



HAL
open science

Some new facts about the unit-Rayleigh distribution with applications

Christophe Chesneau, Farrukh Jamal

► **To cite this version:**

Christophe Chesneau, Farrukh Jamal. Some new facts about the unit-Rayleigh distribution with applications. 2020. hal-02902340

HAL Id: hal-02902340

<https://hal.science/hal-02902340v1>

Preprint submitted on 19 Jul 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Some new facts about the unit-Rayleigh distribution with applications

Christophe Chesneau¹ and Farrukh Jamal²

¹ *Université de Caen Normandie, LMNO, Campus II, Science 3, 14032, Caen, France*

¹ *Department of Statistics, Govt. S.A Postgraduate College Dera Nawab Sahib, Bahawalpur, Punjab, Pakistan*

Emails: ¹christophe.chesneau@gmail.com, ²drfarrukh1982@gmail.com

Abstract

This paper is devoted to new facts about the so-called unit-Rayleigh distribution. More precisely, we contribute by: (i) proving new stochastic ordering results, (ii) bringing new characteristics on the form analysis of its main probabilistic and reliability functions with an emphasis on the mode(s), (iii) exposing manageable expressions of the incomplete and probability weighted moments, (iv) investigating distributional properties of the order statistics, (v) giving the expression of the reliability coefficient, (vi) providing a tractable expression for the Tsallis entropy and (vii) proposing some bivariate unit-Rayleigh distributions. On a practical level, two data sets are analyzed and adjusted, revealing that the unit-Rayleigh distribution can be a better alternative to standard one-parameter unit distributions, such as the one-parameter Kumaraswamy, Topp-Leone, one-parameter beta, power and transmuted distributions.

Keywords: Unit-Rayleigh distribution; hazard rate function; incomplete moments; order statistics; estimation

AMS Subject Classification: 62G07; 62C05; 62E20

1 Introduction

In many applied scenarios, we are often confronted with the uncertainty of a phenomenon which can be quantified in a bounded range of values. Appropriate modelling must take this information into account for optimal use. As an immediate example, it is natural to model characteristics of the proportion type as a random variable (rv) with values in the “unit interval $(0, 1)$ ”, thus following a certain unit distribution, i.e., distribution with support $(0, 1)$. The unit-distributions are of particular interest because of the following argues: (i) any rv X with bounded support of the form $(0, a)$, with $a > 0$, can be rescaled on $(0, 1)$ as $Y = X/a$, and thus Y follows a certain unit-distribution, (ii) the unit-distributions allow to define general and simple families of continuous distributions through the composition techniques; if $F(x)$ denotes a cumulative distribution function (cdf) of a rv following a unit-distribution and $G(x)$ denotes the cdf of any continuous distribution with support denoted by A , then a valid cdf is given as $H(x) = F(G(x))$, $x \in \mathbb{R}$, defining a certain family of distributions with support A (see Al-Hussaini (2012) as pioneer reference, as well as Tahir et Cordeiro (2016) for a complete survey in this regard), and (iii) the unit-distributions play a central role to define regression models having some characteristic with unit-interval values (see Kieschnick and McCullough (2003) for the definitions of such regression models and Ferrari and Cribari-Neto (2004) for a focus on the one of most popular of them: the beta regression model).

Among the most useful unit-distributions with various number of parameters, there are the power distribution, beta distribution, Johnson distribution by Johnson (1949), Topp-Leone distribution by Topp and Leone (1955), Kumaraswamy distribution by Kumaraswamy (1980), unit-gamma distribution by Grassia (1977) and Tadikamalla (1981), unit-logistic distribution by Tadikamalla and Johnson (1982), simplex distribution by Barndorff-Nielsen and Jorgensen (1991), unit-Birnbaum-Saunders distribution by Mazucheli *et al.* (2018), exponentiated Kumaraswamy distribution by Lemonte *et al.* (2013), exponentiated Topp-Leone distribution by Pourdarvish *et al.* (2015), unit-Weibull distribution by Mazucheli *et al.* (2018) and Mazucheli *et al.* (2020), unit-Gompertz by Mazucheli *et al.* (2019), unit-Lindley distribution by Mazucheli *et al.* (2019), unit-inverse Gaussian distribution by Ghitany *et al.* (2019), composite quantile distributions by Rodrigues *et al.* (2019), unit-generalized half normal distribution by Korkmaz (2020), and unit modified Burr-III distribution by Haq *et al.* (2020).

In this paper, we concentrate on a simple one-parameter unit distribution, called the unit-Rayleigh distribution. Technically, this distribution is not new; it corresponds to a special case of the unit-Weibull distribution introduced in Mazucheli *et al.* (2018), and it is briefly presented as such in this reference. However, new researches on the unit-Rayleigh distribution have yielded interesting and surprising mathematical fruits that we share in this paper. Specifically, we provide: (i) new motivations of considering this special distribution, (ii) new stochastic ordering results involving the power distribution and some unlisted distributions of potential interest, (iii) new characteristics on the form analysis of the corresponding probability density function (pdf) (with a closed form for the mode) and hazard rate function (hrf), revealing an unexpected fitting capability of the unit-Rayleigh model, (iv) new manageable expressions and approximations of the incomplete and probability weighted moments involving the complementary error function, (v) some basics on the distributional properties of the order statistics including discussions on their moments, (vi) a simple expression for the reliability coefficient which is useful for estimation purposes, (vii) a tractable expression and approximation of the so-called Tsallis entropy and (viii) some bivariate extensions of the unit-Rayleigh distribution for two-dimensional modeling objectives. On the practical side, we analyze two data sets, showing that the unit-Rayleigh distribution can be a better alternative to standard one-parameter unit distributions, such as the one-parameter Kumaraswamy, Topp-Leone, one-parameter beta, power and transmuted distributions.

The paper organization is planned as follows. Section 2 recalls the unit-Rayleigh distribution, with some new facts related to its main functions. The main technical mathematical results are developed in Section 3. Applications are given in Section 4. Final notes are formulated in Section 5.

2 The unit-Rayleigh distribution

This section discusses some new facts regarding the primary essence of the unit-Rayleigh distribution.

2.1 Corresponding functions

As the basis, the unit-Rayleigh distribution is associated with the cdf given as

$$F(x) = \exp \left\{ -\beta [\log(x)]^2 \right\}, \quad x \in (0, 1), \quad (1)$$

where $\beta > 0$, and $F(x) = 0$ for $x \leq 0$ and $F(x) = 1$ for $x > 1$. Thus defined, it is a special case of the unit-Weibull distribution introduced by Mazucheli *et al.* (2019) with shape parameter

equal to 2. By construction, the unit-Rayleigh distribution is the distribution of the rv $\exp(-Y)$, where Y denotes a rv following the Rayleigh distribution with scale parameter β , i.e., with cdf $F_Y(x) = 1 - \exp(-\beta x^2)$, with $x > 0$, and $F_Y(x) = 0$ otherwise. The Rayleigh distribution also corresponds to the chi-squared distribution with two degrees of freedom. The basics and properties of this distribution can be found in Weisstein (2020).

As a new fact, the unit-Rayleigh distribution is also the distribution of the rv $1/Z$, where Z denotes a rv following the Benini distribution with truncated parameter equal to 1, i.e., with cdf $F_Z(x) = 1 - \exp\{-\beta [\log(x)]^2\}$, with $x > 1$, and $F_Z(x) = 0$ otherwise. The Benini distribution is a long tail distribution that can be viewed as a generalization of the Pareto distribution. We may refer to the former work of Benini (1905).

Based on $F(x)$, the pdf of the unit-Rayleigh distribution is given as

$$f(x) = \frac{d}{dx}F(x) = -\frac{2\beta}{x} \log(x) \exp\{-\beta [\log(x)]^2\}, \quad x \in (0, 1), \quad (2)$$

and $f(x) = 0$ for $x \notin (0, 1)$. The shape properties of this function are fundamental to evaluate the capability of the unit-Rayleigh model to fit data. This aspect will be discussed later.

