):4763–75. doi:10.37256/cm.5420245193.

41. Owoloko EA, Oguntunde PE, Adejumo AO. Performance rating of the transmuted exponential distribution: an analytical approach. SpringerPlus. 2015;4(1):818. doi:10.1186/s40064-015-1590-6.

42. Saha M, Tripathi H, Dey S. Single and double acceptance sampling plans for truncated life tests based on transmuted Rayleigh distribution. J Ind Prod Eng. 2021;38(5):356–68. doi:10.1080/21681015.2021.1893843.

43. Ameeq M, Naz S, Hassan MM, Fatima L, Shahzadi R, Kargbo A. Group acceptance sampling plan for exponential logarithmic distribution: an application to medical and engineering data. Cogent Eng. 2024;11(1):2328386. doi:10.1080/23311916.2024.2328386.

44. Ekemezie D-FN, Alghamdi FM, Aljohani HM, Riad FH, Abd El-Raouf MM, Obulezi OJ. A more flexible Lomax distribution: characterization, estimation, group acceptance sampling plan and applications. Alex Eng J. 2024;109(7):520–31. doi:10.1016/j.aej.2024.09.005.

45. Zhao W, Khosa SK, Ahmad Z, Aslam M, Afify AZ. Type-I heavy tailed family with applications in medicine, engineering and insurance. PLoS One. 2020;15(8):e0237462. doi:10.1371/journal.pone.0237462.

46. Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E. Equation of state calculations by fast computing machines. J Chem Phys. 1953;21(6):1087–92. doi:10.1063/1.1699114.

47. Hastings WK. Monte Carlo sampling methods using Markov chains and their applications. Biometrika. 1970;57(1):97–109. doi:10.1093/biomet/57.1.97.

48. El-Saeed AR, Obulezi OJ, Abd El-Raouf MM. Type II heavy tailed family with applications to engineering, radiation biology and aviation data. J Radiat Res Appl Sci. 2025;18(3):101547. doi:10.1016/j.jrras.2025.101547.

49. Mead M, Nassar MM, Dey S. A generalization of generalized gamma distributions. Pak J Stat Oper Res. 2018;14(1):121–38. doi:10.18187/pjsor.v14i1.1692.

50. Bhaumik DK, Kapur K, Gibbons RD. Testing parameters of a gamma distribution for small samples. Technometrics. 2009;51(3):326–34. doi:10.1198/tech.2009.07038.

51. Nadir S, Aslam M, Anyiam KE, Alshawarbeh E, Obulezi OJ. Group acceptance sampling plan based on truncated life tests for the Kumaraswamy Bell–Rayleigh distribution. Sci Afr. 2025;27(6):e02537. doi:10.1016/j.sciaf.2025.e02537.

52. Johnson NL, Kemp AW, Kotz S. Univariate discrete distributions. Hoboken, NJ, USA: John Wiley & Sons; 2005.

53. Weibull W. A statistical theory of strength of materials. Stockholm, Sweden: Generalstabens Litografiska Anstalts Forlag; 1939.

54. Gumbel EJ. Statistics of extremes. New York, NY, USA: Columbia University Press; 1958.