Vol. 34, No. 4 (2024) | DOI: 10.37190/ord240401

OPEN ACCESS

Operations Research and Decisions

www.ord.pwr.edu.pl

OPERATIONS RESEARCH AND DECISIONS QUARTERLY

A unit Weibull loss distribution with quantile regression and practical applications to actuarial science

Abdul Ghaniyyu Abubakari[1] Suleman Nasiru[1] Christophe Chesneau[2*]

[1]*Department of Statistics and Actuarial Science, School of Mathematical Sciences, C. K. Tedam University of Technology and Applied Sciences, Ghana*

[2]*Department of Mathematics, LMNO, University of Caen, 14032 Caen, France*

[*]*Corresponding author, email address: christophe.chesneau@unicaen.fr*

Abstract

A new bounded distribution called the unit Weibull loss distribution has been studied. The corresponding probability density function plots reveal that it is suitable to analyze data that exhibit right skewness, left skewness, and approximately symmetric and decreasing shapes. Furthermore, the corresponding hazard rate function plots indicate that it is adequate to fit data that have J, bathtub, and modified bathtub hazard rate shapes. This makes the new distribution suitable for modeling data with complex characteristics. Statistical properties such as the quantile, moments, and moment-generating function are determined. Risk measures, including the value-at-risk, tail value-at-risk, and tail variance are also calculated. Furthermore, different principles are derived for the computation of insurance premiums. The parameters of the distribution are estimated using different methods, and their performance is assessed via Monte Carlo simulations. The accuracy of the estimates is thus empirically demonstrated. A quantile regression model with responses following the unit distribution is developed. Applications of the proposed distribution and its corresponding regression model to three insurance data sets are carried out, with their performance compared with other models. The results show that they outperform the competitors. Thus, the new methodology can serve as an alternative option to analyze insurance data.

Keywords: *loss models, claims, risk measures, premium principles, regression models*

1. Introduction

In several fields, accurate data analysis is very important for decision-making. For instance, in actuarial science, the measurement of risk for management purposes and analyses for premium calculations are crucial. Also in the health sciences, understanding the prevalence of a disease is primordial. A notable example is the recent COVID-19 outbreak. Although many distributions have been developed and used, new distributions must be created to model the increasingly complex data, the possibility of distinct

Received 6 January 2024, accepted 26 September 2024, published online 19 December 2024
ISSN 2391-6060 (Online)/

The costs of publishing this issue have been co-financed by the Department of Operations Research and Business Intelligence at the Faculty of Management, Wrocław University of Science and Technology, Wrocław, Poland

features in data from different domains, and the inability of a single distribution to provide a universal model (in the sense "for any type of data").

In particular, several distributions have been developed with positive support or on the whole real line. However, some of the data generated are defined on the unit interval only. These include percentages, ratios, rates, and proportions. Thus, researchers elaborated on several distributions defined on bounded intervals, specifically unit intervals, for modeling such data. Popular unit distributions are the beta and Kumaraswamy [23] distributions. Some other unit distributions include the Topp–Leone (TL) distribution [37], log–Lindley distribution [13], unit Birnbaum–Saunders distribution [25], unit Weibull (UW) distribution [26], unit Lindley (UL) distribution [27], log–extended exponential-geometric distribution [15], unit Rayleigh (UR) distribution [7], unit gamma/Gompertz distribution [8], unit Burr XII distribution [19], unit Chen distribution [18], unit folded normal distribution [20], unit Teissier distribution [22], unit exponentiated Fréchet distribution [1], Marshall-Olkin reduced Kies distribution [2], generalized unit half-logistic geometric distribution [32], and power unit Burr XII distribution [39]. Others include the unit-improved second-degree Lindley distribution [5], unit logistic distribution [29], logit slash distribution [17], unit exponential Pareto distribution [14], arctan power distribution [31], log-cosine-power unit distribution [33], unit Mirra distribution [4], and unit log-log distribution [21]. Also, [34] developed the unit ratio-extended Weibull family of distributions and derived several special distributions.

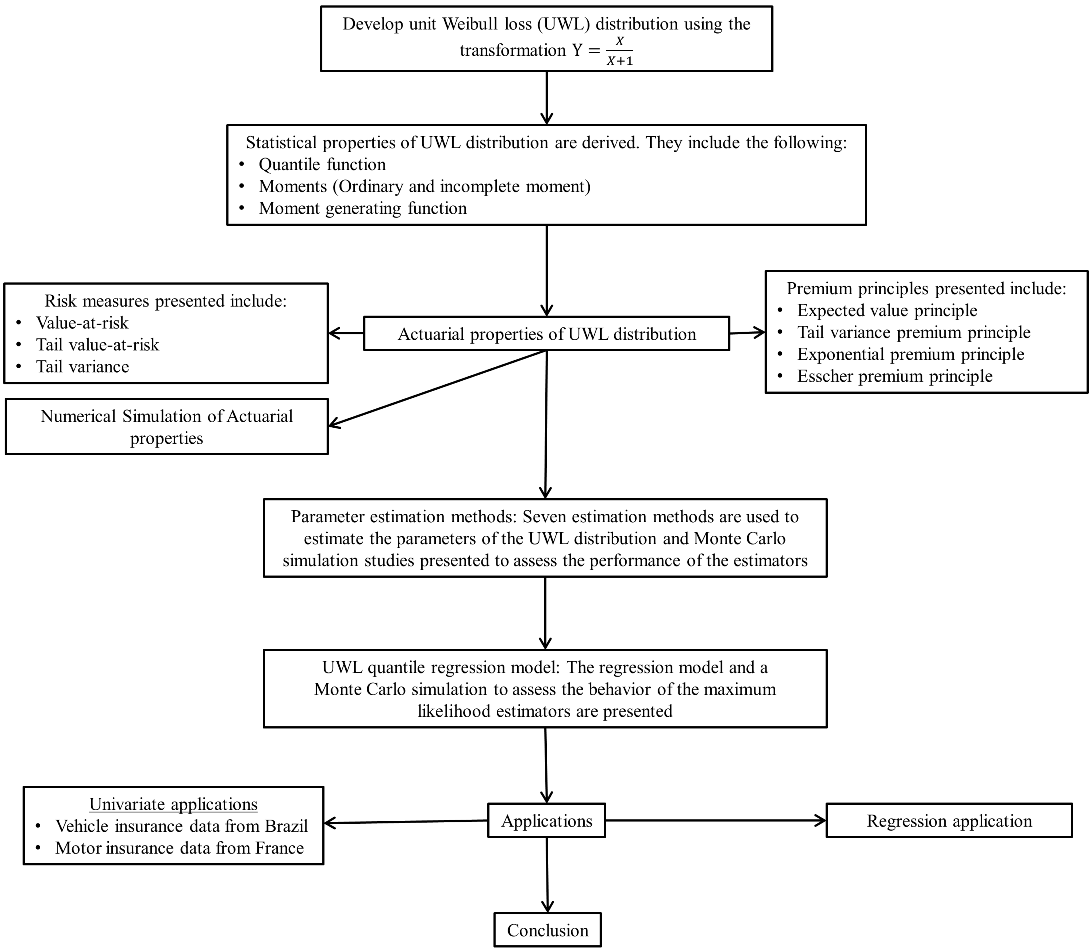

Figure 1. Diagram on the framework of the study

In this study, a new unit distribution called the unit Weibull loss (UWL) distribution is elaborated based on the Weibull loss distribution [3] with an emphasis on actuarial properties and applications. Our motivation for the formulation of this distribution stems from the fact that it can exhibit different desirable shapes, which makes it a competitive solution for analyzing data defined on the unit interval with the least loss of information. In addition, the tractable nature of the corresponding quantile function makes it easier to generate random observations from it and formulate a quantile regression model. Thus, the main aim of this study is to provide a distribution on the unit interval that possesses desirable properties and can also serve as an alternative to other unit distributions, especially for analyzing actuarial data.

The remainder of the article is organized as follows: Section 2 describes the UWL distribution. Some of its statistical measures and properties, including the quantile function and moments, are given in Section 3. Actuarial properties, including risk measures and premium principles, are determined in Section 4. Section 5 presents seven methods for estimating the parameters of the distribution. Also, simulation studies are performed to assess the performance of the estimates in this section. Section 6 focuses on a quantile regression model. Applications of the unit distribution and its regression model are shown in Section 7. Section 8 gives the conclusion of the study. An illustration of the study framework is offered in Figure 1 for a direct and schematic view.

2. Unit Weibull loss distribution

The UWL distribution is derived in this section using the Weibull loss (WL) distribution proposed by [3]. To comprehend its construction, we recall that a random variable X that follows the WL distribution has the following cumulative distribution function (CDF):

$$F_X(x) = 1 - \frac{\alpha \mathrm{e}^{-\beta x^{\lambda}}}{\alpha + \beta x^{\lambda}}, \quad x \geq 0,\ \alpha > 0,\ \beta > 0 \tag{1}$$

Now, let us consider the following transformation of X: $Y = \dfrac{X}{X+1}$. Then the distribution of Y corresponds to the UWL distribution. It is defined with the CDF given as

$$F_Y(y) = 1 - \frac{\alpha \mathrm{e}^{-\beta\left(\frac{y}{1-y}\right)^{\lambda}}}{\alpha + \beta\left(\dfrac{y}{1-y}\right)^{\lambda}}, \quad y \in (0,1),\ \alpha > 0,\ \beta > 0 \tag{2}$$

The probability density function (PDF) is obtained by differentiating Equation (2) with respect to y. This is given as

$$f_Y(y) = \alpha\beta\lambda\left(\frac{y}{1-y}\right)^{\lambda-1}\left(\frac{1}{1-y}\right)^{2} \mathrm{e}^{-\beta\left(\frac{y}{1-y}\right)^{\lambda}} \frac{\left[\alpha + 1 + \beta\left(\dfrac{y}{1-y}\right)^{\lambda}\right]}{\left[\alpha + \beta\left(\dfrac{y}{1-y}\right)^{\lambda}\right]^{2}}, \quad y \in (0,1) \tag{3}$$

The corresponding survival and hazard rate functions are given, respectively, as

$$S_Y(y) = \frac{\alpha e^{-\beta\left(\frac{y}{1-y}\right)^{\lambda}}}{\alpha + \beta\left(\dfrac{y}{1-y}\right)^{\lambda}}, \quad y \in (0,1) \tag{4}$$

and

$$\tau_Y(y) = \beta\lambda\left(\frac{y}{1-y}\right)^{\lambda-1}\left(\frac{1}{1-y}\right)^{2}\frac{\left[\alpha+1+\beta\left(\dfrac{y}{1-y}\right)^{\lambda}\right]}{\left[\alpha+\beta\left(\dfrac{y}{1-y}\right)^{\lambda}\right]}, \quad y \in (0,1) \tag{5}$$

To comprehend the modeling capabilities of the UWL distribution, let us perform a graphical analysis. Figure 2 shows some shapes of the corresponding PDF and hazard rate function for some parameter values. It can be observed that the PDF can have decreasing, left-skewed, right-skewed, and approximately symmetric shapes. The hazard rate function exhibits J, bathtub, and modified bathtub shapes. These observations validate the remarkable flexibility inherent in the UWL distribution.

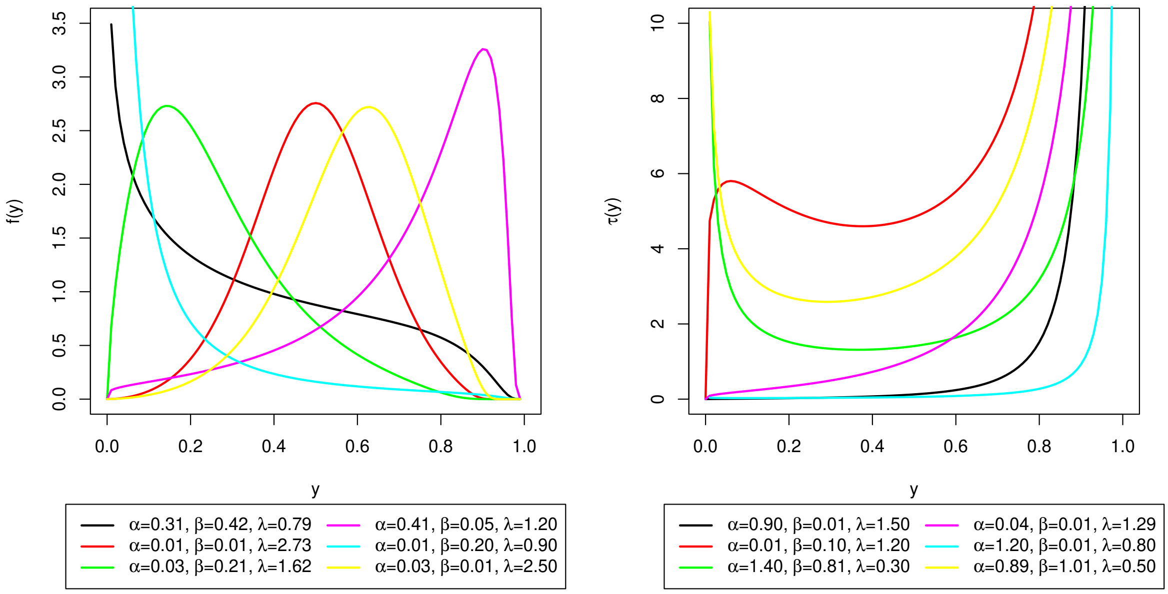

Figure 2. PDF plots (left) and hazard rate function plots (right)

3. Statistical properties

Some statistical properties of the UWL distribution are presented in this section. These include the quantile function, ordinary moments, incomplete moments, and moment generating function.