As main reliability functions, the survival function is obtained by

$$\bar{F}(x) = 1 - F(x) = 1 - \exp\{-\beta [\log(x)]^2\}, \quad x \in (0, 1),$$

and $\bar{F}(x) = 1$ for $x \leq 0$ and $\bar{F}(x) = 0$ for $x > 1$, the cumulative hrf is specified as

$$H(x) = -\log[\bar{F}(x)] = -\log\left[1 - \exp\{-\beta [\log(x)]^2\}\right], \quad x \in (0, 1),$$

and $H(x) = 0$ for $x \leq 0$ and $H(x) = +\infty$ for $x > 1$, and the hrf is given as

$$h(x) = \frac{d}{dx}H(x) = -\frac{2\beta}{x} \log(x) \frac{\exp\{-\beta [\log(x)]^2\}}{1 - \exp\{-\beta [\log(x)]^2\}}, \quad x \in (0, 1), \quad (3)$$

and $h(x) = 0$ for $x \notin (0, 1)$. The shape properties of the hrf are precious indicators on some features of the unit-Rayleigh model. This point will be discussed later.

We end this part by specifying the quantile function of the unit-Rayleigh distribution obtained as

$$Q(x) = F^{-1}(x) = \exp\left\{-\left[-\frac{1}{\beta} \log(x)\right]^{1/2}\right\}, \quad x \in (0, 1). \quad (4)$$

2.2 Analysis of the cdf

Basically, $F(x)$ is an increasing and derivable function with respect to x for $x \in (0, 1)$. We can express it as a function of the power distribution cdf as

$$F(x) = F_*(x)^{-\log(x)},$$

where $F_*(x) = x^\beta$ for $x \in (0, 1)$, $F_*(x) = 0$ for $x \leq 0$ and $F_*(x) = 1$ for $x \geq 1$. The following inequalities can be deduced: For any $x \in (0, \exp(-1))$, we have $F(x) \leq x^\beta$, and for $x \in (\exp(-1), 1)$, the reverse inequality holds: $F(x) \geq x^\beta$.

In addition, we can remark that $F(x)$ is a decreasing function with respect to β ; by setting $F(x; \beta) = F(x)$, for any $\beta_2 \geq \beta_1$, we have

$$F(x; \beta_2) \leq F(x; \beta_1).$$

This inequality reveals a basic first-order stochastic dominance of the unit-Rayleigh distribution.

Moreover, one can note that, for any $x \in (0, 1)$,

$$\frac{\partial^2 F(x; \beta)}{\partial \beta^2} = [\log(x)]^4 \exp \left\{ -\beta [\log(x)]^2 \right\} > 0,$$

implying that $F(x; \beta)$ is a convex function with respect to β .

2.3 Analysis of the pdf

In this section, we analyze $f(x)$ as described in (2), also performing a mode(s) analysis. A such global analysis has been performed in Mazucheli *et al.* (2020) for the unit-Weibull distribution, in full generality. Here, we provide more specific details on this aspect for the unit-Rayleigh distribution, including the expression of the mode and its compartment when β varied.

As an alpha remark, let us note that, for any $\beta > 0$,

$$\lim_{x \rightarrow 0} f(x) = \lim_{x \rightarrow 1} f(x) = 0,$$

with the equivalence $f(x) \sim 2\beta(1-x)$ when $x \rightarrow 1$. Since it is positive, the function $f(x)$ is not monotonic; the points 0 and 1 are not modes of $f(x)$. Now, for $x \in (0, 1)$, we have

$$\frac{d}{dx} f(x) = \frac{2\beta}{x^2} \exp \left\{ -\beta [\log(x)]^2 \right\} \left[2\beta [\log(x)]^2 + \log(x) - 1 \right].$$

Therefore, a critical point for $f(x)$, say x_0 , satisfies $x_0 \in (0, 1)$ and $2\beta [\log(x_0)]^2 + \log(x_0) - 1 = 0$. After developments, we get

$$x_0 = \exp \left\{ -\frac{1}{4\beta} \left[\sqrt{8\beta + 1} + 1 \right] \right\}. \quad (5)$$

Let us now study the nature of this critical point. For any $x \in (0, 1)$, we have

$$\frac{d^2}{dx^2} f(x) = -\frac{2\beta}{x^3} \exp \left\{ -\beta [\log(x)]^2 \right\} (2\beta \log(x) + 1) \left[2\beta [\log(x)]^2 + 2\log(x) - 3 \right].$$

Since $(2\beta/x^3) \exp \left\{ -\beta [\log(x)]^2 \right\} > 0$, the sign of $d^2 f(x)/dx^2|_{x=x_0}$ is the one of

$$\eta = -(2\beta \log(x_0) + 1) \left[2\beta [\log(x_0)]^2 + 2\log(x_0) - 3 \right].$$

After developments, we obtain $\eta = -\sqrt{8\beta + 1} < 0$. We conclude that the point x_0 as defined by (5) is a maximum for the function $f(x)$; it is the (unique) mode of the unit-Rayleigh distribution. Therefore, the pdf of the unit-Rayleigh distribution is “more or less bell shape”.

Let us now discuss the behavior of this mode. By setting $x_0(\beta) = x_0$, we have

$$\frac{\partial}{\partial \beta} x_0(\beta) = \exp \left\{ -\frac{1}{4\beta} \left[\sqrt{8\beta + 1} + 1 \right] \right\} \frac{4\beta + \sqrt{8\beta + 1} + 1}{4\beta^2 \sqrt{8\beta + 1}} > 0,$$

implying that $x_0(\beta)$ is an increasing function with respect to β . Furthermore, we have

$$\lim_{\beta \rightarrow 0^+} x_0(\beta) = 0, \quad \lim_{\beta \rightarrow +\infty} x_0(\beta) = 1,$$

meaning that the mode can take all the values of the interval $(0, 1)$. This result indicates a certain flexibility of the unit-Rayleigh distribution regarding its mode. A graphical illustration of the possible shapes of $f(x)$ is provided in Figure 1, considering the following values for β : 0.1, 0.5, 0.8, 1.5 and 4.

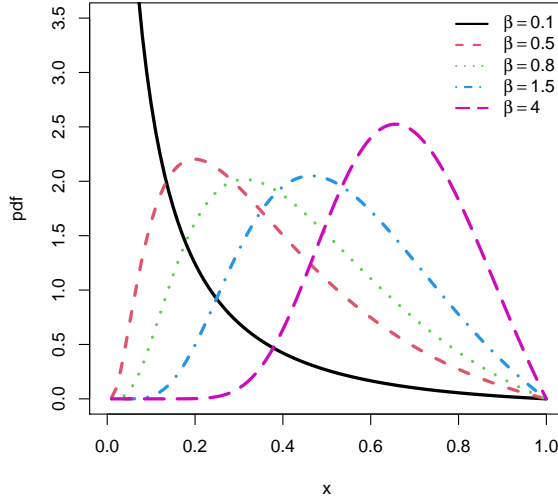


Figure 1: Several curves of the pdf of the unit-Rayleigh distribution.

We see in Figure 1 the “more or less bell shape” of the pdf, with an increasing mode according to β . Note that the black curve, corresponding to the pdf defined with $\beta = 0.1$, is “highly spiked”; the increasing curve does not appear in the figure because it is too sharp. In some senses, this graphical analysis completes the one performed in (Mazucheli *et al.*, 2018, Figure 1, last subfigure).

2.4 Analysis of the hrf

In this section, we analyze $h(x)$ as specified in (3). As far as we know, this aspect has been explored only graphically in (Mazucheli *et al.*, 2018, Figure 2, last subfigure). Some new facts are discussed below. Firstly, for any $\beta > 0$, we have

$$\lim_{x \rightarrow 0} h(x) = 0, \quad \lim_{x \rightarrow 1} h(x) = +\infty,$$

with the equivalence $h(x) \sim 2/(1-x)$ when $x \rightarrow 1$. Since $h(x)$ is positive, the point 0 is a minimum for $h(x)$. Now, for $x \in (0, 1)$, we have

$$\frac{d}{dx} h(x) = 2\beta \frac{w(x)}{x^2 \left[\exp \left\{ \beta [\log(x)]^2 \right\} - 1 \right]^2},$$

where

$$w(x) = 2\beta [\log(x)]^2 \exp\left\{\beta [\log(x)]^2\right\} - (1 - \log(x)) \left[\exp\left\{\beta [\log(x)]^2\right\} - 1\right],$$

being the difference of two positive functions.