3.1. Quantile function

The quantile function of the UWL distribution is useful for the generation of random numbers and some characterization of the distribution. It is obtained as the inverse function of the CDF of the UWL distribution given in Equation (2). Let us determine this function, step by step. Equating the CDF of the UWL distribution to $u \in (0,1)$ and simplifying yields

$$1 - \frac{\alpha \mathrm{e}^{-\beta\left(\frac{y}{1-y}\right)^{\lambda}}}{\alpha + \beta\left(\frac{y}{1-y}\right)^{\lambda}} = u$$

so that

$$\alpha \mathrm{e}^{-\beta\left(\frac{y}{1-y}\right)^{\lambda}} + (1-u)\log \mathrm{e}^{-\beta\left(\frac{y}{1-y}\right)^{\lambda}} = \alpha(1-u)$$

Letting $z = \mathrm{e}^{-\beta\left(\frac{y}{1-y}\right)^{\lambda}}$, we obtain

$$\alpha z + (1-u)\log z = \alpha(1-u)$$

Dividing by $(1-u)$ gives

$$\left(\frac{\alpha}{1-u}\right) z + \log z = \alpha \quad \text{thus} \quad \log \mathrm{e}^{\left(\frac{\alpha}{1-u}\right)z} + \log z = \alpha$$

Simplifying further, we obtain

$$z \mathrm{e}^{\left(\frac{\alpha}{1-u}\right)z} = \mathrm{e}^{\alpha}$$

Using the Lambert function $W(x)$ defined as $W(x)\mathrm{e}^{W(x)} = x$, we have

$$z = \left(\frac{1-u}{\alpha}\right) W\left[\left(\frac{\alpha}{1-u}\right)\mathrm{e}^{\alpha}\right]$$

Noting that $z = \mathrm{e}^{-\beta\left(\frac{y}{1-y}\right)^{\lambda}}$ and making y the subject gives the quantile function of the UWL distribution as

$$Q_Y(u) = \left\{1 + \left[-\frac{1}{\beta}\log\left\{\left(\frac{1-u}{\alpha}\right) W\left[\left(\frac{\alpha}{1-u}\right)\mathrm{e}^{\alpha}\right]\right\}\right]^{-\frac{1}{\lambda}}\right\}\mathrm{e}^{-1}, \quad u \in (0,1) \tag{6}$$

The skewness and kurtosis of the UWL distribution can be obtained via its quantile function. This can be achieved using Moor's kurtosis and Bowley's skewness, respectively, indicated as

$$MK = \frac{Q_Y\left(\frac{7}{8}\right) - Q_Y\left(\frac{5}{8}\right) + Q_Y\left(\frac{3}{8}\right) - Q_Y\left(\frac{1}{8}\right)}{Q_Y\left(\frac{6}{8}\right) - Q_Y\left(\frac{2}{8}\right)}$$

and

$$BS = \frac{Q_Y\left(\frac{3}{4}\right) + Q_Y\left(\frac{2}{4}\right) - 2Q_Y\left(\frac{2}{4}\right)}{Q_Y\left(\frac{3}{4}\right) - Q_Y\left(\frac{1}{4}\right)}$$

We can evaluate them numerically or graphically. To this end, Figure 3 displays Moor's kurtosis and Bowley's skewness plots for $\alpha = 0.5$ and different values of β and λ. It is seen that the UWL distribution can exhibit both right- and left-skewed shapes and different degrees of kurtosis.

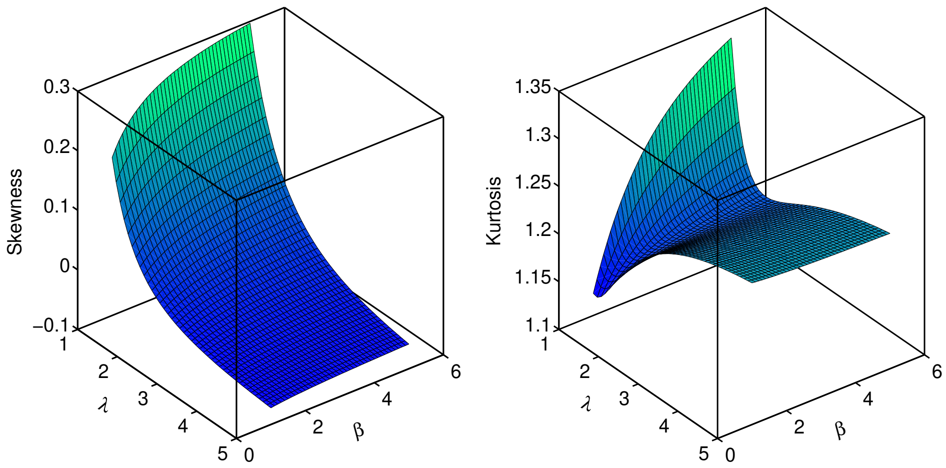

Figure 3. Skewness plot (left) and kurtosis plot (right)

3.2. Moments

The characterization of a distribution through the moments of random variables is a well-established approach. In this subsection, we present the ordinary moment of the UWL distribution as a key element in understanding its statistical properties.

3.2.1. Ordinary moment

For any positive integer r, the rth ordinary moment of a random variable Y with support on the unit interval, PDF $f_Y(y)$, and CDF $F_Y(y)$, can be expressed as

$$\mu_r' = E(Y^r) = \int_0^1 y^r f_Y(y)dy = -\int_0^1 y^r d\left(1 - F_Y(y)\right) = r\int_0^1 y^{r-1}\left(1 - F_Y(y)\right)dy \tag{7}$$

Thus, the rth ordinary moment associated with the UWL distribution is obtained by substituting its survival function in Equation (4) into the definition in Equation (7). This gives

$$\mu_r' = \alpha r \int_0^1 y^{r-1} \frac{\mathrm{e}^{-\beta\left(\frac{y}{1-y}\right)^\lambda}}{\alpha + \beta\left(\dfrac{y}{1-y}\right)^\lambda} dy \tag{8}$$

The integral in Equation (8) can be determined numerically using various software such as R, Python, Mathematica, and Matlab.

The mean of Y, denoted by μ, is obtained by putting $r = 1$ into Equation (8). Various central moments, defined as $E[(Y-\mu)^r]$ can be obtained using the non-central moment. For $r = 2, 3$ and 4, they are given as $E[(Y-\mu)^2] = \sigma^2 = \mu_2' - \mu^2$, $E[(Y-\mu)^3] = \mu_3' - 3\mu_2'\mu + 2\mu^3$ and $E[(Y-\mu)^4] = \mu_4' - 4\mu_3'\mu + 6\mu_2'\mu^2 - 3\mu^4$, respectively. These moments are useful for obtaining the coefficients of skewness (CS) and kurtosis (CK) of the distribution, given, respectively, as

$$CS = \frac{E[(Y-\mu)^3]}{\sigma^3} \quad \text{and} \quad CK = \frac{E[(Y-\mu)^4]}{\sigma^4}$$

As a numerical indicator, Table 1 presents the first four ordinary moments, the standard deviation (SD), CS, and CK of the UWL distribution. It is noticed that the distribution can exhibit both right and left skewness and various degrees of kurtosis.

Table 1. Moments, SD, CS and CK of the UWL distribution

μ'_r	$\alpha = 0.03, \beta = 4.5, \lambda = 2.8$	$\alpha = 1.1, \beta = 2.45, \lambda = 2.5$	$\alpha = 1.5, \beta = 1.8, \lambda = 2.1$
μ'_1	0.15245	0.31889	0.33215
μ'_2	0.02856	0.11206	0.12447
μ'_3	0.00633	0.04213	0.05047
μ'_4	0.00161	0.01668	0.02168
SD	0.07292	0.10183	0.11895
CS	0.90765	-0.20339	-0.15992
CK	3.97506	2.62565	2.50537

3.3. Moment generating function

The moment generating function (MGF) of Y is defined as $M_Y(t) = E(\mathrm{e}^{tY})$ and it can be presented using Taylor series expansion as

$$M_Y(t) = \sum_{r=0}^{\infty} \frac{t^r}{r!}\mu'_r$$

where μ'_r is the associated rth ordinary moment. Substituting the rth ordinary moment associated with the UWL distribution in Equation (8) into the definition gives

$$M_Y(t) = \sum_{r=1}^{\infty} \frac{\alpha t^r}{(r-1)!} \int_0^1 y^{r-1} \frac{\mathrm{e}^{-\beta\left(\frac{y}{1-y}\right)^\lambda}}{\alpha + \beta\left(\frac{y}{1-y}\right)^\lambda} dy \tag{9}$$

provided it converges in the series sense.

4. Actuarial properties

In this section, several actuarial properties are derived with respect to the UWL distribution. These properties include risk measures and premium principles. For each of these notions, the related general formula is recalled by taking into account a random variable Y with PDF $f_Y(y)$, CDF $F_Y(y)$, and quantile function $Q_Y(y)$.

4.1. Risk measures

Risk measures are used to quantify risks. Among the most useful are the value-at-risk, tail value-at-risk, and tail variance. These risk measures are presented in this section when the loss distribution follows the UWL distribution.

4.1.1. Value-at-risk

Given a probability p, the value-at-risk (VaR) measures losses that will not be exceeded. VaR can be defined as the quantile function of a distribution, that is, $VaR_p = Q_Y(p)$. Thus, for the UWL distribution, owing to Equation (6), the quantile function is defined as

$$\text{Var}_p = \left\{1 + \left[-\frac{1}{\beta}\log\left\{\left(\frac{1-p}{\alpha}\right)W\left[\left(\frac{\alpha}{1-p}\right)\mathrm{e}^{\alpha}\right]\right\}\right]^{-\frac{1}{\lambda}}\right\}\mathrm{e}^{-1}, \quad p \in (0,1) \tag{10}$$

4.1.2. Tail value-at-risk

The tail value-at-risk ($TVaR$) measures losses above VaR. Given that the probability of losses being less than or equal to VaR is p, then the $TVaR$ measures the expected value of $1-p$ of the losses, which is the worst case of the loss. The $TVaR$ is therefore defined as

$$TVaR = \frac{1}{1-p}\int_{VaR_p}^{1} y f_Y(y)dy \tag{11}$$

Substituting the PDF of the UWL distribution in Equation (3) into Equation (11) gives

$$TVaR = \frac{\alpha\beta\lambda}{1-p}\int_{VaR_p}^{1} y\left(\frac{y}{1-y}\right)^{\lambda-1}\left(\frac{1}{1-y}\right)^{2}\mathrm{e}^{-\beta\left(\frac{y}{1-y}\right)^{\lambda}}\frac{\left[\alpha+1+\beta\left(\frac{y}{1-y}\right)^{\lambda}\right]}{\left[\alpha+\beta\left(\frac{y}{1-y}\right)^{\lambda}\right]^{2}}dy \tag{12}$$

4.1.3. Tail variance

The tail variance (TV) measures the variability of losses beyond the VaR. This is very important for the overall measure of the risk of a company. The TV is defined as

$$TV = E(Y^2 \mid Y > VaR_p) - (TVaR_p)^2 = \frac{1}{1-p}\int_{VaR_p}^{1} y^2 f_Y(y)dy - (TVaR_p)^2 \tag{13}$$

Again, substituting the PDF of the UWL distribution in Equation (3) and $TVaR$ in Equation (12) into Equation (13) gives

$$TV = \frac{\alpha\beta\lambda}{1-p}\int\limits_{VaR_p}^{1}\left(\frac{y}{1-y}\right)^{\lambda-1}\left(\frac{y}{1-y}\right)^{2}\mathrm{e}^{-\beta\left(\frac{y}{1-y}\right)^{\lambda}}\frac{\left[\alpha+1+\beta\left(\frac{y}{1-y}\right)^{\lambda}\right]}{\left[\alpha+\beta\left(\frac{y}{1-y}\right)^{\lambda}\right]^{2}}dy$$
$$-\left(\frac{\alpha\beta\lambda}{1-p}\int\limits_{VaR_p}^{1}y\left(\frac{y}{1-y}\right)^{\lambda-1}\left(\frac{1}{1-y}\right)^{2}\mathrm{e}^{-\beta\left(\frac{y}{1-y}\right)^{\lambda}}\frac{\left[\alpha+1+\beta\left(\frac{y}{1-y}\right)^{\lambda}\right]}{\left[\alpha+\beta\left(\frac{y}{1-y}\right)^{\lambda}\right]^{2}}dy\right)^{2} \tag{14}$$

4.2. Premium principles

Premium principles are used to obtain premiums for insurance coverage for events, taking into consideration the level of risk of the event. Several premium principles have been developed over the decades. Some of them are presented in this subsection with a loss distribution following the UWL distribution. In this subsection, let ρ denote the non-negative risk loading parameter.

4.2.1. Expected value principle

The expected value principle (EVP) is defined as

$$\mathrm{EVP} = (1+\rho)E(Y)$$

Therefore, substituting the expected value of the UWL distribution, obtained by letting $r = 1$ in Equation (8), into the definition yields the EVP of the distribution as

$$\mathrm{EVP} = (1+\rho)\alpha\int\limits_{0}^{1}\frac{\mathrm{e}^{-\beta\left(\frac{y}{1-y}\right)^{\lambda}}}{\alpha+\beta\left(\frac{y}{1-y}\right)^{\lambda}}dy$$

That is, the EVP defines the premium as being proportional to the expected value of the loss. When $\rho = 0$, the simplest premium principle, known as the equivalence premium principle, is obtained.

4.2.2. Tail variance premium principle

The tail variance premium (TVP) principle was developed to take into consideration tail losses. The TVP was proposed by [12] and is defined as

$$TVP = TVaR + \rho TV$$

In the context of the UWL distribution, $TVaR$ and TV are given by equations (12) and (14), respectively.

4.2.3. Exponential premium principle

The exponential principle (ExP) is obtained by solving for ExP in the equation $u(w-\mathrm{ExP}) = E(w-Y)$, where w represents the wealth of an individual and $u(y) = -\mathrm{e}^{-\rho y}$ is the exponential utility function. Therefore, the ExP is obtained as

$$\mathrm{ExP} = \frac{1}{\rho} M_Y(\rho),$$

where $M_Y(\rho)$ is the MGF of Y. Substituting the MGF of the UWL distribution given in Equation (9) into the definition gives

$$\mathrm{ExP} = \frac{\alpha}{\rho} \sum_{r=0}^{\infty} \frac{\rho^r}{(r-1)!} \int_0^1 y^{r-1} \frac{\mathrm{e}^{-\beta\left(\frac{y}{1-y}\right)^\lambda}}{\alpha + \beta\left(\frac{y}{1-y}\right)^\lambda} dy.$$

4.2.4. Esscher premium principle

The Esscher principle (EsP) is defined as

$$\mathrm{EsP} = \frac{E(Y\mathrm{e}^{\rho Y})}{M_Y(\rho)}$$

where $M_Y(\rho)$ is the MGF of Y. The EsP associated with the UWL distribution is specified as

$$\mathrm{EsP} = \frac{\alpha\beta\lambda \displaystyle\int_0^1 \mathrm{e}^{\rho Y} \left(\frac{y}{1-y}\right)^{\lambda-1} \left(\frac{1}{1-y}\right)^2 \mathrm{e}^{-\beta\left(\frac{y}{1-y}\right)^\lambda} \frac{\left[\alpha + 1 + \beta\left(\frac{y}{1-y}\right)^\lambda\right]}{\left[\alpha + \beta\left(\frac{y}{1-y}\right)^\lambda\right]^2} dy}{\displaystyle\sum_{r=1}^{\infty} \frac{\alpha\rho^r}{(r-1)!} \int_0^1 y^{r-1} \frac{\mathrm{e}^{-\beta\left(\frac{y}{1-y}\right)^\lambda}}{\alpha + \beta\left(\frac{y}{1-y}\right)^\lambda} dy}.$$

4.3. Numerical simulation of measures

This section presents a Monte Carlo simulation of the risk measures and premium principles. The following process is used to carry out the simulation:

1. Using the quantile function of the UWL distribution given in Equation (6), generate a random sample of size 100.
2. Estimate the parameters of the distribution using the maximum likelihood (ML) method.
3. Repeat steps 1 and 2 2000 times. Compute the risk measures VaR, $TVaR$, and TV and the premium principles EVP, TVP, ExP, and EsP for each repetition.
4. Repeat steps 1 to 3 for the confidence levels $p = (0.50, 0.55, 0.60, \ldots, 0.999)$ for the risk measures and risk loading $\rho = (0.3, 0.5, 0.7, 1.0, 1.5, 2.0, 2.5)$ for the premium principles.
5. Repeat steps 1 to 4 for the three sets of parameter values: $(\alpha, \beta, \lambda) = (1.1, 0.8, 0.5)$, $(\alpha, \beta, \lambda) = (1.5, 1.2, 0.8)$ and $(\alpha, \beta, \lambda) = (2.1, 1.8, 0.3)$.

Figure 4 exposes the plots of the risk measures. It can be remarked that as the confidence level increases, VaR and $TVaR$ increase while TV decreases. Hence, it is expected that more capital should be set aside to be able to offset losses with a higher level of confidence. Also, the risk associated with the losses, measured by TV, is expected to decrease with a higher level of confidence.

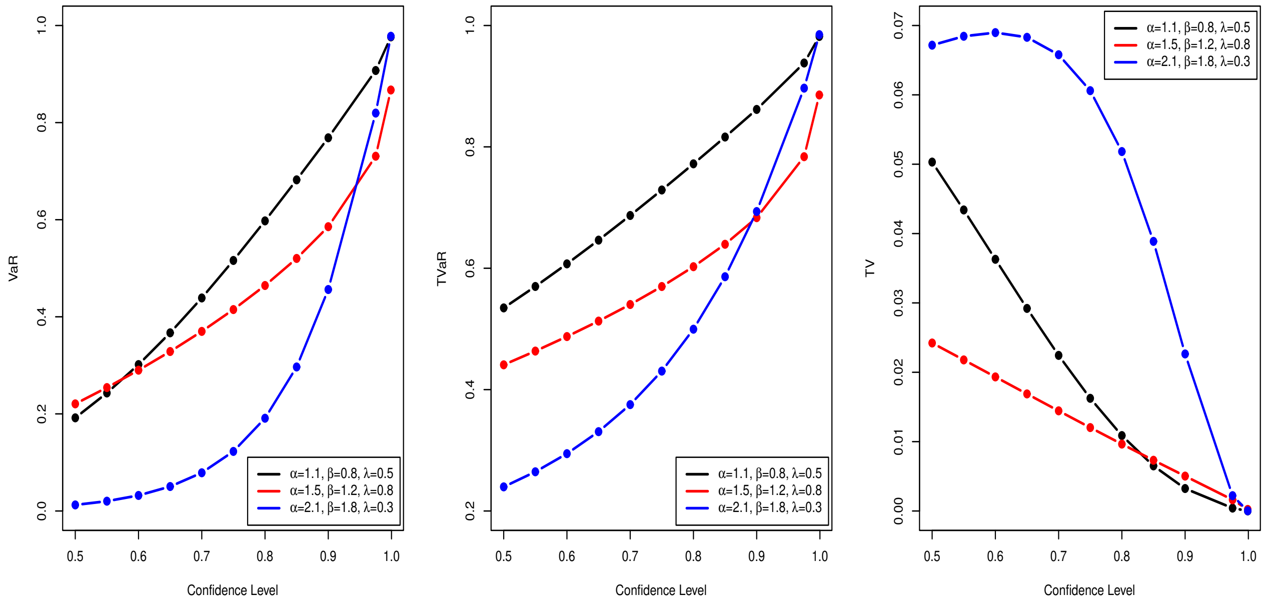

Figure 4. Plots of the risk measures

Figure 5 shows plots of the premium principles. It can be noticed that they are all increasing with an increase in the risk associated with the loss. This is also a desirable characteristic of loss distributions.

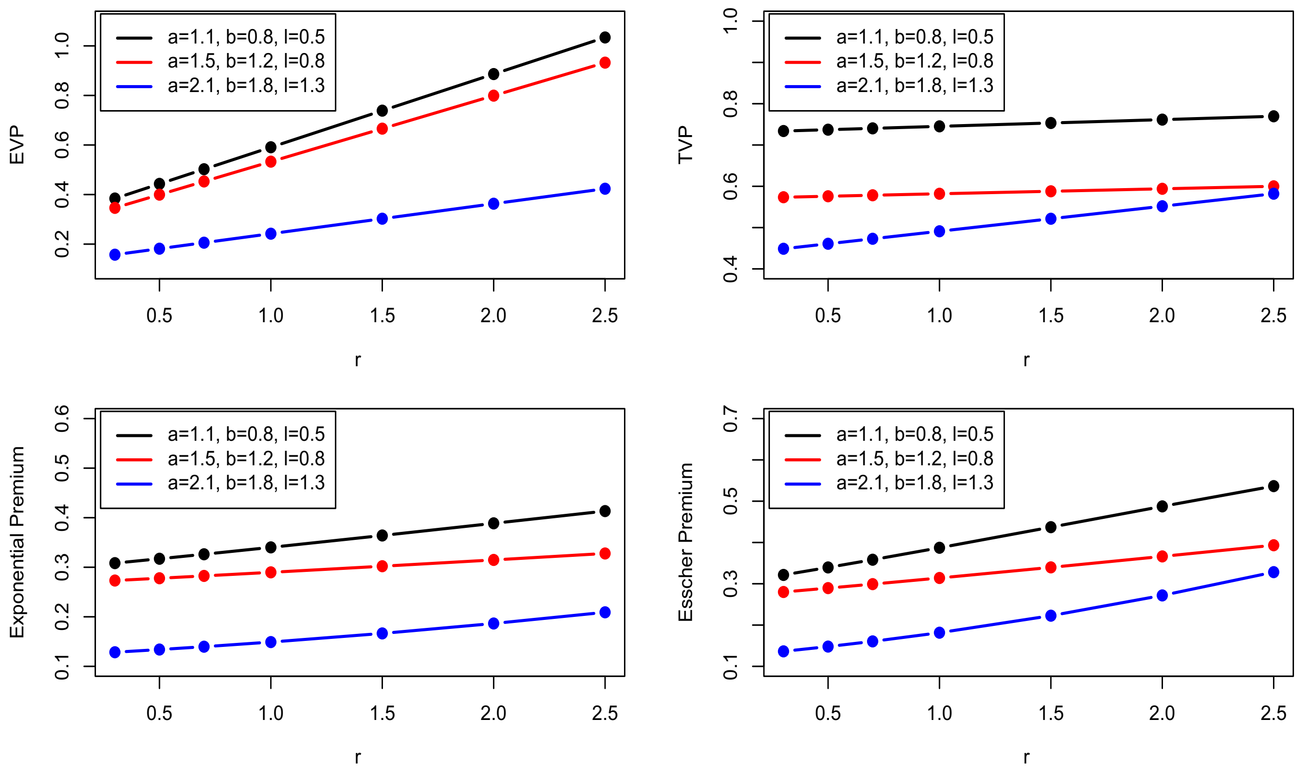

Figure 5. Plots of the premium principles

5. Parameter estimation methods

This section presents seven methods for estimating the parameters of the UWL distribution, denoted $\Theta = (\alpha, \beta, \lambda)'$. These methods are the ML, maximum product spacing (MPS) [9, 35], ordinary least squares (OLS) and weighted least squares (WLS) [36], Anderson–Darling (AD) [6], Cramér–von Mises (CVM) [24] and percentile (PC) [16] estimation methods.

5.1. Maximum likelihood method

Suppose $y_1, y_2, \ldots, y_n$ are values from a random variable Y that follows the UWL distribution. The set of parameters that maximize the log-likelihood function defined as $\ell = \log\left[\prod_{i=1}^{n} f_Y(y_i)\right]$ gives the ML estimates of the parameters. To this end, we consider the log-likelihood function of the UWL distribution, which is given as

$$\ell(\Theta) = n\log(\alpha\beta\lambda) - \beta\sum_{i=1}^{n}\left(\frac{y_i}{1-y_i}\right)^{\lambda} + (\lambda-1)\sum_{i=1}^{n}\log\left(\frac{y_i}{1-y_i}\right) - 2\sum_{i=1}^{n}\log(1-y_i)$$
$$+\sum_{i=1}^{n}\log\left[\alpha+1+\beta\left(\frac{y_i}{1-y_i}\right)^{\lambda}\right] - 2\sum_{i=1}^{n}\log\left[\alpha+\beta\left(\frac{y_i}{1-y_i}\right)^{\lambda}\right] \tag{15}$$

5.2. Maximum product spacing method

Let $y_{(1)}, y_{(2)}, \ldots, y_{(n)}$ be the ordered values $y_1, y_2, \ldots, y_n$ in increasing order. For any $i = 1, \ldots, n+1$, let D_i be the uniform spacing defined as

$$D_i = F_Y(y_{(i)}) - F_Y(y_{(i-1)})$$

where $F_Y(y_{(0)}) = 0$, $F_Y(y_{(n+1)}) = 1$ and $\sum_{i=0}^{n+1} D_i = 1$. For the UWL distribution, the uniform spacing is obtained as

$$D_i = \frac{\alpha\left(1-y_{(i)}\right)^{\lambda} e^{-\beta\left(\frac{y_{(i)}}{1-y_{(i)}}\right)^{\lambda}}}{\alpha\left(1-y_{(i)}\right)^{\lambda} + \beta y_{(i)}^{\lambda}} - \frac{\alpha\left(1-y_{(i-1)}\right)^{\lambda} e^{-\beta\left(\frac{y_{(i-1)}}{1-y_{(i-1)}}\right)^{\lambda}}}{\alpha\left(1-y_{(i-1)}\right)^{\lambda} + \beta y_{(i-1)}^{\lambda}}$$

The MPS parameter estimates are obtained by maximizing the function

$$G(\Theta) = \frac{1}{n+1}\sum_{i=1}^{n+1}\log(D_i)$$

with respect to the parameters.

5.3. Ordinary and weighted least squares method

The OLS and WLS estimates of the parameters of the UWL distribution are obtained by minimizing the following functions, respectively:

$$L(\Theta)=\sum_{i=1}^{n}\left[1-\frac{\alpha\left(1-y_{(i)}\right)^{\lambda}\mathrm{e}^{-\beta\left(\frac{y_{(i)}}{1-y_{(i)}}\right)^{\lambda}}}{\alpha\left(1-y_{(i)}\right)^{\lambda}+\beta y_{(i)}^{\lambda}}-\frac{i}{n+1}\right]^{2}$$

and

$$W(\Theta)=\sum_{i=1}^{n}\frac{(n+1)^{2}(n+2)}{i(n+1-i)}\left[1-\frac{\alpha\left(1-y_{(i)}\right)^{\lambda}\mathrm{e}^{-\beta\left(\frac{y_{(i)}}{1-y_{(i)}}\right)^{\lambda}}}{\alpha\left(1-y_{(i)}\right)^{\lambda}+\beta y_{(i)}^{\lambda}}-\frac{i}{n+1}\right]^{2}$$

5.4. Anderson–Darling method

The AD estimates of the parameters of the UWL distribution are obtained by minimizing, with respect to the parameters, the following function:

$$A(\Theta)=-n-\frac{1}{n}\sum_{i=1}^{n}(2i-1)\left[\log(\alpha)+\log\left(1-\frac{\alpha\left(1-y_{(i)}\right)^{\lambda}\mathrm{e}^{-\beta\left(\frac{y_{(i)}}{1-y_{(i)}}\right)^{\lambda}}}{\alpha\left(1-y_{(i)}\right)^{\lambda}+\beta y_{(i)}^{\lambda}}\right)\right.$$

$$\left.-\beta\left(\frac{y_{(n+1-i)}}{1-y_{(n+1-i)}}\right)^{\lambda}-\log\left(\alpha+\beta\left(\frac{y_{(n+1-i)}}{1-y_{(n+1-i)}}\right)^{\lambda}\right)\right]$$

5.5. Cramér–von Mises method

The parameter estimates of the UWL distribution, via the CVM method, are obtained by minimizing the difference between the empirical and estimated CDFs. Thus, minimizing the following function, the estimates of the UWL distribution are determined:

$$C(\Theta)=\frac{1}{12n}+\sum_{i=1}^{n}\left[1-\frac{\alpha\left(1-y_{(i)}\right)^{\lambda}\mathrm{e}^{-\beta\left(\frac{y_{(i)}}{1-y_{(i)}}\right)^{\lambda}}}{\alpha\left(1-y_{(i)}\right)^{\lambda}+\beta y_{(i)}^{\lambda}}-\frac{2i-1}{2n}\right]^{2}$$

5.6. Percentile method

The PC estimates of the parameters of the UWL distribution are obtained by minimizing the function

$$P(\Theta)=\sum_{i=1}^{n}\left[y_{(i)}-\left\{1+\left[-\frac{1}{\beta}\log\left\{\left(\frac{1-q_i}{\alpha}\right)W\left[\left(\frac{\alpha}{1-q_i}\right)\mathrm{e}^{\alpha}\right]\right\}\right]^{\frac{1}{\lambda}}\right\}\mathrm{e}^{-1}\right]^{2}$$

with respect to the parameters, where $q_i = i/(n+1)$ is an unbiased estimator of $F_Y(y_{(i)})$.