In view of the denominator term, a critical point for $h(x)$, say x_1 , satisfies $x_1 \in (0, 1)$ and $w(x_1) = 0$. The study of this equation is not obvious from the analytical side. As results, we propose some alternative arguments showing that the hrf can be various forms.

Firstly, due the presence of the β in factor of the first term, by dominance, for any $x \in (0, 1)$, we have $\lim_{\beta \rightarrow +\infty} w(x) = +\infty$. This implies the existence of a β_* such that, for $\beta > \beta_*$, we have $w(x) > 0$ and, a fortiori, $dh(x)/dx > 0$. Hence, for $\beta > \beta_*$, $h(x)$ is an increasing function with respect to x . Now, assume that β is small, say $\beta \rightarrow 0$. Then, standard equivalences gives

$$w(x) \sim \beta [\log(x)]^2 (1 + \log(x)),$$

and the equivalence function is equal to 0 if $x = \exp(-1) \in (0, 1)$. Therefore, in the case where β is small enough, at least one critical point exists. This new fact reveals that the hrf is not only an increasing function, as one can think at first sight of (Mazucheli *et al.*, 2018, Figure 2, last subfigure); non monotonic shapes are possible for “small or not too small” β .

We illustrate this new fact by some plots of $h(x)$ in Figure 1, considering the following values for β : 0.1, 0.5, 0.8, 1.5 and 4.

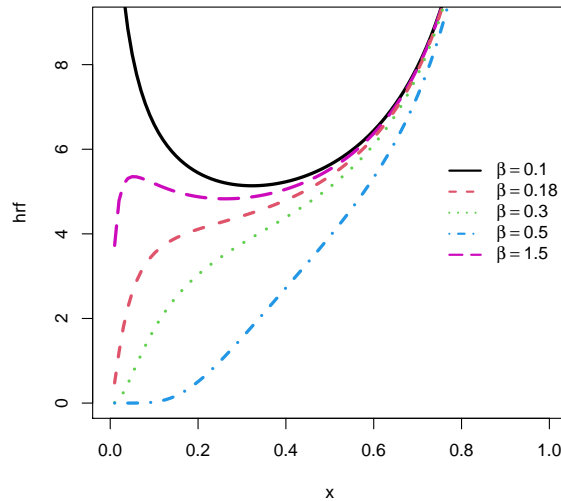


Figure 2: Several curves of the hrf of the unit-Rayleigh distribution.

We see in Figure 2 that the hrf can be increasing with convex and concave properties. Also, for the black line corresponding to the hrf defined with $\beta = 0.1$, a bathtub shape is observed. These observations confirm the flexible hazard rate of the unit-Rayleigh distribution.

3 New results

More mathematical results are developed in this section, all new.

3.1 Stochastic order results

We have already presented some stochastic order results involving the cdfs of the unit-Rayleigh and power distributions. More technical ones are described in the result below.

Proposition 3.1 *The following inequality holds:*

$$F_o(x) \leq F(x) \leq F_{oo}(x),$$

where

- $F_o(x) = x^{\beta(1/x-1)}$ for $x \in (0, 1)$, $F_o(x) = 0$ for $x \leq 0$ and $F_o(x) = 1$ for $x \geq 1$,
- $F_{oo}(x) = x^{\beta(1-x)}$ for $x \in (0, 1)$, $F_{oo}(x) = 0$ for $x \leq 0$ and $F_{oo}(x) = 1$ for $x \geq 1$,

both being cdfs of unit-distributions.

Proof. The inequalities are immediate for $x \leq 0$ and $x \geq 1$. For $x \in (0, 1)$, we can express $F(x)$ as $F(x) = x^{\beta[-\log(x)]}$. The desired result is a consequence of the following well-known inequalities: For $x \in (0, 1)$, we have $1 - 1/x \leq \log(x) \leq x - 1$. The claims that $F_o(x)$ and $F_{oo}(x)$ are valid cdfs are developed below. We have $\lim_{x \rightarrow 0} F_o(x) = \lim_{x \rightarrow 0} F_{oo}(x) = 0$, $\lim_{x \rightarrow 1} F_o(x) = \lim_{x \rightarrow 1} F_{oo}(x) = 1$, both are derivable for $x \in (0, 1)$ with, for $x \in (0, 1)$,

$$\frac{d}{dx} F_o(x) = \beta x^{\beta(1/x-1)-2} [1 - x - \log(x)] > 0, \quad \frac{d}{dx} F_{oo}(x) = \beta x^{\beta(1-x)-1} [1 - x - x \log(x)] > 0.$$

This concludes the proof of Proposition 3.1. □

As far as we know, the cdfs $F_o(x)$ and $F_{oo}(x)$ described in Proposition 3.1 are not listed in the literature. They can be of independent interest for purposes out the scope of this paper (modelling of proportion-type characteristics, constructions of new general families of continuous distributions...).

3.2 Incomplete moments

The incomplete moments of the unit-Rayleigh remain unexplored. We now fill this gap by providing their analytical expressions via comprehensive functions.

Proposition 3.2 *Let r be a nonnegative integer and X be a rv following the unit-Rayleigh distribution. Then, the r^{th} incomplete moment of X at $t \in (0, 1)$ is given as*

$$\begin{aligned} m_r(t) &= E(X^r I\{X \leq t\}) \\ &= t^r \exp\left\{-\beta [\log(t)]^2\right\} - \exp\left(\frac{r^2}{4\beta}\right) \frac{r}{\sqrt{\beta}} \frac{\sqrt{\pi}}{2} \operatorname{erfc}\left(-\sqrt{\beta} \log(t) + \frac{r}{2\sqrt{\beta}}\right), \end{aligned}$$

where $\operatorname{erfc}(a)$ is the complementary error function defined by $\operatorname{erfc}(a) = (2/\sqrt{\pi}) \int_a^{+\infty} \exp(-y^2) dy$, with $a \in \mathbb{R}$.

Proof. We recall that the unit-Rayleigh distribution corresponds to the one of the rv $\exp(-Y)$, where Y is a rv following the Rayleigh distribution with scale parameter β , i.e., with pdf $f_Y(x) =$

$2\beta x \exp(-\beta x^2)$, with $x > 0$, and $f_Y(x) = 0$ otherwise. Therefore, we have

$$\begin{aligned} m_r(t) &= E(\exp(-rY)I\{\exp(-Y) \leq t\}) = E(\exp(-rY)I\{Y \geq -\log(t)\}) \\ &= \int_{-\log(t)}^{+\infty} \exp(-rx) f_Y(x) dx = 2\beta \int_{-\log(t)}^{+\infty} x \exp(-rx - \beta x^2) dx \\ &= 2\beta \exp\left(\frac{r^2}{4\beta}\right) \int_{-\log(t)}^{+\infty} x \exp\left\{-\left(\sqrt{\beta}x + \frac{r}{2\sqrt{\beta}}\right)^2\right\} dx. \end{aligned}$$

By applying the change of variable $y = \sqrt{\beta}x + r/(2\sqrt{\beta})$, that is $x = [y - r/(2\sqrt{\beta})]/\sqrt{\beta}$, and performing some calculus, we obtain

$$\begin{aligned} m_r(t) &= 2\beta \exp\left(\frac{r^2}{4\beta}\right) \int_{-\sqrt{\beta}\log(t)+r/(2\sqrt{\beta})}^{+\infty} \frac{1}{\sqrt{\beta}} \left(y - \frac{r}{2\sqrt{\beta}}\right) \exp(-y^2) \frac{1}{\sqrt{\beta}} dy \\ &= 2 \exp\left(\frac{r^2}{4\beta}\right) \left\{ \left[-\frac{1}{2} \exp(-y^2)\right]_{-\sqrt{\beta}\log(t)+r/(2\sqrt{\beta})}^{+\infty} - \frac{r}{2\sqrt{\beta}} \int_{-\sqrt{\beta}\log(t)+r/(2\sqrt{\beta})}^{+\infty} \exp(-y^2) dy \right\} \\ &= \exp\left(\frac{r^2}{4\beta}\right) \exp\left\{-\left(-\sqrt{\beta}\log(t) + \frac{r}{2\sqrt{\beta}}\right)^2\right\} - \exp\left(\frac{r^2}{4\beta}\right) \frac{r}{\sqrt{\beta}} \frac{\sqrt{\pi}}{2} \operatorname{erfc}\left(-\sqrt{\beta}\log(t) + \frac{r}{2\sqrt{\beta}}\right) \\ &= t^r \exp\left\{-\beta [\log(t)]^2\right\} - \exp\left(\frac{r^2}{4\beta}\right) \frac{r}{\sqrt{\beta}} \frac{\sqrt{\pi}}{2} \operatorname{erfc}\left(-\sqrt{\beta}\log(t) + \frac{r}{2\sqrt{\beta}}\right). \end{aligned}$$