5.7. Monte Carlo simulations

The estimates established for the parameters of the UWL distribution are assessed using a Monte Carlo simulation study. The process used for this study is given as follows:

1. Using the quantile function of the UWL distribution given in Equation (6), a sample of size n is generated.
2. The estimates of the parameters $\hat{\Theta} = (\hat{\alpha}, \hat{\beta}, \hat{\lambda})$ are obtained using the ML, MPS, OLS, WLS, CVM, AD, and PC estimation methods.
3. Steps 1 to 2 are repeated $N = 3000$ times.
4. The average estimate (AE), absolute bias (AB) and the root mean square error (RMSE) of the estimates are computed.
5. Steps 1 to 4 are repeated for the sample sizes $n = 25, 75, 150, 300, 700, 1000$ and the parameter sets $(\alpha, \beta, \lambda) = (1.2, 0.2, 0.3)$ and $(\alpha, \beta, \lambda) = (1.8, 0.8, 1.1)$.

Tables 2 and 3 present the results of the Monte Carlo simulation. All the estimates from the methods can be observed to be consistent, as AE tends to approach the true values, while AB and RMSE tend to decrease with increasing sample sizes. It can further be noticed that ML method performed better than the other estimation methods. Hence, this method is therefore considered for the estimation of the parameters of the distribution.

The performance of the methods is ranked to determine the best of them. Only the AB and RMSE are used for the ranking. Tables 4 and 5 present the results. It can be observed that the ML method has the lowest sum of ranks and has the first position for both simulation results. Therefore, it is the best estimation method for estimating the parameters of the UWL distribution. Suppose $y_1, y_2, \ldots, y_n$ are values from a random variable Y that follows the UWL distribution. Then, the regression structure is defined as

$$h(\eta_i) = \boldsymbol{x}_i' \boldsymbol{\delta}, \quad i = 1, 2, \ldots,$$

where $\boldsymbol{x}_i = (1, x_{i1}, \ldots, x_{in})'$ is the vector of independent variables and $\boldsymbol{\delta} = (\delta_0, \delta_1, \ldots, \delta_p)'$ is a vector of the parameters of the model. In this notation, $h(\eta_i)$ is used to connect the independent variables to the response variable, and it is known as the link function.

6. Unit Weibull loss quantile regression model

A quantile regression model with a random response variable following the UWL distribution is introduced in this section. Firstly, we make α the subject in the quantile function of the UWL distribution and substitute it into its PDF in equation (3) to obtain a re-parameterized PDF of the distribution. It is determined as

$$f_Y(y) = \frac{\lambda(1-u)\left(\frac{\eta}{1-\eta}\right)^{\lambda}\left\{(1-u)\left[\beta\left(\frac{\eta}{1-\eta}\right)^{\lambda} - \beta\left(\frac{y}{1-y}\right)^{\lambda} - 1\right] + \left[\beta\left(\frac{y}{1-y}\right)^{\lambda} + 1\right]\mathrm{e}^{-\beta\left(\frac{\eta}{1-\eta}\right)^{\lambda}}\right\}}{(1-y)^2\left(\frac{y}{1-y}\right)^{1-\lambda}\mathrm{e}^{\beta\left(\frac{y}{1-y}\right)^{\lambda}}\left\{(1-u)\left[\left(\frac{\eta}{1-\eta}\right)^{\lambda} - \left(\frac{y}{1-y}\right)^{\lambda}\right] + \left(\frac{y}{1-y}\right)^{\lambda}\mathrm{e}^{-\beta\left(\frac{\eta}{1-\eta}\right)^{\lambda}}\right\}\mathrm{e}^2} \quad (16)$$

where $y \in (0,1), \eta \in (0,1), u \in (0,1), \beta > 0$ and $\lambda > 0$.

Table 2. Parameter estimates of the UWL distribution for $(\alpha, \beta, \lambda) = (1.2, 0.2, 0.3)$

Measure	Parameter	n	ML	MPS	OLS	WLS	AD	CVM	PC
AE	α	25	2.0282	3.1338	3.1508	3.2541	3.0869	3.0767	4.2952
		75	1.7899	2.7245	2.7004	2.6675	2.6025	2.6600	3.7489
		150	1.6495	2.4173	2.3983	2.3954	2.3196	2.3596	3.4291
		300	1.6500	2.3339	2.5460	2.4257	2.3844	2.5275	3.1663
		700	1.5556	1.9553	2.0905	1.9562	1.9483	2.0791	2.8048
		1000	1.5744	1.9521	2.0536	1.9801	1.9616	2.0445	2.7724
	β	25	0.2064	0.2016	0.2242	0.2317	0.2233	0.2108	0.2684
		75	0.1954	0.1993	0.2020	0.2078	0.2062	0.1968	0.2355
		150	0.1914	0.1969	0.1949	0.2006	0.1993	0.1917	0.2205
		300	0.1978	0.2045	0.2063	0.2094	0.2082	0.2046	0.2137
		700	0.2010	0.2052	0.2046	0.2066	0.2064	0.2037	0.2005
		1000	0.2042	0.2086	0.2084	0.2101	0.2094	0.2077	0.2019
	λ	25	0.3265	0.3586	0.3072	0.3040	0.3131	0.3286	0.2663
		75	0.3166	0.3288	0.3161	0.3111	0.3127	0.3233	0.2896
		150	0.3137	0.3206	0.3155	0.3117	0.3122	0.3192	0.3036
		300	0.3070	0.3105	0.3077	0.3052	0.3057	0.3095	0.3023
		700	0.3032	0.3048	0.3041	0.3025	0.3026	0.3049	0.3059
		1000	0.3020	0.3030	0.3022	0.3011	0.3014	0.3027	0.3051
AB	α	25	1.4192	2.5703	2.5948	2.6321	2.5069	2.5613	3.3422
		75	1.2476	2.2094	2.2290	2.1401	2.0784	2.2091	2.9958
		150	1.0940	1.8829	1.9147	1.8610	1.7855	1.8902	2.7795
		300	0.9971	1.6897	1.9574	1.7712	1.7323	1.9457	2.5784
		700	0.7838	1.1905	1.4020	1.2020	1.1933	1.3959	2.3290
		1000	0.7574	1.1406	1.2871	1.1641	1.1527	1.2816	2.2703
	β	25	0.0923	0.1018	0.1175	0.1143	0.1122	0.1147	0.1080
		75	0.0800	0.0889	0.1013	0.0947	0.0920	0.1001	0.0961
		150	0.0705	0.0811	0.0884	0.0845	0.0830	0.0876	0.0950
		300	0.0579	0.0678	0.0792	0.0717	0.0710	0.0788	0.0931
		700	0.0446	0.0502	0.0599	0.0521	0.0518	0.0598	0.0946
		1000	0.0413	0.0472	0.0525	0.0474	0.0474	0.0523	0.0922
	λ	25	0.0545	0.0722	0.0570	0.0524	0.0542	0.0633	0.0747
		75	0.0346	0.0397	0.0420	0.0364	0.0363	0.0442	0.0469
		150	0.0282	0.0319	0.0355	0.0314	0.0312	0.0366	0.0388
		300	0.0192	0.0210	0.0243	0.0215	0.0214	0.0245	0.0287
		700	0.0130	0.0140	0.0159	0.0138	0.0138	0.0160	0.0215
		1000	0.0114	0.0123	0.0131	0.0120	0.0121	0.0131	0.0193
RMSE	α	25	1.5094	2.9626	2.9640	3.0025	2.9093	2.9385	3.5088
		75	1.3772	2.6972	2.6871	2.6161	2.5648	2.6715	3.2700
		150	1.2632	2.4324	2.4255	2.4015	2.3181	2.4020	3.1101
		300	1.1867	2.2712	2.4776	2.3330	2.2990	2.4679	2.9665
		700	1.0070	1.7815	1.9866	1.7617	1.7520	1.9805	2.7721
		1000	0.9820	1.7279	1.8901	1.7623	1.7428	1.8843	2.7236
	β	25	0.1122	0.1223	0.1402	0.1371	0.1337	0.1368	0.1330
		75	0.0947	0.1024	0.1159	0.1092	0.1068	0.1147	0.1104
		150	0.0828	0.0925	0.1020	0.0962	0.0948	0.1014	0.1061
		300	0.0673	0.0772	0.0899	0.0820	0.0813	0.0896	0.1026
		700	0.0524	0.0600	0.0701	0.0622	0.0619	0.0700	0.1030
		1000	0.0483	0.0563	0.0620	0.0570	0.0569	0.0619	0.1014
	λ	25	0.0729	0.0950	0.0740	0.0668	0.0701	0.0845	0.0931
		75	0.0484	0.0555	0.0574	0.0494	0.0495	0.0610	0.0596
		150	0.0382	0.0425	0.0470	0.0404	0.0403	0.0487	0.0485
		300	0.0259	0.0284	0.0336	0.0289	0.0286	0.0342	0.0365
		700	0.0167	0.0179	0.0211	0.0181	0.0181	0.0213	0.0277
		1000	0.0141	0.0151	0.0166	0.0148	0.0149	0.0167	0.0252

Table 3. Parameter estimates of the UWL distribution for $(\alpha, \beta, \lambda) = (1.8, 0.8, 1.1)$

Measure	Parameter	n	ML	MPS	OLS	WLS	AD	**CVM**	**PC**
AE	α	25	2.4736	3.3141	3.4105	3.2787	3.2309	3.3417	3.2841
		75	2.1521	2.8359	2.7422	2.7736	2.6895	2.7065	2.5842
		150	2.0930	2.8809	2.8811	2.8407	2.7696	2.8552	2.6623
		300	2.0086	2.8345	2.8614	2.7592	2.7248	2.8423	2.6717
		700	1.9596	2.5454	2.5778	2.5245	2.5056	2.5659	2.4375
		1000	1.9841	2.5596	2.6100	2.5427	2.5428	2.5999	2.4514
	β	25	0.8579	0.8463	0.7476	0.7614	0.7692	0.7587	0.7609
		75	0.7933	0.7525	0.6928	0.7192	0.7180	0.6933	0.6841
		150	0.7823	0.7728	0.7345	0.7537	0.7512	0.7342	0.7317
		300	0.7739	0.7920	0.7627	0.7764	0.7747	0.7620	0.7600
		700	0.7829	0.8048	0.7772	0.7915	0.7918	0.7765	0.7772
		1000	0.7887	0.8115	0.7945	0.8032	0.8033	0.7939	0.7920
	λ	25	1.1248	1.3594	1.1508	1.1517	1.1817	1.2289	1.1327
		75	1.1301	1.2474	1.1907	1.1811	1.1844	1.2174	1.1825
		150	1.1346	1.1950	1.1632	1.1569	1.1601	1.1762	1.1582
		300	1.1220	1.1465	1.1346	1.1287	1.1298	1.1411	1.1306
		700	1.1080	1.1148	1.1167	1.1101	1.1101	1.1195	1.1132
		1000	1.1054	1.1090	1.1084	1.1042	1.1046	1.1104	1.1066
AB	α	25	1.1617	2.5090	2.5442	2.4212	2.4048	2.5252	2.4047
		75	1.0370	2.2059	2.2126	2.1613	2.0874	2.2107	2.0828
		150	0.9980	2.0961	2.1650	2.0504	1.9818	2.1570	1.9440
		300	0.9402	1.8854	2.0088	1.8366	1.8036	2.0023	1.8422
		700	0.7899	1.4056	1.6597	1.4808	1.4563	1.6566	1.4917
		1000	0.7559	1.3491	1.5569	1.3882	1.3879	1.5533	1.3659
	β	25	0.2105	0.3235	0.2932	0.2828	0.2820	0.3100	0.2736
		75	0.1522	0.2515	0.2815	0.2543	0.2463	0.2867	0.2595
		150	0.1316	0.2175	0.2378	0.2154	0.2114	0.2410	0.2164
		300	0.1208	0.1788	0.2029	0.1788	0.1771	0.2044	0.1850
		700	0.0990	0.1283	0.1657	0.1406	0.1394	0.1664	0.1471
		1000	0.0936	0.1213	0.1453	0.1258	0.1257	0.1456	0.1283
	λ	25	0.0947	0.2875	0.2036	0.1900	0.1955	0.2322	0.1744
		75	0.0801	0.1768	0.1648	0.1521	0.1510	0.1773	0.1564
		150	0.0703	0.1211	0.1168	0.1071	0.1076	0.1219	0.1090
		300	0.0516	0.0721	0.0800	0.0696	0.0699	0.0818	0.0751
		700	0.0380	0.0436	0.0567	0.0472	0.0469	0.0573	0.0524
		1000	0.0356	0.0398	0.0470	0.0412	0.0414	0.0471	0.0443
RMSE	α	25	1.1854	2.6724	2.7003	2.6124	2.6008	2.6799	2.6059
		75	1.0936	2.4362	2.4288	2.3977	2.3471	2.4257	2.3245
		150	1.0614	2.3692	2.4095	2.3380	2.2783	2.4035	2.2387
		300	1.0146	2.2331	2.3157	2.1803	2.1584	2.3090	2.1649
		700	0.8935	1.8395	2.0187	1.8781	1.8471	2.0152	1.8665
		1000	0.8715	1.7989	1.9508	1.8228	1.8229	1.9474	1.7796
	β	25	0.2695	0.3958	0.3690	0.3551	0.3517	0.3853	0.3449
		75	0.1927	0.3022	0.3501	0.3156	0.3059	0.3536	0.3253
		150	0.1655	0.2574	0.2921	0.2631	0.2602	0.2941	0.2679
		300	0.1486	0.2085	0.2452	0.2149	0.2137	0.2464	0.2215
		700	0.1220	0.1535	0.1931	0.1663	0.1647	0.1937	0.1720
		1000	0.1143	0.1439	0.1703	0.1502	0.1501	0.1708	0.1531
	λ	25	0.1076	0.3751	0.2608	0.2492	0.2600	0.3047	0.2275
		75	0.0880	0.2349	0.2203	0.2017	0.2008	0.2364	0.2068
		150	0.0784	0.1588	0.1581	0.1421	0.1424	0.1653	0.1458
		300	0.0615	0.0996	0.1100	0.0959	0.0966	0.1129	0.1014
		700	0.0473	0.0569	0.0739	0.0618	0.0609	0.0749	0.0680
		1000	0.0435	0.0505	0.0604	0.0524	0.0525	0.0609	0.0567