The result of Proposition 3.2 is obtained. \square

From Proposition 3.2, by taking $r = 0$, we obtain $m_0(t) = F(t) = \exp\left\{-\beta [\log(t)]^2\right\}$ with $t \in (0, 1)$. Also, the r^{th} raw moments of X can be derived as

$$m_r = E(X^r) = \lim_{t \rightarrow 1} m_r(t) = 1 - \exp\left(\frac{r^2}{4\beta}\right) \frac{r}{\sqrt{\beta}} \frac{\sqrt{\pi}}{2} \operatorname{erfc}\left(\frac{r}{2\sqrt{\beta}}\right).$$

We thus rediscover the formula in (Mazucheli *et al.*, 2018, Subsection 2.2).

In addition, the incomplete moments of X allow us to define an arsenal of interesting measures and functions involving the unit-Rayleigh distribution, such as mean deviations, mean residual life function, variance residual life function, reversed mean residual life function, Zenga curve, and so on. The complete list can be found in the book of Cordeiro *et al.* (2020), among others.

For approximation purposes of the incomplete moments of X , for any $a \in \mathbb{R}$, we can express and approximate $\operatorname{erfc}(a)$ as

$$\operatorname{erfc}(a) = 1 - \frac{2}{\sqrt{\pi}} \sum_{j=0}^{+\infty} \frac{(-1)^j}{(2j+1)j!} a^{2j+1} \approx 1 - \frac{2}{\sqrt{\pi}} \sum_{j=0}^J \frac{(-1)^j}{(2j+1)j!} a^{2j+1},$$

where J denotes a large integer. Let us mention that some more simple approximations of $\operatorname{erfc}(a)$ exist, with the assumptions that x is “small enough” or “large enough” (see Decker (1975)). On the other side, $\operatorname{erfc}(a)$ is implemented in all the modern mathematical softwares, making the computations of $m_r(t)$ straightforward.

The following result presents a new and simple series expansion of the incomplete moments, with direct integration; no existing results on $\operatorname{erfc}(x)$ is used. It thus provides an alternative expression to the one presented in Proposition 3.2.

Proposition 3.3 Under the setting of Proposition 3.2, for any $t \in (0, 1)$, the following series expansion holds:

$$m_r(t) = \sum_{j=0}^{+\infty} \frac{1}{j!} (-1)^j r^j \beta^{-j/2} \Gamma\left(\frac{j}{2} + 1, \beta[\log(t)]^2\right),$$

where $\Gamma(a, x)$ is the incomplete upper gamma function defined by $\Gamma(a, x) = \int_x^{+\infty} t^{a-1} \exp(-t) dt$, with $a, x > 0$.

Proof. By making the change of variable $x = Q(y)$ as defined as (4), we get

$$m_r(t) = \int_0^t x^r f(x) dx = \int_0^{F(t)} [Q(y)]^r dy = \int_0^{\exp\{-\beta[\log(t)]^2\}} \exp\left\{-r \left[-\frac{1}{\beta} \log(y)\right]^{1/2}\right\} dy.$$

By applying the Taylor series expansion of the exponential function, we obtain

$$m_r(t) = \sum_{j=0}^{+\infty} \frac{1}{j!} (-1)^j r^j \beta^{-j/2} \int_0^{\exp\{-\beta[\log(t)]^2\}} [-\log(y)]^{j/2} dy.$$

Now, the change of variable $y = \exp(-z)$ yields

$$\int_0^{\exp\{-\beta[\log(t)]^2\}} [-\log(y)]^{j/2} dy = \int_{\beta[\log(t)]^2}^{+\infty} z^{j/2} \exp(-z) dz = \Gamma\left(\frac{j}{2} + 1, \beta[\log(t)]^2\right).$$

The proof of Proposition 3.3 follows from the above equalities. □

From Proposition 3.3, by taking $r = 0$ with the convention $0^0 = 1$ in the sum, we rediscover $m_0(t) = F(t) = 1 \times \Gamma(1, \beta[\log(t)]^2) = \exp\{-\beta[\log(t)]^2\}$, with $t \in (0, 1)$. Also, the r^{th} raw moments of X can be derived as

$$m_r = E(X^r) = \lim_{t \rightarrow 1} m_r(t) = \sum_{j=0}^{+\infty} \frac{1}{j!} (-1)^j r^j \beta^{-j/2} \Gamma\left(\frac{j}{2} + 1\right),$$

where $\Gamma(a)$ is the standard gamma function defined by $\Gamma(a) = \int_0^{+\infty} t^{a-1} \exp(-t) dt$, with $a > 0$.

Also, the following simple finite sum approximation holds:

$$m_r(t) \approx \sum_{j=0}^J \frac{1}{j!} (-1)^j r^j \beta^{-j/2} \Gamma\left(\frac{j}{2} + 1, \beta[\log(t)]^2\right),$$

where J denotes a large integer.

3.3 Probability weighted moments

The probability weighted moments can be viewed as generalizations of raw moments. They appear quite naturally when we deal with the raw moments of order statistics. The closed forms of the probability weighted moments for the unit-Rayleigh distribution are given below.

Proposition 3.4 Let r and s be two nonnegative integers and X be a rv following the unit-Rayleigh distribution. Then, the $(r, s)^{th}$ probability weighted moment of X is given as

$$m_{r,s} = E[X^r F(X)^s] = \frac{\beta}{\beta + s} \left[1 - \exp\left(\frac{r^2}{4(\beta + s)}\right) \frac{r}{\sqrt{\beta + s}} \frac{\sqrt{\pi}}{2} \operatorname{erfc}\left(\frac{r}{2\sqrt{\beta + s}}\right) \right],$$

$\operatorname{erfc}(x)$ being the complementary error function.

Proof. First of all, based on (1) and (2), let us notice that

$$F(x)^s f(x) = -\frac{2\beta}{x} \log(x) \exp\left\{-\beta [\log(x)]^2\right\} = \frac{\beta}{\beta + s} f_*(x),$$

where $f_*(s)$ denotes the pdf of a rv Z following the unit-Rayleigh distribution with scale parameter $\beta + s$. Therefore

$$m_{r,s} = \int_0^1 x^r F(x)^s f(x) dx = \frac{\beta}{\beta + s} \int_0^1 x^r f_*(x) dx = \frac{\beta}{\beta + s} E(Z^r).$$

Owing to Proposition 3.2 with $\beta + s$ instead of β and $t \rightarrow 1$, we have

$$E(Z^r) = 1 - \exp\left(\frac{r^2}{4(\beta + s)}\right) \frac{r}{\sqrt{\beta + s}} \frac{\sqrt{\pi}}{2} \operatorname{erfc}\left(\frac{r}{2\sqrt{\beta + s}}\right).$$

By combining the two equalities above, we conclude the proof of Proposition 3.4. \square

Clearly, we have $m_r = m_{r,0}$. The probability weighted moments will find applications in the next section.

3.4 Order statistics

The modelling of several physical systems involved the use of order statistics. In this section, the basic properties of the order statistics of the unit-Rayleigh distribution are discussed. The theory and details on order statistics in a general setting can be found in David and Nagaraja (2003).

First, the pdf of the u^{th} order statistic of X in a random sample of size n from the unit-Rayleigh distribution, say $X_{(u)}$, can be expressed as

$$f_{X_{(u)}}(x) = \frac{n!}{(u-1)!(n-u)!} f(x) F(x)^{u-1} [1 - F(x)]^{n-u}, \quad x \in \mathbb{R}. \quad (6)$$

Hence, based on (1) and (2), for $x \in (0, 1)$, we can write

$$f_{X_{(u)}}(x) = -\frac{n!}{(u-1)!(n-u)!} \frac{2\beta}{x} \log(x) \exp\left\{-\beta u [\log(x)]^2\right\} \left[1 - \exp\left\{-\beta [\log(x)]^2\right\}\right]^{n-u}.$$

In particular, for the minimum and maximum order statistics, we get

$$f_{X_{(1)}}(x) = -n \frac{2\beta}{x} \log(x) \exp\left\{-\beta [\log(x)]^2\right\} \left[1 - \exp\left\{-\beta [\log(x)]^2\right\}\right]^{n-1}$$

and

$$f_{X_{(n)}}(x) = -n \frac{2\beta}{x} \log(x) \exp\left\{-\beta n [\log(x)]^2\right\},$$

respectively. The raw moments of $X_{(u)}$ can be simply expressed via the probability weighted moments of the former unit-Rayleigh distribution. Indeed, from (6) and the binomial formula, we can write

$$f_{X_{(u)}}(x) = \frac{n!}{(u-1)!(n-u)!} \sum_{j=0}^{n-u} \binom{n-u}{j} (-1)^j f(x) F(x)^{j+u-1}$$

and the r^{th} raw moment of $X_{(u)}$ is specified by

$$\begin{aligned} m_{(u),r} &= E(X_{(u)}^r) = \int_0^1 x^r f_{X_{(u)}}(x) dx \\ &= \frac{n!}{(u-1)!(n-u)!} \sum_{j=0}^{n-u} \binom{n-u}{j} (-1)^j \int_0^1 x^r f(x) F(x)^{j+u-1} dx \\ &= \frac{n!}{(u-1)!(n-u)!} \sum_{j=0}^{n-u} \binom{n-u}{j} (-1)^j m_{r,j+u-1}, \end{aligned}$$

where, by Proposition 3.4,

$$\begin{aligned} m_{r,j+u-1} &= \frac{\beta}{\beta + j + u - 1} \times \\ &\left[1 - \exp\left(\frac{r^2}{4(\beta + j + u - 1)}\right) \frac{r}{\sqrt{\beta + j + u - 1}} \frac{\sqrt{\pi}}{2} \operatorname{erfc}\left(\frac{r}{2\sqrt{\beta + j + u - 1}}\right) \right]. \end{aligned}$$

From the raw moments of $X_{(u)}$, several measures can be derived such as the skewness and kurtosis coefficients, L-moments, allowing to define the L-scale, L-skewness and L-kurtosis, among others.

3.5 Reliability coefficient

The reliability coefficient allows to study the behavior of various random systems. It is defined as the probability that a hierarchy exists between two characteristics of the system with unknown values a priori. All the details can be found in Surles and Padgett (2001). Here, we show that the reliability coefficient can be expressed in a simple manner for the unit-Rayleigh distribution.

Proposition 3.5 *Let U and V be two independent rvs following the unit-Rayleigh distribution with scale parameters β and β_* , respectively. Then, the corresponding reliability coefficient is defined by $R = P(U \leq V)$ and*

$$R = \frac{\beta_*}{\beta + \beta_*}.$$

Proof. Let $F(x; \beta) = F(x)$ be the cdf of U , $f(x; \beta_*) = f(x)$ be the pdf of V and $f_*(x)$ be the pdf of the unit-Rayleigh distribution with scale parameter $\beta + \beta_*$. Then, we have

$$\begin{aligned} R &= P(U \leq V) = \int_0^1 F(x; \beta) f(x; \beta_*) dx = \int_0^1 -\frac{2\beta_*}{x} \log(x) \exp\left\{-\frac{2(\beta + \beta_*)}{x} [\log(x)]^2\right\} dx \\ &= \frac{\beta_*}{\beta + \beta_*} \int_0^1 f_*(x) dx = \frac{\beta_*}{\beta + \beta_*}. \end{aligned}$$

This proved Proposition 3.5. □

From Proposition 3.5, we clearly have $R < 1/2$ for $\beta_* < \beta$, $R = 1/2$ for $\beta_* = \beta$, and $R > 1/2$ for $\beta_* > \beta$. The simple expression of R is useful for statistical aims. In particular, by the invariance property, maximum likelihood estimates of the parameters β and β_* , say $\hat{\beta}$ and $\hat{\beta}_*$, respectively, provide the maximum likelihood estimate for R given as

$$\hat{R} = \frac{\hat{\beta}_*}{\hat{\beta} + \hat{\beta}_*}.$$

The simple expression and the nice theoretical guarantees on this estimate are advantages.

3.6 Tsallis entropy

Let X be a rv following the unit-Rayleigh distribution. Then, for $\tau \neq 1$ and $\tau > 0$, the Tsallis entropy of X is given as

$$T_\tau = \frac{1}{\tau - 1} \left[1 - \int_0^1 f(x)^\tau dx \right].$$

Commonly, T_τ is a measure of randomness of X . One can refer to the study of Amigo *et al.* (2018) for discussions on the roles of various entropy measures in applied sciences, including the Tsallis entropy. The following result concerns a series expansion of this entropy measure.

Proposition 3.6 *Let $\tau \neq 1$ and $\tau > 0$. Then, we can express T_τ as*

$$T_\tau = \frac{1}{\tau - 1} \left[1 - 2^{\tau-1} \beta^{(\tau-1)/2} \tau^{-(\tau+1)/2} \sum_{j=0}^{+\infty} \frac{1}{j!} (\tau-1)^j \beta^{-j/2} \tau^{-j/2} \Gamma\left(\frac{j+\tau+1}{2}\right) \right],$$

where J denotes a large integer.

Proof. We only need to treat the integral term in the definition of T_τ . Owing to (2), we have

$$\int_0^1 f(x)^\tau dx = 2^\tau \beta^\tau \int_0^1 x^{-\tau} [-\log(x)]^\tau \exp\left\{-\beta\tau [\log(x)]^2\right\} dx.$$

Therefore, by making the change of variable $x = \exp(-y)$, i.e., $y = -\log(x)$, and by introducing a rv Y following the Rayleigh distribution with scale parameter $\beta\tau$, we get

$$\begin{aligned} \int_0^1 f(x)^\tau dx &= 2^\tau \beta^\tau \int_0^{+\infty} \exp[(\tau-1)y] y^\tau \exp(-\beta\tau y^2) dy \\ &= 2^{\tau-1} \beta^{\tau-1} \tau^{-1} \int_0^{+\infty} \exp[(\tau-1)y] y^{\tau-1} [2\beta\tau y \exp(-\beta\tau y^2)] dy \\ &= 2^{\tau-1} \beta^{\tau-1} \tau^{-1} E\left\{\exp[(\tau-1)Y] Y^{\tau-1}\right\}. \end{aligned}$$

Now, by using the Taylor expansion series of the exponential function and the following well-known moment properties of the Rayleigh distribution: For any $v > -2$, $E(Y^v) = \beta^{-v/2} \tau^{-v/2} \Gamma(v/2 + 1)$, we have

$$\begin{aligned} E\left\{\exp[(\tau-1)Y] Y^{\tau-1}\right\} &= \sum_{j=0}^{+\infty} \frac{1}{j!} (\tau-1)^j E(Y^{j+\tau-1}) \\ &= \beta^{-(\tau-1)/2} \tau^{-(\tau-1)/2} \sum_{j=0}^{+\infty} \frac{1}{j!} (\tau-1)^j \beta^{-j/2} \tau^{-j/2} \Gamma\left(\frac{j+\tau+1}{2}\right). \end{aligned}$$

By putting the above equalities together, we obtain

$$\int_0^1 f(x)^\tau dx = 2^{\tau-1} \beta^{(\tau-1)/2} \tau^{-(\tau+1)/2} \sum_{j=0}^{+\infty} \frac{1}{j!} (\tau-1)^j \beta^{-j/2} \tau^{-j/2} \Gamma\left(\frac{j+\tau+1}{2}\right).$$

The desired result follows by substituting this series expansion into the former definition of T_r . \square

Also, from Proposition 3.6, the following approximation holds:

$$T_r \approx \frac{1}{\tau-1} \left[1 - 2^{\tau-1} \beta^{(\tau-1)/2} \tau^{-(\tau+1)/2} \sum_{j=0}^J \frac{1}{j!} (\tau-1)^j \beta^{-j/2} \tau^{-j/2} \Gamma\left(\frac{j+\tau+1}{2}\right) \right],$$

where J is a large enough integer.