Table 4. Parameter estimates of the UWL distribution for $(\alpha, \beta, \lambda) = (1.2, 0.2, 0.3)$

Measure	Parameter	n	ML	MPS	OLS	WLS	AD	CVM	PC
AB	α	25	1	4	5	6	2	3	7
		75	1	5	6	3	2	4	7
		150	1	4	6	3	2	5	7
		300	1	2	6	4	3	5	7
		700	1	2	6	4	3	5	7
		1000	1	2	6	4	3	5	7
		Sum of ranks	6	19	35	24	15	27	42
		Rank	1	3	6	4	2	5	7
	β	25	1	2	7	5	4	6	3
		75	1	2	7	4	3	6	5
		150	1	2	6	4	3	5	7
		300	1	2	6	4	3	5	7
		700	1	2	6	4	3	5	7
		1000	1	2	6	3	3	5	7
		Sum of ranks	6	12	38	24	19	32	36
		Rank	1	2	7	4	3	5	6
	λ	25	3	6	4	1	2	5	7
		75	1	4	5	3	2	6	7
		150	1	4	5	3	2	6	7
		300	1	2	5	4	3	6	7
		700	1	4	5	2	2	6	7
		1000	1	4	5	2	3	5	7
		Sum of ranks	8	24	29	15	14	34	42
		Rank	1	4	5	3	2	6	7
RMSE	α	25	1	4	5	6	2	3	7
		75	1	6	5	3	2	4	7
		150	1	6	5	3	2	4	7
		300	1	2	6	4	3	5	7
		700	1	4	6	3	2	5	7
		1000	1	2	6	4	3	5	7
		Sum of ranks	6	24	33	23	14	26	42
		Rank	1	4	6	3	2	5	7
	β	25	1	2	7	6	4	5	3
		75	1	2	7	4	3	6	5
		150	1	2	6	4	3	5	7
		300	1	2	6	4	3	5	7
		700	1	2	6	4	3	5	7
		1000	1	2	6	4	3	5	7
		Sum of ranks	6	12	38	26	19	31	36
		Rank	1	2	7	4	3	5	6
	λ	25	3	7	4	1	2	5	6
		75	1	4	5	2	3	7	6
		150	1	4	5	3	2	7	6
		300	1	2	5	4	3	6	7
		700	1	2	5	3	3	6	7
		1000	1	4	5	2	3	6	7
		Sum of ranks	8	23	29	15	16	37	39
		Rank	1	4	5	2	3	6	7

In this study, the logit link is employed. With it, we can write

$$\eta_i = \frac{e^{\boldsymbol{x}_i'\boldsymbol{\delta}}}{1 + e^{\boldsymbol{x}_i'\boldsymbol{\delta}}}.$$

Thus, the final form of the PDF of the regression model for purposes of inference can be obtained by substituting η_i into Equation (16). The parameters of the regression model can thus be obtained using the PDF of the regression model.

Table 5. Parameter estimates of the UWL distribution for $(\alpha, \beta, \lambda) = (1.8, 0.8, 1.1)$

Measure	Parameter	n	ML	MPS	OLS	WLS	AD	CVM	PC
AB	α	25	1	5	7	4	3	6	2
		75	1	5	7	4	3	6	2
		150	1	5	7	4	3	6	2
		300	1	5	7	3	2	6	4
		700	1	2	7	4	3	6	5
		1000	1	2	7	5	4	6	3
		Sum of ranks	6	24	42	24	18	36	18
		Rank	1	4	7	4	2	6	2
	β	25	1	7	5	4	3	6	2
		75	1	3	6	4	2	7	5
		150	1	5	6	3	2	7	4
		300	1	4	6	3	2	7	5
		700	1	2	6	4	3	7	5
		1000	1	2	6	4	3	7	5
		Sum of ranks	6	23	35	22	15	41	26
		Rank	1	4	6	3	2	7	5
	λ	25	1	7	5	3	4	6	2
		75	1	6	5	3	2	7	4
		150	1	6	5	2	3	7	4
		300	1	4	6	2	3	7	5
		700	1	2	6	4	3	7	5
		1000	1	2	6	3	4	7	5
		Sum of ranks	6	27	33	17	19	41	25
		Rank	1	5	6	2	3	7	4
RMSE	α	25	1	5	7	4	2	6	3
		75	1	7	6	4	3	5	2
		150	1	5	7	4	3	6	2
		300	1	5	7	4	2	6	3
		700	1	2	7	5	3	6	4
		1000	1	3	7	4	5	6	2
		Sum of ranks	6	27	41	25	18	35	16
		Rank	1	5	7	4	3	6	2
	β	25	1	7	5	4	3	6	2
		75	1	2	6	4	3	7	5
		150	1	2	6	4	3	7	5
		300	1	2	6	4	3	7	5
		700	1	2	6	4	3	7	5
		1000	1	2	6	4	3	7	5
		Sum of ranks	6	17	35	24	18	41	27
		Rank	1	2	6	4	3	7	5
	λ	25	1	7	5	3	4	6	2
		75	1	6	5	3	2	7	4
		150	1	6	5	2	3	7	4
		300	1	4	6	2	3	7	5
		700	1	2	6	4	3	7	5
		1000	1	2	6	3	4	7	5
		Sum of ranks	6	27	33	17	19	41	25
		Rank	1	5	6	2	3	7	4

6.1. Monte Carlo simulation for regression

A simulation study is carried out to demonstrate the behavior of the subjacent estimators of the parameters in the quantile regression model. Two independent variables are used to perform the simulation with the following regression structure:

$$\operatorname{logit}\left(\frac{\eta_i}{1-\eta_i}\right) = \delta_0 + \delta_1 x_{i1} + \delta_2 x_{i2} \tag{17}$$

The steps used for the simulation for the parameter values $(\beta, \lambda, \delta_0, \delta_1, \delta_2) = (0.8, 1.1, 1.2, 0.7, 0.8)$ are as follows:

1. Generate samples of sizes $n = 50, 100, 350, 500$ and 700 based on the UWL distribution.
2. Generate the covariates x_1 and x_2 from a uniform $U(0, 1)$ distribution.
3. Estimate the parameters of the model $(\alpha, \beta, \delta_0, \delta_1, \delta_2)$ using the ML estimation method.
4. Repeat steps 1 to 3 $N = 3000$ times.
5. Compute the AE, AB, and RMSE of the estimated parameters.

The process is repeated for the lower quantile $(u = 0.25)$, the median $(u = 0.5)$ and the upper quantile $(u = 0.75)$. Table 6 shows the results of the simulation study. The ML estimators are consistent, as increasing the sample size results in the AE approaching the true parameter values while the AB and RMSE decrease.

Table 6. Simulation results for the UWL quantile regression model

Parameter	n	$u = 0.25$			$u = 0.50$			$u = 0.75$		
		AV	AB	RMSE	AV	AB	RMSE	AV	AB	RMSE
β	50	1.1328	0.3333	0.4039	1.1313	0.3320	0.4028	1.1465	0.3469	0.4143
	100	1.0585	0.2585	0.3346	1.0591	0.2592	0.3330	1.0902	0.2902	0.3545
	350	0.8907	0.0907	0.1888	0.8974	0.0974	0.1955	0.9251	0.1251	0.2201
	500	0.8512	0.0512	0.1407	0.8544	0.0544	0.1444	0.8710	0.0710	0.1663
	700	0.8223	0.0223	0.0923	0.8251	0.0251	0.0984	0.8383	0.0383	0.1210
λ	50	1.1046	0.0787	0.1113	1.1086	0.0797	0.1112	1.1066	0.0850	0.1175
	100	1.0950	0.0449	0.0692	1.0963	0.0459	0.0700	1.0956	0.0510	0.0741
	350	1.0955	0.0099	0.0247	1.0954	0.0101	0.0244	1.0930	0.0134	0.0284
	500	1.0973	0.0044	0.0148	1.0967	0.0052	0.0166	1.0956	0.0066	0.0186
	700	1.0988	0.0018	0.0089	1.0986	0.0023	0.0105	1.0975	0.0034	0.0126
δ_0	50	2.6085	1.5936	2.1828	2.6100	1.6122	2.2067	2.6710	1.6826	2.2630
	100	2.5086	1.3644	2.0337	2.5261	1.3751	2.0355	2.6508	1.5226	2.1520
	350	1.7639	0.5653	1.3047	1.8112	0.6128	1.3756	1.9747	0.7768	1.5378
	500	1.5304	0.3305	0.9984	1.5504	0.3512	1.0327	1.6534	0.4554	1.1706
	700	1.3429	0.1430	0.6477	1.3672	0.1672	0.7107	1.4479	0.2479	0.8650
δ_1	50	0.9552	2.3955	3.5493	1.0180	2.4336	3.6293	1.0494	2.5652	3.7143
	100	0.6229	1.6459	2.7262	0.6528	1.7414	2.8703	0.7365	2.0121	3.1418
	350	0.5245	0.5255	1.4400	0.5271	0.5705	1.5043	0.5048	0.7519	1.7156
	500	0.5975	0.3086	1.1002	0.6011	0.3189	1.1023	0.5139	0.3886	1.1646
	700	0.6451	0.1170	0.6315	0.6311	0.1331	0.6724	0.6163	0.2311	0.9534
δ_2	50	0.9786	2.3751	3.5358	0.9192	2.3745	3.5073	0.9483	2.5081	3.6059
	100	0.6333	1.7423	2.8562	0.6327	1.7287	2.8436	0.6589	1.9447	3.0267
	350	0.6258	0.5481	1.4743	0.6058	0.6149	1.5834	0.5754	0.7784	1.7732
	500	0.6672	0.2807	0.9946	0.6761	0.3245	1.1011	0.6672	0.4221	1.2853
	700	0.7598	0.1397	0.7627	0.7386	0.1439	0.7064	0.7191	0.2097	0.8841

7. Applications

This section demonstrates the usefulness of the UWL distribution and the UWL quantile regression model. It compares their performance with that of other competing models.

7.1. Univariate applications

The application of the UWL distribution to analyze claims data sets from France and Brazil is demonstrated in the following subsections. The comparison of the performance of the UWL distribution with other distributions, with regards to the data, is carried out using the Akaike information criterion (AIC), Bayesian information criterion (BIC), Cramér von–Mises (CVM), Anderson–Darling (AD), and Kolmogrov–Smirnov (KS) goodness-of-fit measures. The distribution with the least measure and highest corresponding p-value for CVM, AD, and KS is considered to provide a good fit to the data set. The following distributions are used for the comparison: unit Weibull (UW) distribution, beta distribution, Kumaraswamy (Kum) distribution, unit gamma/Gompertz (UGG) distribution, Topp–Leone (TL) distribution, and unit Rayleigh (UR) distribution. Finally, the Weibull–loss (WL) distribution is also considered for the comparison.

7.1.1. Data set I. Vehicle insurance data from Brazil

The first data set consists of vehicle insurance claim amounts, specifically for collisions, from Brazil in 2011. The data set can be obtained from the CASdatasets package [10] in the R program with the name brvehins2a. The first 50 observations with claims greater than zero, divided by 1000,000, were used for this study.

A histogram of the data set with a total test time (TTT) plot is given in Figure 6. It can be remarked that the data are extremely right skewed. Also, the TTT plot is convex in shape, suggesting that the hazard rate of the data is decreasing. The plots suggest that the UWL distribution is adapted to the situation.

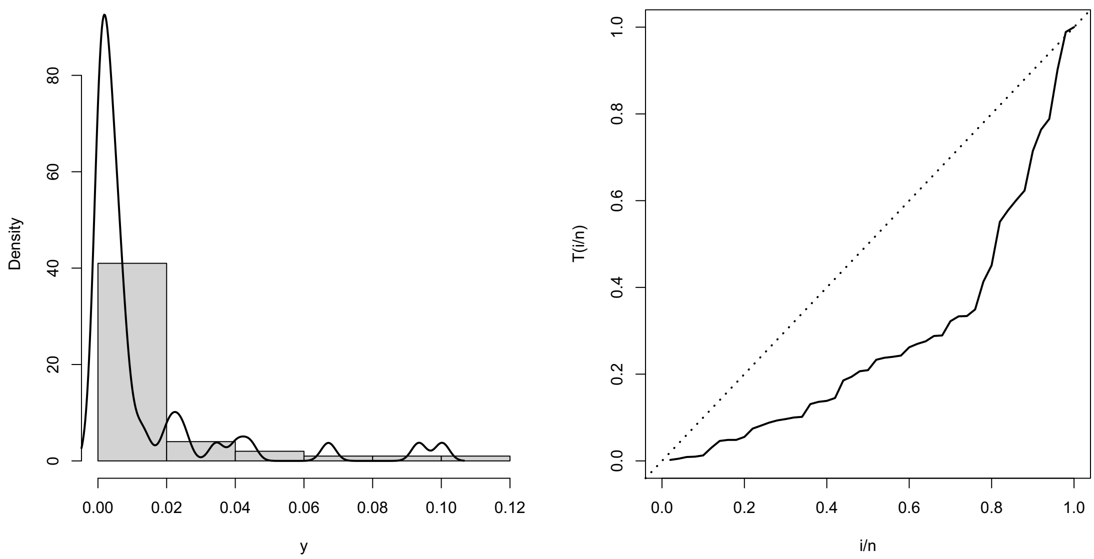

Figure 6. Histogram and TTT plot of data set I

The parameter estimates of the fitted distributions are given in Table 7 with their corresponding standard errors in brackets.