3.7 Some bivariate unit-Rayleigh distributions

Now, we present some motivated ideas to construct bivariate unit-Rayleigh distributions, which are of interest for the modelling of conjoint characteristics with values on the unit interval. In order to keep a control on the structure of the marginal rvs, we propose to use the special probabilistic functions called copulas (see Nelsen (2006)).

As a first approach, we can define the Farlie-Gumbel-Morgenstern unit-Rayleigh distribution by the following cdf:

$$F(x, y) = F(x; \beta)F(y; \beta_*) + \lambda F(x; \beta)F(y; \beta_*) [1 - F(x; \beta)] [1 - F(y; \beta_*)], \quad (x, y) \in \mathbb{R}^2,$$

where $\lambda \in [-1, 1]$, $F(x; \beta)$ and $F(y; \beta_*)$ are defined as (1) with the scale parameters β and β_* , respectively. Hence, for $(x, y) \in (0, 1)^2$, we have

$$F(x, y) = \exp\left\{-\beta [\log(x)]^2 - \beta_* [\log(y)]^2\right\} + \lambda \exp\left\{-\beta [\log(x)]^2 - \beta_* [\log(y)]^2\right\} \times \\ \left[1 - \exp\left\{-\beta [\log(x)]^2\right\}\right] \left[1 - \exp\left\{-\beta_* [\log(y)]^2\right\}\right].$$

By taking $\lambda = 0$, the marginal rvs are independent.

Similarly, one can also define a bivariate unit-Rayleigh distribution by using the Clayton copula. Thus, we can define the Clayton unit-Rayleigh distribution by the following cdf:

$$F_*(x, y) = \left[\max\left(F(x; \beta)^{-\lambda} + F(y; \beta_*)^{-\lambda} - 1, 0\right) \right]^{-1/\lambda}, \quad (x, y) \in \mathbb{R}^2,$$

where $\lambda \geq -1$ and $\lambda \neq 0$, $F(x; \beta)$ and $F(y; \beta_*)$ are defined as (1) with the scale parameters β and β_* , respectively. Thus, for $(x, y) \in (0, 1)^2$, we have

$$F_*(x, y) = \left[\max\left(\exp\left\{\lambda\beta [\log(x)]^2\right\} + \exp\left\{\lambda\beta_* [\log(y)]^2\right\} - 1, 0\right) \right]^{-1/\lambda}.$$

As last example, we can define the Gumbel unit-Rayleigh distribution by the following cdf:

$$F_{**}(x, y) = \exp\left\{-\left[(-\log[F(x; \beta)])^\lambda + (-\log[F(y; \beta_*)])^\lambda\right]^{1/\lambda}\right\}, \quad (x, y) \in \mathbb{R}^2,$$

where $\lambda \geq 1$, $F(x; \beta)$ and $F(y; \beta_*)$ are defined as (1) with the scale parameters β and β_* , respectively. Hence, for $(x, y) \in (0, 1)^2$, we have

$$F_{**}(x, y) = \exp\left\{-\left[\beta^\lambda [\log(x)]^{2\lambda} + \beta_*^\lambda [\log(y)]^{2\lambda}\right]^{1/\lambda}\right\}.$$

These bivariate extensions need further investigations in an applied setting, that we leave for a future work.

4 Applications

This section shows the applicability behavior of the unit-Rayleigh distribution in a data analysis framework, which has not received a particular attention in Mazucheli *et al.* (2018) or Mazucheli *et al.* (2020). We estimate the parameter β by the maximum likelihood method, as done in Mazucheli *et al.* (2018) but by putting the shape parameter equal to 2; it is not to be estimated. In this case, from n observations of a rv following the unit-Rayleigh distribution, say x_1, \dots, x_n , the maximum likelihood estimate of β is defined by

$$\hat{\beta} = \left\{ \frac{1}{n} \sum_{i=1}^n [\log(x_i)]^2 \right\}^{-1}.$$

Also, based on (1), (2) and (3), the estimated cdf, pdf and hrf are obtained by substituting β by $\hat{\beta}$ in their own expressions.

Thus, with the maximum likelihood method, we aim to compare the fit behavior of the unit-Rayleigh distribution with those of the following well-known one-parameter competitors.

- the one-parameter Kumaraswamy (Ku) distribution (or special Lehmann type II power distribution) defined by the cdf given as

$$F_{Ku}(x) = 1 - (1 - x)^\alpha, \quad x \in (0, 1),$$

$F_{Ku}(x) = 0$ for $x \leq 0$ and $F_{Ku}(x) = 1$ for $x \geq 1$, with $\alpha > 0$. See Kumaraswamy (1980).

- the Topp-Leone (TL) distribution defined by the cdf specified by

$$F_{TL}(x) = x^\theta(2 - x)^\theta, \quad x \in (0, 1),$$

$F_{TL}(x) = 0$ for $x \leq 0$ and $F_{TL}(x) = 1$ for $x \geq 1$, where $\theta > 0$. See Topp and Leone (1955).

- the one-parameter beta (B) distribution defined by the cdf given as

$$F_B(x) = \frac{1}{B(\mu, 2)} \int_0^x t^{\mu-1}(1-t)dt, \quad x \in (0, 1),$$

$F_B(x) = 0$ for $x \leq 0$ and $F_B(x) = 1$ for $x \geq 1$, where $\mu > 0$ and $B(a, b) = \int_0^1 t^{a-1}(1-t)^{b-1}dt$, $a, b > 0$.

- the power (P) distribution defined by the cdf expressed by

$$F_P(x) = x^\eta, \quad x \in (0, 1),$$

$F_P(x) = 0$ for $x \leq 0$ and $F_P(x) = 1$ for $x \geq 1$, where $\eta > 0$.

- the transmuted (TM) distribution defined by the cdf expressed by

$$F_{TM}(x) = (1 + \lambda)x - \lambda x^2, \quad x \in (0, 1),$$

$F_{TM}(x) = 0$ for $x \leq 0$ and $F_{TM}(x) = 1$ for $x \geq 1$, where $\lambda \in [-1, 1]$. We may refer to Shaw and Buckley (2009) all the characteristics of the transmuted distribution.

These distributions can be assimilated to semi-parametric statistical models for adjustment purposes. The following classic criteria are used to compare the fits: minus estimated log-likelihood ($-\hat{\ell}$), consistent Akaike information criterion (CAIC), Hannan-Quinn information criterion (HQIC), Akaike information criterion (AIC), Bayesian information criterion (BIC), Cramer-von Mises criterion (W) and Anderson-Darling criterion (A) are computed. The lower the values of these criteria, the better the fit. The R software developed by R Development Core Team (2005) is used, with the help of the R function `goodness.fit` function from the package `AdequacyModel` (see *Marinho et al.* (2019)).

First of all, we consider data of times to infection of kidney dialysis patients as described by Klein and Moeschberger (2006). These “times of infection” data set is: {2.5, 2.5, 3.5, 3.5, 3.5, 4.5, 5.5, 6.5, 6.5, 7.5, 7.5, 7.5, 7.5, 8.5, 9.5, 10.5, 11.5, 12.5, 12.5, 13.5, 14.5, 14.5, 21.5, 21.5, 22.5, 22.5, 25.5, 27.5}. Now, we make a normalization operation by divided these data by 30, to get data between 0 and 1. The transformed data set becomes: {0.08333333, 0.08333333, 0.11666667, 0.11666667, 0.11666667, 0.15000000, 0.18333333, 0.21666667, 0.21666667, 0.25000000, 0.25000000, 0.25000000, 0.25000000, 0.28333333, 0.31666667, 0.35000000, 0.38333333, 0.41666667, 0.41666667, 0.45000000, 0.48333333, 0.48333333, 0.71666667, 0.71666667, 0.75000000, 0.75000000, 0.85000000, 0.91666667}.

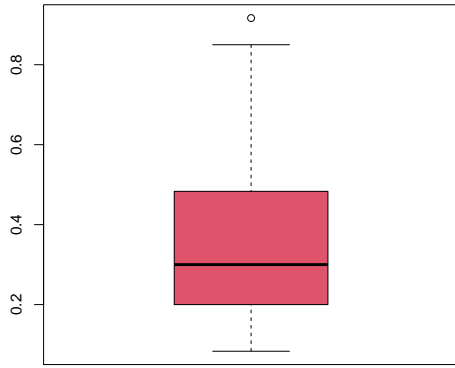
The second data set concerns the failure times of the air conditioning system of an air plane as reported in Linhart and Zucchini (1986). These “failure times” data set is: {23, 261, 87, 7, 120, 14, 62, 47, 225, 71, 246, 21, 42, 20, 5, 12, 120, 11, 3, 14, 71, 11, 14, 11, 16, 90, 1, 16, 52, 95}. Again, we make a normalization operation by divided these data by 265, to get data between 0 and 1. That is, we work with the following data set: {0.086792453, 0.984905660, 0.328301887, 0.026415094, 0.452830189, 0.052830189, 0.233962264, 0.177358491, 0.849056604, 0.267924528, 0.928301887, 0.079245283, 0.158490566, 0.075471698, 0.018867925, 0.045283019, 0.452830189, 0.041509434, 0.011320755, 0.052830189, 0.267924528, 0.041509434, 0.052830189, 0.041509434, 0.060377358, 0.339622642, 0.003773585, 0.060377358, 0.196226415, 0.358490566}.

The two transformed data sets are basically analyzed in Table 1.

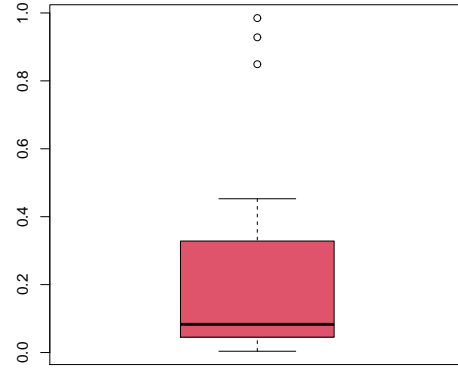
Table 1: Descriptive analyzes for the times of infection and failure times data sets.

| | n | Mean | Median | Variance | Skewness | Kurtosis | Min | Max |
|--------------------|-----|------|--------|----------|----------|----------|-------|------|
| Times of infection | 28 | 0.38 | 0.3 | 0.06 | 0.72 | -0.75 | 0.08 | 0.92 |
| Failure times | 30 | 0.22 | 0.08 | 0.07 | 1.61 | 1.64 | 0.003 | 0.98 |

Table 1 indicates that the times of infection data is right-skewed, with small dispersion and negative kurtosis. This point means that the curve of the unknown pdf behind these data is flatter than a normal pdf. Concerning the failure times data, we can say that is “significantly” right-skewed, with small dispersion and “significant” kurtosis. So the nature of the two data sets differ on numerous aspects. This is also illustrated through the corresponding boxplots in Figure 3, presenting different quantiles characteristics. Note that some extreme points are present.



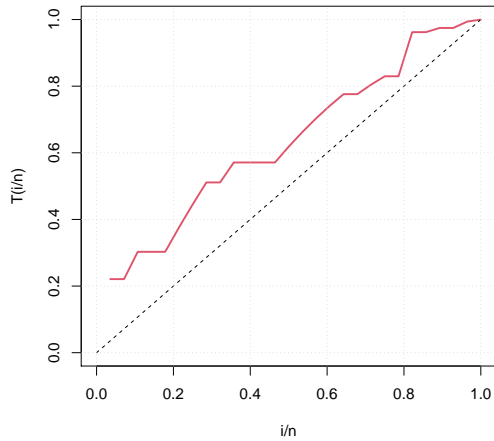
(a)



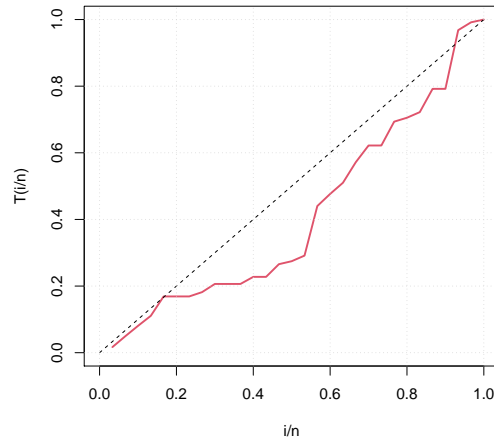
(b)

Figure 3: Boxplots of the (a) times of infection data set and (b) failure times data set.

We complete the first statistical analysis by the total time on test (TTT) plots of the two data sets in Figure 4.



(a)



(b)

Figure 4: TTT plots of the (a) times of infection data set and (b) failure times data set.

We see that the TTT curve for the times of infection data set is concave indicating an increasing failure intensity, whereas the TTT curve for the failure times data set is convex then concave suggesting a U-shape failure intensity. Thus, these TTT plots highlighted the different nature of the failure intensity of these two data sets. It should also be noted that the increasing and U-shaped failure intensities are covered by the unit-Rayleigh model, which makes it suitable for more suitable analyzes of these data sets.

The quality of fit measurements for the models, as well as the MLEs and standard errors (SEs) of the parameters involved are collected in Tables 2 and 3 for the times of infection and failure times data sets, respectively.

Table 2: Goodness-of-fit measures, MLEs and SEs for the times of infection data set.

| Model | $-\hat{\ell}$ | CAIC | HQIC | AIC | BIC | W | A | MLEs (SEs) |
|---------------|---------------|---------|---------|---------|---------|--------|--------|------------|
| UR | -4.4825 | -6.8111 | -6.557 | -6.9650 | -5.6328 | 0.0556 | 0.3832 | 0.5221 |
| (β) | | | | | | | | (0.0986) |
| Ku | -3.0686 | -3.9834 | -3.7300 | -4.1373 | -2.8051 | 0.1109 | 0.6897 | 1.6615 |
| (α) | | | | | | | | (0.3140) |
| TL | -3.8524 | -5.551 | -5.2975 | -5.704 | -4.3726 | 0.1066 | 0.6678 | 1.3778 |
| (θ) | | | | | | | | (0.2603) |
| B | -3.7584 | -5.3629 | -5.1095 | -5.5168 | -4.1846 | 0.1097 | 0.6839 | 1.3085 |
| (μ) | | | | | | | | (0.2151) |
| TM | -2.9334 | -3.7131 | -3.4596 | -3.8669 | -2.5347 | 0.0963 | 0.6172 | 0.7936 |
| (λ) | | | | | | | | (0.2721) |

Table 3: Goodness-of-fit measures, MLEs and SEs for the failure times data set.

| Model | $-\hat{\ell}$ | CAIC | HQIC | AIC | BIC | W | A | MLEs (SEs) |
|-------------|---------------|----------|----------|----------|----------|---------|--------|------------|
| UR | -12.7730 | -23.4033 | -23.0979 | -23.5461 | -22.1449 | 0.1253 | 0.7933 | 0.1497 |
| (β) | | | | | | | | (0.0273) |
| Ku | -7.5378 | -12.9330 | -12.627 | -13.0759 | -11.6747 | 0.2153 | 1.3759 | 2.2333 |
| (α) | | | | | | | | (0.4077) |
| TL | -11.9801 | -21.8175 | -21.5121 | -21.9603 | -20.5591 | 0.2379 | 1.5102 | 0.6017 |
| (θ) | | | | | | | | (0.1098) |
| B | -12.0261 | -21.9094 | -21.6040 | -22.0523 | -20.6511 | 0.23084 | 1.4687 | 0.6228 |
| (μ) | | | | | | | | (0.1061) |
| P | -12.7018 | -23.2607 | -22.9553 | -23.4036 | -22.0024 | 0.2068 | 1.3212 | 0.4501 |
| (η) | | | | | | | | (0.0821) |
| TM | -8.4186 | -14.6944 | -14.3890 | -14.8373 | -13.4361 | 0.1764 | 1.1390 | 0.8688 |
| (λ) | | | | | | | | (0.1318) |

From Tables 2 and 3, the unit-Rayleigh model can be considered as the best model for the two data sets, because it has the smallest values for the CAIC, HQIC, AIC, BIC, W and A statistics. Figures 5 and 6 confirm this claim through a graphical approach. In them, we plot the estimated pdfs over the adequate histograms and the estimated cdfs over the adequate empirical cdfs for the times of infection and failure times data sets, respectively.