Table 7. Parameter estimates with standard errors for data set I

Distribution	Parameter estimates (standard errors)		
	α	β	λ
UWL	0.9429 (0.1563)	0.0062 (0.0358)	1.2650 (6.3483)
UW	0.0032 (0.0038)	3.1416 (0.5839)	
Beta	0.4803 (0.0790)	38.8041 (10.0788)	
Kum	0.5920 (0.0684)	16.7297 (5.4496)	
UGG	79.1894 (36.6476)	1.0386 (0.5077)	0.7873 (0.1558)
UL	79.1120 (11.0520)		
UR	0.0277 (0.0039)		
WL	0.8831 (0.1539)	9.4006 (7.2982)	0.0756 (0.1165)

Table 8 displays the goodness-of-fit measures to assess the fit of each distribution to the data. It can be noticed that the UWL distribution has the least AIC, BIC, CVM, AD, and KS measures with the highest corresponding p-values of the CVM, AD, and KS measures. This means that the UWL distribution best describes the data compared to the other distributions. The WL distribution closely follows in terms of modeling performance.

Table 8. Goodness-of-fit measures for data set I

Distribution	AIC	BIC	CVM (p-value)	AD (p-value)	KS (p-value)
UWL	–365.3575	–362.6214	0.0299 (0.9774)	0.2189 (0.9843)	0.0631 (0.9879)
UW	–365.0906	–362.1256	0.0925 (0.6250)	0.5507 (0.6949)	0.1105 (0.5746)
Beta	–360.5053	–356.6812	0.2528 (0.1849)	1.2697 (0.2423)	0.1657 (0.1284)
Kum	–364.7511	–360.9271	0.1147 (0.5191)	0.6406 (0.6097)	0.1218 (0.4486)
UGG	–362.4933	–356.7572	0.1046 (0.5647)	0.6491 (0.6020)	0.1180 (0.4896)
UL	–336.3184	–334.4063	1.6597 (0.0001)	9.8311 (0.0000)	0.3068 (0.0002)
UR	222.1444	224.0564	0.7026 (0.0122)	3.7531 (0.0116)	0.2297 (0.0102)
WL	–365.3100	–360.5750	0.0337 (0.9640)	0.2297 (0.9801)	0.0688 (0.9721)

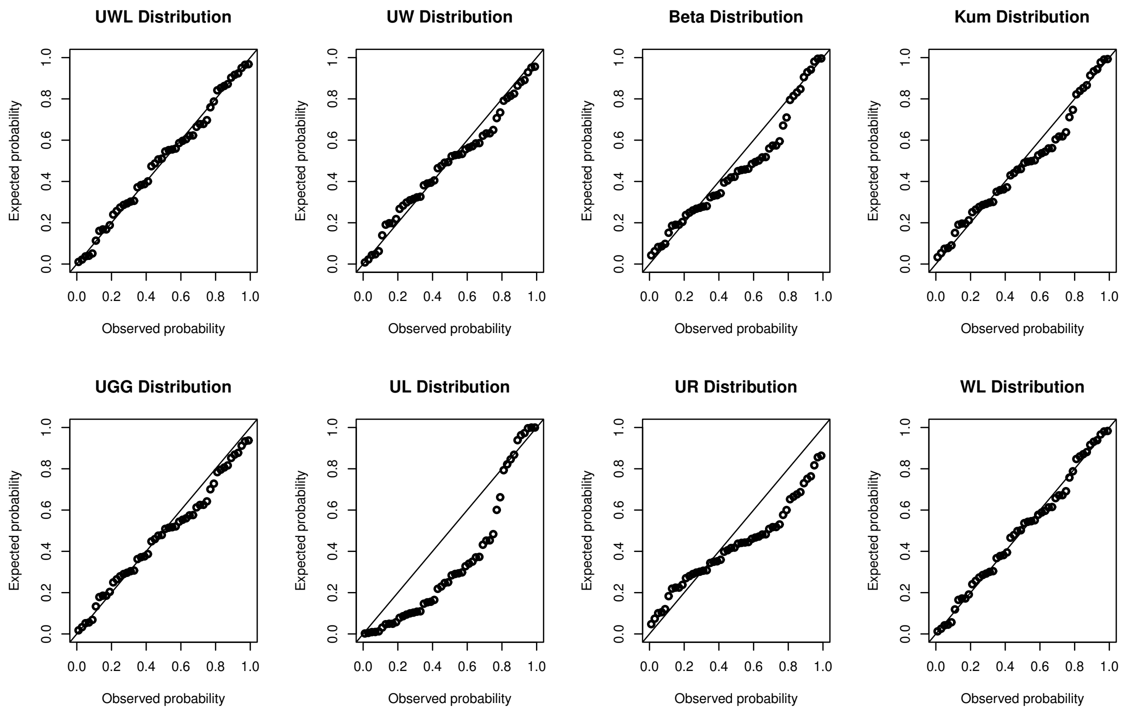

Figure 7. P–P plots of the fitted distributions for data set I

The probability–probability (P–P) plots of the fitted models are given in Figure 7 to illustrate the fit of the distributions to the data. The UWL and WL distributions closely describe the data, as the plots cluster along the diagonal lines.

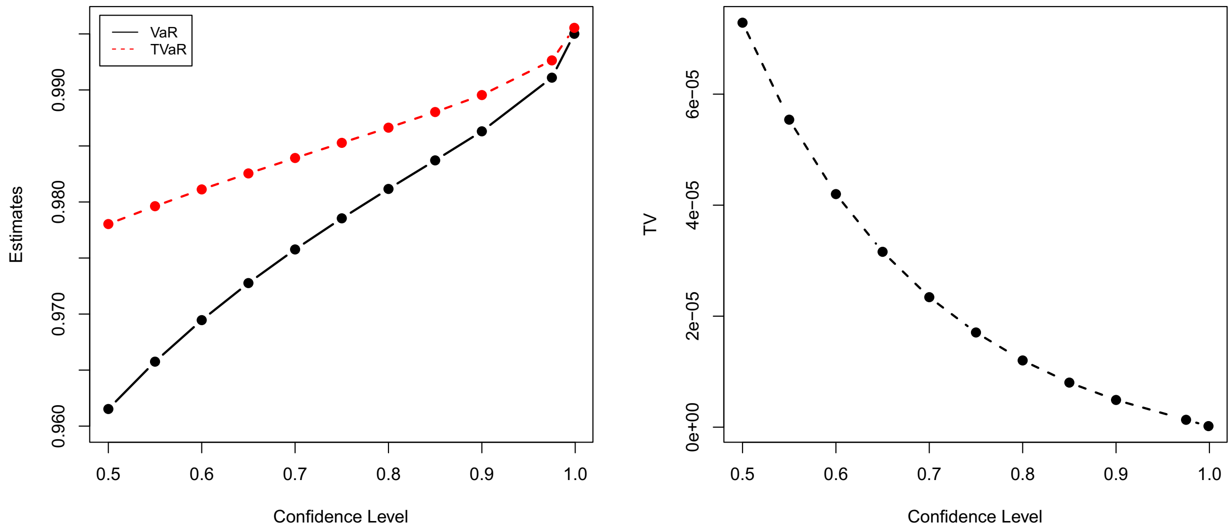

Figure 8. Estimates of risk measures, including VaR, $TVaR$, and TV associated with the UWL distribution for data set I

Using the estimated parameter values of the UWL distribution, some risk measures and premium principles are estimated. Figure 8 exposes the plots of the estimated risk measures, including VaR, $TVaR$, and TV. VaR and $TVaR$ increase with an increase in significance level, while TV decreases. This is generally expected in actuarial practice. As the confidence level increases, VaR and $TVaR$, which may represent some capital requirement, are expected to increase while the risk associated with that, represented by TV, decreases.

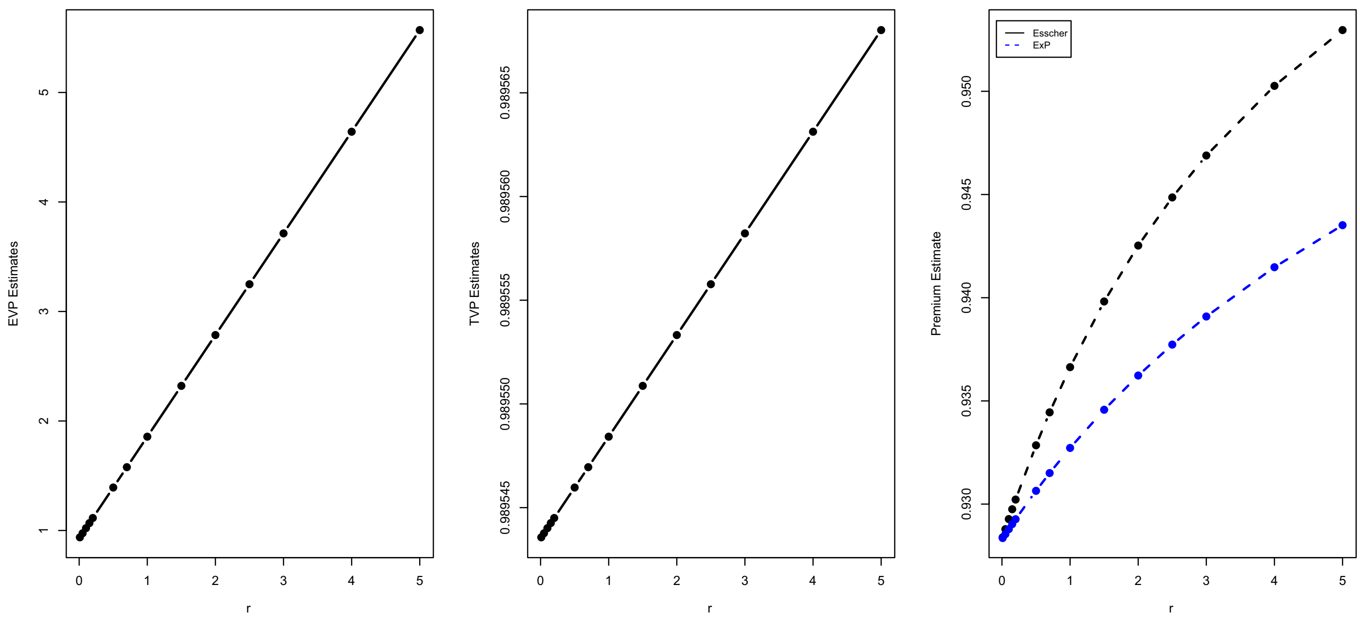

Figure 9. Estimates of risk measures, including EVP, TVP, ExP, and Esscher principles, associated with the UWL distribution for data set I

Figure 9 shows the plots of estimated premiums, including EVP, TVP, ExP, and Esscher principles. It can be seen that all the premium estimates increase with increasing risk loading. This is consistent with actuarial practice; as the risk associated with the event increases, the premium is expected to also increase. It can be observed that EVP estimates are the highest and ExP estimates are the least, with the same corresponding risk loading.

7.1.2. Data set II. Motor insurance data from France

The second data set consists of claims for private motor insurance from France for the period 2003 to 2004. The data set can be obtained from the *CASdatasets* package [10] in R with the name *fremotor1sev0304a*. The first 50 claim amounts greater than zero, divided by 100,000, were used for the study.

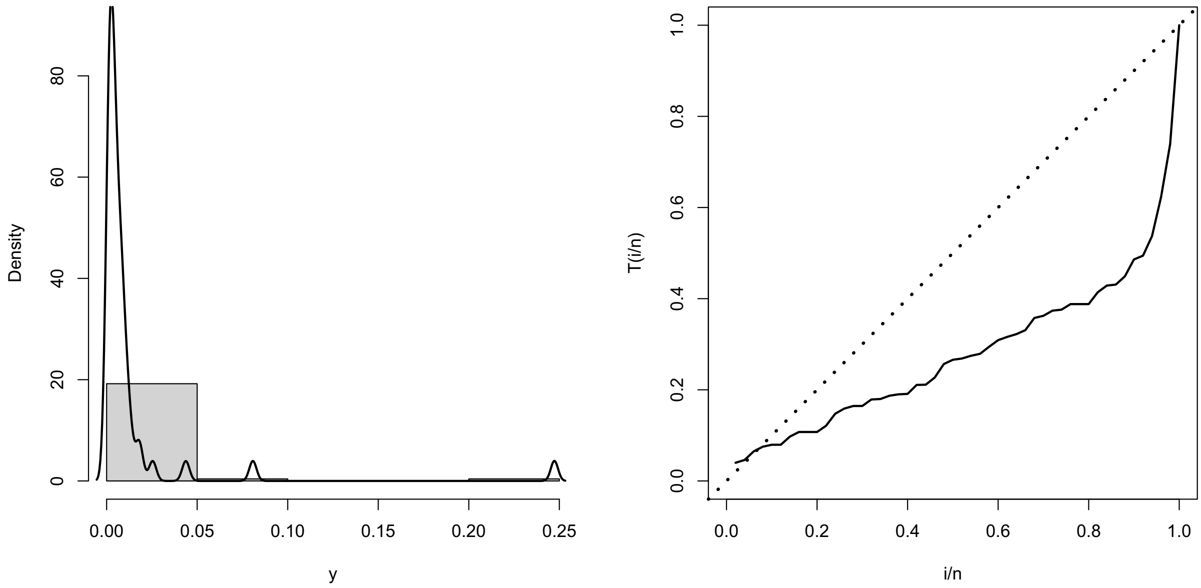

Figure 10. Histogram and TTT Plot of Motor Insurance data

Figure 10 displays the histogram and TTT plot of the data set. It can be noticed that the data are extremely right skewed. Also, the TTT plot has a convex shape, suggesting that the data have a decreasing hazard rate. Thus, the UWL distribution can be used to fit these data.

Table 9 shows the parameter estimates of the fitted distributions with their corresponding standard errors.

Table 9. Parameter estimates (standard errors) of distributions for data set II

Distribution	α	β	λ
UWL	1.4343 (0.1792)	0.0020 (0.0021)	4.7639 (0.0001)
UW	0.0024 (0.0019)	3.4895 (0.4466)	
Beta	0.5819 (0.0974)	42.9454 (10.6451)	
Kum	0.6704 (0.0675)	22.5613 (6.9591)	
UGG	305.0186 (136.4060)	2.2037 (1.0728)	0.9026 (0.1165)
UL	69.2445 (9.6571)		
UR	0.0331 (0.0047)		
WL	1.4122 (0.2984)	7.0428 (0.0135)	0.0035 (0.0060)

Table 10 displays the goodness-of-fit measures to assess the fit of each distribution to data set II. It can be observed that the UWL distribution has the least AIC and BIC. It also has the least CVM, AD, and KS measures with correspondingly higher p-values. Again, the UWL distribution is closely followed by the WL distribution.

The P–P plots of the fitted models are given in Figure 11. It can be noticed that the UWL distribution best describes the data set.

Table 10. Goodness-of-fit measures for data set I

Distribution	AIC	BIC	**CVM (*p*-value)**	**AD (*p*-value)**	**KS (*p*-value)**
UWL	–372.7223	–366.9863	0.0249 (0.9904)	0.2696 (0.9589)	0.0634 (0.9880)
UW	–363.5235	–359.6994	0.5355 (0.0320)	2.7887 (0.0353)	0.2117 (0.0226)
Beta	–339.1632	–335.3392	0.7296 (0.0105)	3.9474 (0.0093)	0.2377 (0.0070)
Kum	–349.0566	–345.2326	0.4059 (0.0693)	2.5844 (0.0450)	0.1761 (0.0901)
UGG	–368.2395	–362.5034	0.2731 (0.1612)	1.5469 (0.1658)	0.1637 (0.1372)
UL	–323.0120	–321.1000	2.0265 (<0.0001)	9.9510 (<0.0001)	0.3391 (<0.0001)
UR	208.8915	210.8035	1.6686 (0.0001)	8.2085 (0.0001)	0.3381 (<0.0001)
WL	–372.6986	–366.9626	0.0324 (0.9689)	0.2981 (0.9393)	0.0744 (0.9448)

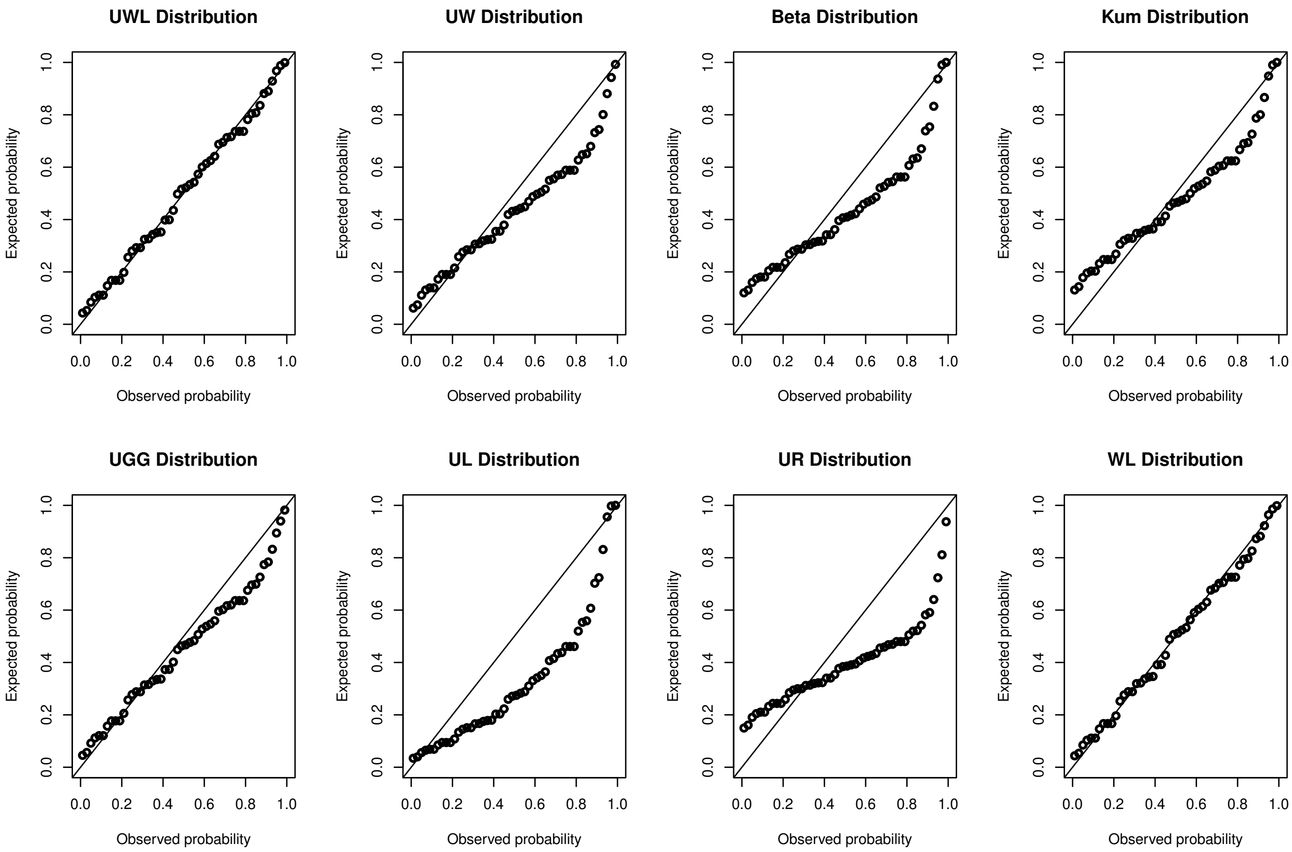

Figure 11. P–P plots of the fitted distributions for data set II

Furthermore, risk measures and premiums are estimated for the data set. Figure 12 exposes the estimates of VaR, $TVaR$, and TV. It can be remarked that VaR and $TVaR$ decrease as the significance level increases, while TV decreases.

Premium estimates using the EVP, TVP, ExP, and Esscher principles are shown in Figure 13. It can be observed that all the estimates increase with increasing risk loading. Again, with the same corresponding risk loading, EVP premiums are the highest while ExP premiums are the lowest.

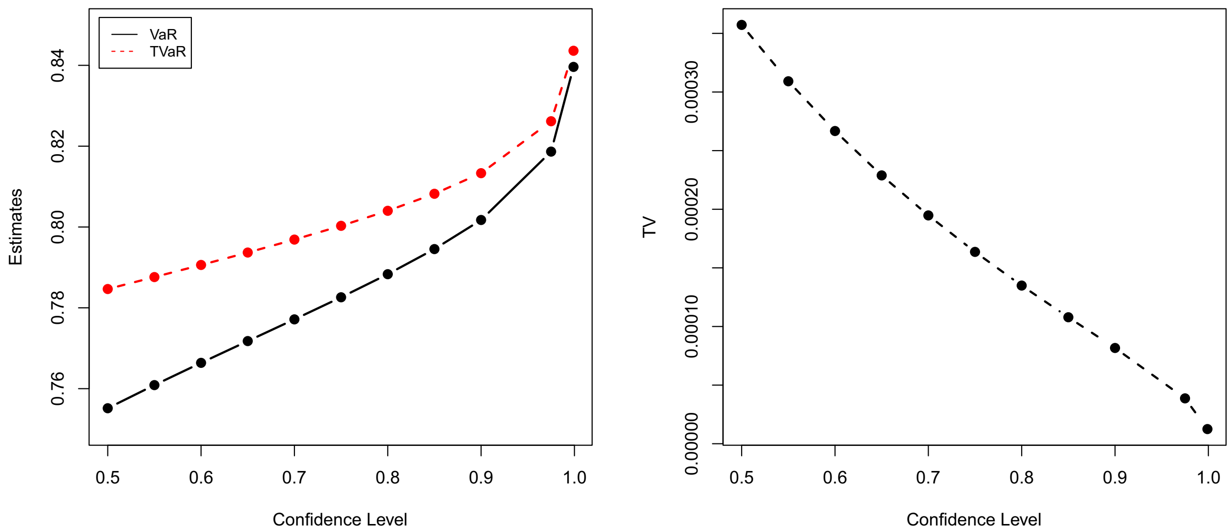

Figure 12. Estimates of risk measures, including VaR, $TVaR$, and TV associated with the UWL distribution for data set II

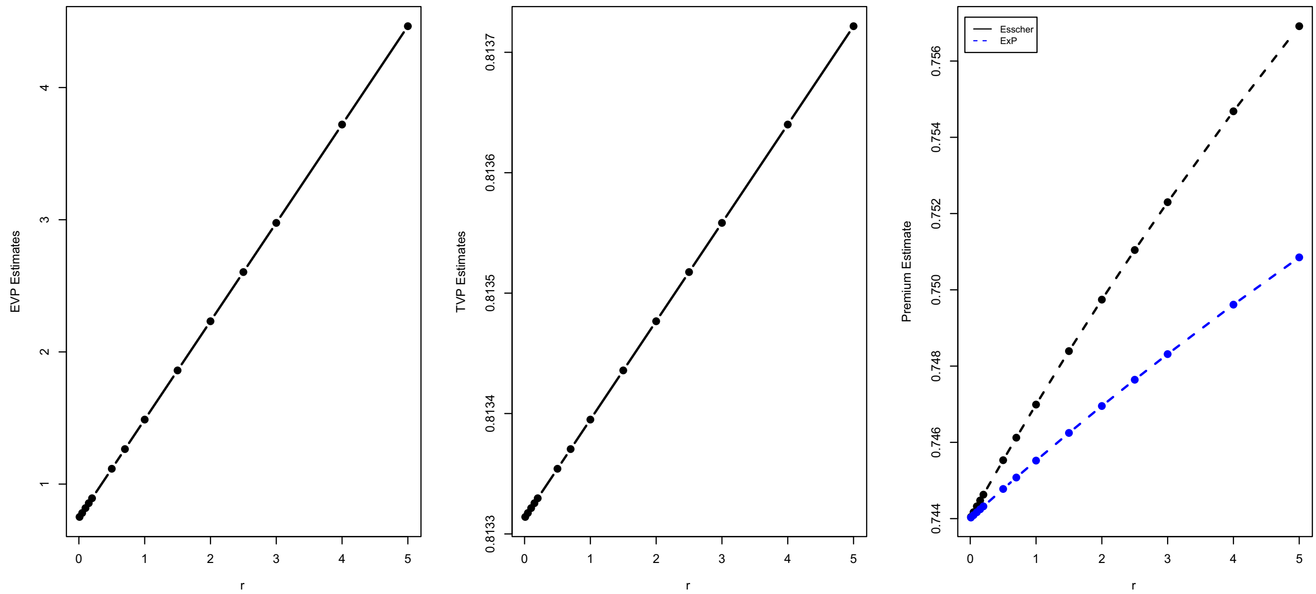

Figure 13. Estimates of risk measures, including EVP, TVP, ExP, and Esscher principles, associated with the UWL distribution for data set II

7.2. Regression application

The application of the UWL quantile regression model is demonstrated in this section. The data set used is from the Swedish insurance company Wasa and can be obtained from the *insuranceData* package [38] in the R program with the name *dataOhlsson*. Aggregated data on all insurance policies and claims from 1994 to 1998 are used. Also, the dependent variable is the claim cost in 10 billion Swedish krona. For this study, the claims were divided by 1 million. The independent variables used are the vehicle age and MC class, a classification by the so-called EV ratio, defined as (engine power in kW 100)/(vehicle weight in kg+75), rounded to the nearest lower integer, where 75kg represents the average driver weight. The EV ratios are divided into seven classes.

The independent variable, MC class, is a categorical variable with seven (7) levels. For the regression model application, it is coded using an indicator variable into $k-1$ indicator variables, where k is the number of levels and one of the variables is used as a reference level. In this study, level 6 is chosen as a reference level because it has the highest number of occurrences. With this, the following regression structure is used:

$$\log \eta_i = \delta_0 + \delta_1 \text{Vehicle age}_i + \delta_2 \text{Level } 1_i + \delta_3 \text{Level } 2_i + \delta_4 \text{Level } 3_i + \delta_5 \text{Level } 4_i + \delta_6 \text{Level } 5_i + \delta_7 \text{Level } 7_i, \quad i = 1, \ldots, n$$

The performance of the UWL regression model is compared with that of the UW regression model [28], Kumaraswamy (Kum) regression model [30], UL regression model [28], beta regression model [11], and unit improved second-degree Lindley (UISDL) regression model [5].

Table 11. Regression parameter estimates

Model	Parameter	β	λ	δ_0	δ_1	δ_2	δ_3	δ_4	δ_5	δ_6	δ_7
UWL	Estimate	7.3347	0.9544	–3.6925	–0.0737	–0.4548	–0.4615	–0.4566	–0.3968	–0.2494	0.2908
	SE	1.4567	0.0502	0.1558	0.0109	0.2762	0.2433	0.1903	0.2179	0.1877	0.6648
	p-value	< 0.0001	< 0.0001	< 0.0001	< 0.0001	0.0997	0.0578	0.0164	0.0687	0.1840	0.6618
UW	Estimate		3.2528	–4.0195	–0.0552	–0.3046	–0.0480	–0.4032	–0.3652	–0.2726	0.6205
	SE		0.0959	0.1334	0.0100	0.2473	0.2177	0.1594	0.1868	0.1632	0.5316
	p-value		< 0.0001	< 0.0001	< 0.0001	0.2181	0.8257	0.0115	0.0506	0.0949	0.2431
Kum	Estimate		0.7146	–3.9241	–0.0547	–0.0347	–0.3864	0.1046	–0.1442	–0.1521	0.2090
	SE		0.0232	0.1205	0.0063	0.2375	0.2137	0.1547	0.1793	0.1575	0.6283
	p-value		< 0.0001	< 0.0001	< 0.0001	0.8838	0.0705	0.4990	0.4213	0.3340	0.7394
UL	Estimate			–3.3249	–0.0503	0.1303	–0.3342	0.2733	–0.0782	–0.1141	0.1578
	SE			0.0820	0.0042	0.1703	0.1516	0.1082	0.1264	0.1112	0.4408
	p-value			< 0.0001	< 0.0001	0.4443	0.0275	0.0116	0.5360	0.3045	0.7204
Beta	Estimate		24.1531	–3.2928	–0.0333	–0.1555	–0.1314	–0.1489	–0.1631	–0.1027	0.2164
	SE		1.5927	0.0858	0.0059	0.1498	0.1371	0.0978	0.1146	0.1001	0.3766
	p-value		< 0.0001	< 0.0001	< 0.0001	0.2994	0.3377	0.1281	0.1546	0.3052	0.5656
UISDL	Estimate			–3.3246	–0.0503	0.1303	–0.3342	0.2735	–0.0783	–0.1141	0.1575
	SE			0.0819	0.0042	0.1702	0.1515	0.1082	0.1263	0.1111	0.4404
	p-value			< 0.0001	< 0.0001	0.4438	0.0274	0.0115	0.5354	0.3046	0.7206

Table 11 displays the parameter estimates of the regression models with the corresponding standard errors and p-values. It can be noticed that all the true parameters (β, λ) of the models are significant at 5% significance level. Also, vehicle age is significant for all the models. Furthermore, the MC levels 1, 4, 5, and 7 are insignificant for all the models, while level 3 is only insignificant for Kum and beta regression models. Finally, level 2 is significant for the UL and UISDL regression models.

As age is a significant contributor to claims, the interpretation of its estimated coefficient in the regression model is given as follows: For an estimated coefficient of the UWL quantile regression model, $\hat{\delta}_i$, associated with a continuous independent variable, then a unit increase in the independent variable will result in a $100\%(\mathrm{e}^{\hat{\delta}_i} - 1)$ change in the conditional quantile of the dependent variable whiles holding other independent variables constant. Thus, holding all other independent variables constant, a unit increase in vehicle age would result in a statistically significant decrease of 7.10% in the median claims.

Table 12 shows the information criteria of the fitted models. It can be observed that the UWL regression model has the least of the measures. This indicates that it fits the data better than the other models.

To evaluate the performance of the fitted models, a Cox–Snell residuals analysis is conducted. In this analysis, the residuals should follow a standard exponential distribution if the model fits well.

Table 12. Information criteria of fitted regression models

Regression model	$-\ell$	AIC	BIC
UWL	–1921.4457	–3822.8915	–3777.8187
UW	–1911.4811	–3804.9622	–3764.3967
Kum	–1912.0681	–3806.1362	–3765.5707
UL	–1838.3727	–3660.7455	–3624.6873
Beta	–1879.7858	–3741.5717	–3701.0062
UISDL	–1838.0897	–3660.1795	–3624.1212

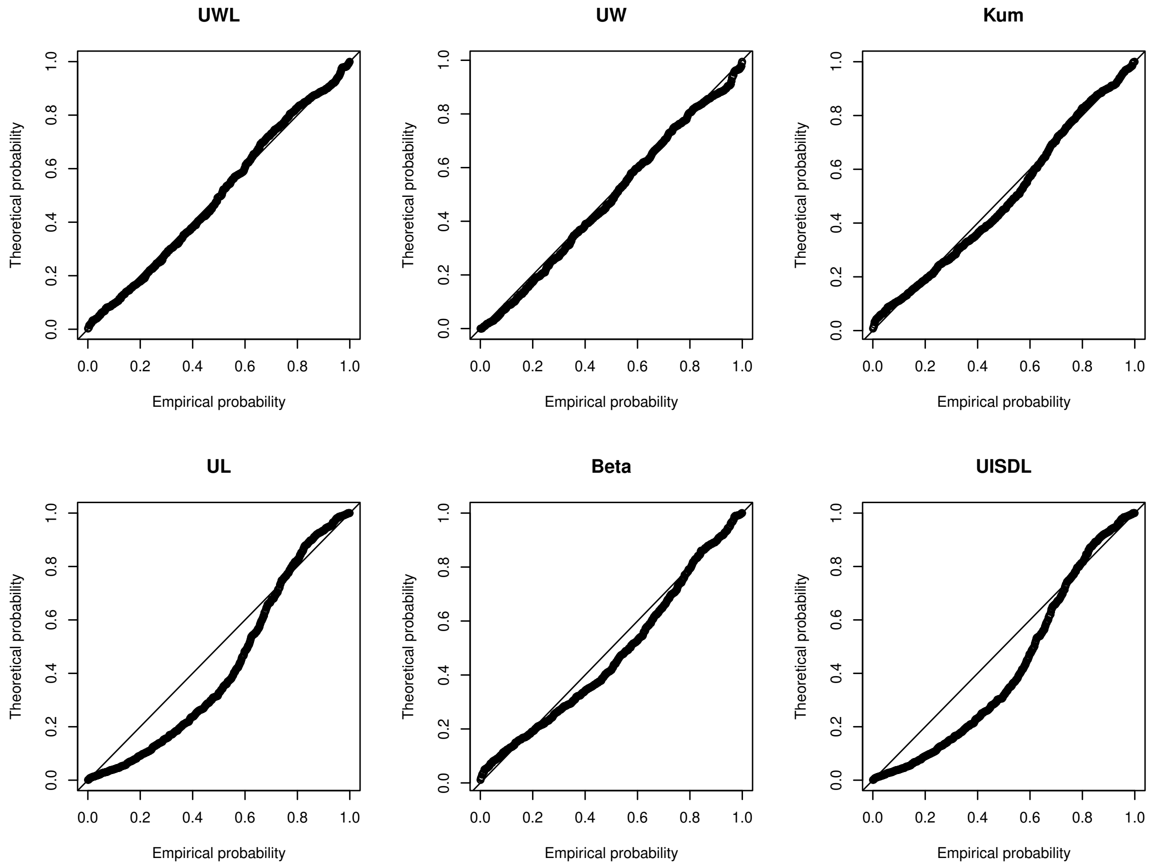

Figure 14. Cox–Snell residuals P–P plots

This is graphically represented by Figure 14, with the plots of the empirical probabilities of the residuals against the theoretical probabilities from the standard exponential distribution. From this figure, it can be noticed that the UWL regression model performs better in describing the data than the other models.

8. Conclusion

This article presented a new distribution defined on the bounded interval, known as the unit Weibull loss distribution. Several statistical and actuarial properties were developed. Various plots, including PDF, hazard, skewness, and kurtosis, indicate that it exhibits desirable properties. Several parameter estimation methods were used to obtain accurate estimates. The subjacent estimators were all shown to be consistent using Monte Carlo simulations. However, the ML estimation method performed better in estimating the involved parameters. Also, a quantile regression model based on responses following the UWL distribution was developed. Fitting applications and the corresponding regression model were performed to demonstrate the usefulness of the distribution. The results reveal that the UWL distribution and its regression model can be used as alternatives to other models in describing insurance data.

Funding

The authors declare that no funding was received for this research.

References

[1] Abubakari, A. G., Luguterah, A., and Nasiru, S. Unit exponentiated Fréchet distribution: Actuarial measures, quantile regression and applications. *Journal of the Indian Society for Probability and Statistics 23*, 2 (2022), 387–424.

[2] Afify, A. Z., Nassar, M., Kumar, D., and Cordeiro, G. M. A new unit distribution: properties, inference, and applications. *Electronic Journal of Applied Statistical Analysis 15*, 2 (2022), 460–484.

[3] Ahmad, Z., Mahmoudi, E., and Hamedani, G. G. A family of loss distributions with an application to the vehicle insurance loss data. *Pakistan Journal of Statistics and Operation Research 15*, 3 (2019), 731–744.

[4] Al-Omari, A. I., Alanzi, A. R. A., and Alshqaq, S. S. The unit two parameters Mirra distribution: Reliability analysis, properties, estimation and applications. *Alexandria Engineering Journal 92* (2024), 238–253.

[5] Altun, E., and Cordeiro, G. M. The unit-improved second-degree Lindley distribution: inference and regression modeling. *Computational Statistics 35*, 1 (2020), 259–279.

[6] Anderson, T. W., and Darling, D. A. Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes. *The Annals of Mathematical Statistics 23*, 2 (1952), 193–212.

[7] Bantan, R. A. R., Chesneau, C., Jamal, F., Elgarhy, M., Tahir, M. H., Ali, A., Zubair, M., and Anam, S. Some new facts about the unit-Rayleigh distribution with applications. *Mathematics 8*, 11 (2020), 1954.

[8] Bantan, R. A. R., Jamal, F., Chesneau, C., and Elgarhy, M. Theory and applications of the unit gamma/Gompertz distribution. *Mathematics 9*, 16 (2021), 1850.

[9] Cheng, R., and Amin, N. *Maximum product of spacings estimation with application to the lognormal distribution*. Mathematical report 79-1, University of Wales IST, Cardif, UK, 1979.

[10] Dutang, C., and Charpentier, A. (11 December 2020) Package 'CASdatasets' (accessed on 26 November 2022).

[11] Ferrari, S., and Cribari-Neto, F. Beta regression for modelling rates and proportions. *Journal of Applied Statistics 31*, 7 (2004), 799–815.

[12] Furman, E., and Landsman, Z. Tail variance premium with applications for elliptical portfolio of risks. *ASTIN Bulletin: The Journal of the IAA 36*, 2 (2006), 433–462.

[13] Gómez-Déniz, E., Sordo, M. A., and Calderín-Ojeda, E. The Log–Lindley distribution as an alternative to the beta regression model with applications in insurance. *Insurance: Mathematics and Economics 54* (2014), 49–57.

[14] Haj Ahmad, H., Almetwally, E. M., Elgarhy, M., and Ramadan, D. A. On unit exponential Pareto distribution for modeling the recovery rate of COVID-19. *Processes 11*, 1 (2023), 232.

[15] Jodrá, P., and Jiménez-Gamero, M. D. A quantile regression model for bounded responses based on the exponential-geometric distribution. *REVSTAT – Statistical Journal 18*, 4 (2020), 415–436 .

[16] Kao, J. H. K. A graphical estimation of mixed Weibull parameters in life test electron tube. *Technometrics 1*, 4 (1959), 389–407.

[17] Korkmaz, M. Ç. A new heavy-tailed distribution defined on the bounded interval: the logit slash distribution and its application. *Journal of Applied Statistics 47*, 12 (2020), 2097–2119.

[18] Korkmaz, M. Ç., Altun, E., Chesneau, C., and Yousof, H. M. On the Unit-Chen distribution with associated quantile regression and applications. *Mathematica Slovaca 72*, 3 (2022), 765–786.

[19] Korkmaz, M. Ç., and Chesneau, C. On the unit Burr-XII distribution with the quantile regression modeling and applications. *Computational and Applied Mathematics 40*, (2021), 29.

[20] Korkmaz, M. Ç., Chesneau, C., and Korkmaz, Z. S. The unit folded normal distribution: A new unit probability distribution with the estimation procedures, quantile regression modeling and educational attainment applications. *Journal of Reliability and Statistical Studies 15*, 1 (2022), 261–298.

[21] Korkmaz, M. Ç., and Korkmaz, Z. S. The unit log–log distribution: a new unit distribution with alternative quantile regression modeling and educational measurements applications. *Journal of Applied Statistics 50*, 4 (2023), 889–908.

[22] Krishna, A., Maya, R., Chesneau, C., and Irshad, M. R. The unit Teissier distribution and its applications. *Mathematical and Computational Applications 27*, 1 (2022), 12.

[23] Kumaraswamy, P. A generalized probability density function for double-bounded random processes. *Journal of Hydrology 46*, 1-2 (1980), 79–88.

[24] Macdonald, P. D. M. Comments and Queries Comment on "An Estimation Procedure for Mixtures of Distributions" by Choi and Bulgren. *Journal of the Royal Statistical Society: Series B (Methodological) 33*, 2 (1971), 326–329.

[25] Mazucheli, J., Menezes, A. F. B., and Dey, S. The unit-Birnbaum-Saunders distribution with applications. *Chilean Journal of Statistics 9*, 1 (2018), 47–57.

[26] Mazucheli, J., Menezes, A. F. B., and Ghitany, M. E. The unit-Weibull distribution and associated inference. *Journal of Applied Probability and Statistics 13*, 2 (2018), 1–22.

[27] Mazucheli, J., Menezes, A. F. B., and Chakraborty, S. On the one parameter unit-Lindley distribution and its associated regression model for proportion data. *Journal of Applied Statistics 46*, 4 (2019), 700–714.

[28] Mazucheli, J., Menezes, A. F. B., Fernandes, L. B., de Oliveira, R. P., and Ghitany, M. E. The unit-Weibull distribution as an alternative to the Kumaraswamy distribution for the modeling of quantiles conditional on covariates. *Journal of Applied Statistics 47*, 6 (2020), 954–974.

[29] Menezes, A. F. B., Mazucheli, J., and Dey, S. The unit-logistic distribution: Different methods of estimation. *Pesquisa Operacional 38*, 3 (2018), 555–578.

[30] Mitnik, P. A., and Baek, S. The Kumaraswamy distribution: median-dispersion re-parameterizations for regression modeling and simulation-based estimation. *Statistical Papers 54*, 1 (2013), 177–192.

[31] Nasiru, S., Abubakari, A. G., and Chesneau, C. The arctan power distribution: Properties, quantile and modal regressions with applications to biomedical data. *Mathematical and Computational Applications 28*, 1 (2023), 25.

[32] Nasiru, S., Chesneau, C., Abubakari, A. G., and Angbing, I. D. Generalized unit half-logistic geometric distribution: Properties and regression with applications to insurance. *Analytics 2*, 2 (2023), 438–462.

[33] Nasiru, S., Chesneau, C., and Ocloo, S. K. The log-cosine-power unit distribution: A new unit distribution for proportion data analysis. *Decision Analytics Journal 10* (2024), 100397.

[34] Peña-Ramírez, F. A., Guerra, R. R., and Mafalda, C. P. The unit ratio-extended Weibull family and the dropout rate in Brazilian undergraduate courses. *PLOS ONE 18*, 11 (2023), e0290885.

[35] Ranneby, B. The maximum spacing method: An estimation method related to the maximum likelihood method. *Scandinavian Journal of Statistics 11*, 2 (1984), 93–112.

[36] Swain, J. J., Venkatraman, S., and Wilson, J. R. Least-Squares estimation of distribution functions in Johnson's translation system. *Journal of Statistical Computation and Simulations 29*, 4 (1988), 271–297.

[37] Topp, C. W., and Leone, F. C. A family of J-shaped frequency functions. *Journal of the American Statistical Association 50*, 269 (1955), 209–219.

[38] Wolny-Dominiak, A., and Trzęsiok, M. (4 September 2014) insurancedata: A collection of insurance datasets useful in risk classification in non-life insurance. *R package version 1.0* (accessed on 12 October 2022).

[39] Yildirim, E., Ilikkan, E. S., Gemeay, A. M., Makumi, N., Bakr, M. E., and Balogun, O. S. Power unit Burr-XII distribution: Statistical inference with applications. *AIP Advances 13*, 10 (2023), 105107.