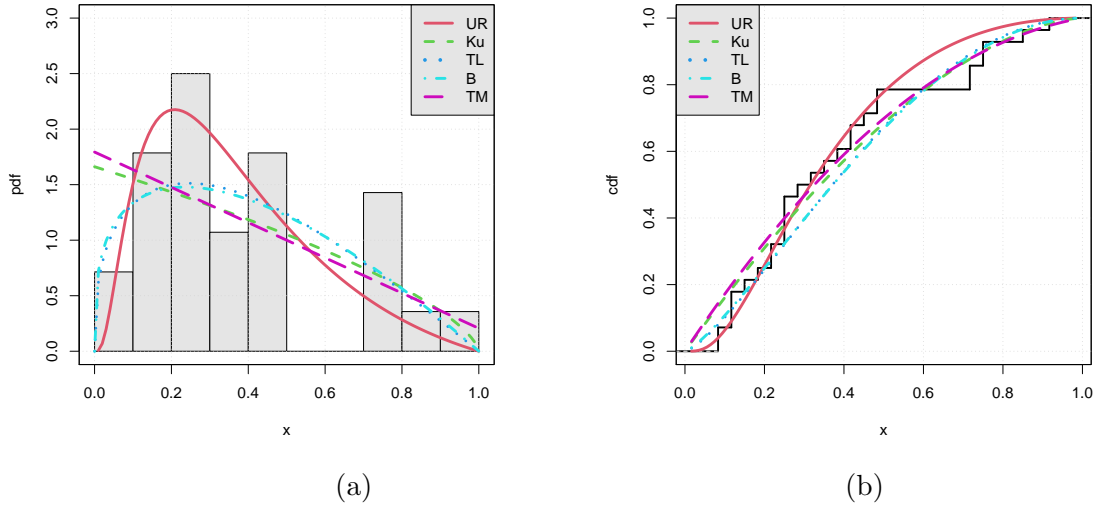


Figure 5: Plots of the (a) estimated pdfs and (b) estimated cdfs for the times of infection data set.

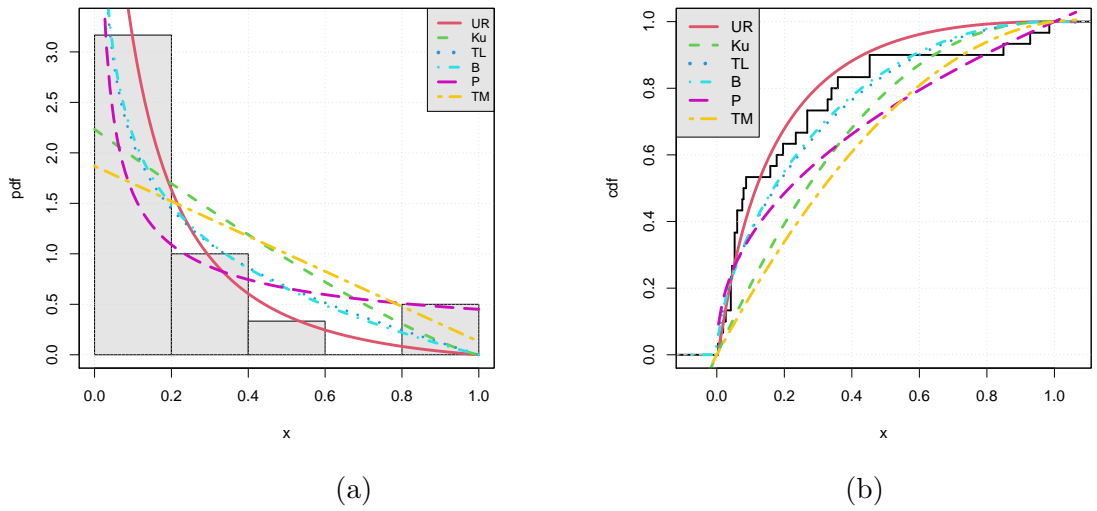


Figure 6: Plots of the (a) estimated pdfs and (b) estimated cdfs for the failure times data set.

Figures 5 and 6 show the nice fits of the unit-Rayleigh model, which has captured the main characteristics of the data contrary to most of the competitors. We complete this graphical analysis by plotting the estimated hrfs of the unit-Rayleigh model only in Figure 7.

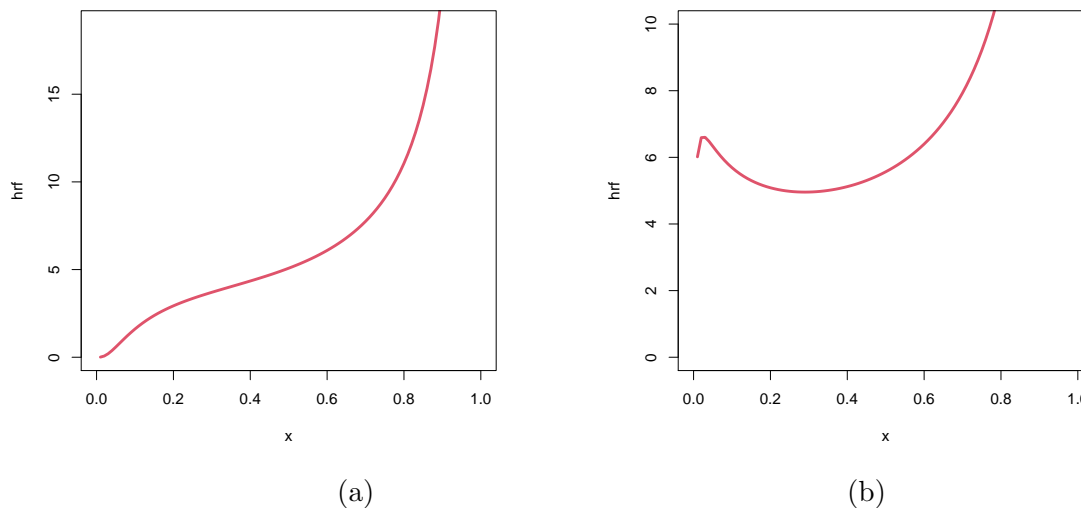


Figure 7: Plots of the estimated hrfs for the (a) times of infection data set and (b) failure times data set.

As expected with the TTT plots in Figure 4, based on the unit-Rayleigh model, Figure 7 shows an increasing estimated hrf for the times of infection data set, and a U-shape estimated hrf for the failure time data set. We thus see the importance of the possible U-shape of the hrf of the unit-Rayleigh distribution as evoked above for such a modelling.

All the preceding points highlight the undeniable capacities of the unit-Rayleigh model in the adjustment of various data.

5 Final notes

In this article, we highlight new facts about the unit-Rayleigh distribution, including new motivations, interests, theoretical results, and practical benefits. The most relevant facts are a detailed analysis of the main functions, results on some stochastic ordering, the expressions of the incomplete and probability weighted moments, as well as those of the Tsallis entropy and reliability coefficient, various properties on the order statistics, and a list of potential bivariate extensions. An applied work show how the unit-Rayleigh distribution can be used in practice. Also, we show that it can be superior to others well-reputed one-parameter unit distributions for some data sets.

A possible continuation of this work may be the use of the unit-Rayleigh distribution for the construction of general families of distributions, through composition techniques or others, and the construction of regression models including characteristics with values on the unit interval. All of these studies remain to be developed; we leave it for future investigations.

References

Al-Hussaini, E.K. (2012). Composition of cumulative distribution functions, *Journal of Statistical Theory and Applications*, 11, 333-336.

- Amigo, J.M., Balogh, S.G. and Hernandez, S. (2018). A brief review of generalized entropies, *Entropy*, 20, 813.
- Barndorff-Nielsen, O. and Jorgensen, B. (1991). Some parametric models on the Simplex, *Journal of Multivariate Analysis*, 39, 1, 106-116.
- Benini, R. (1905). I diagrammi a scala logaritmica (a proposito della graduazione per valore delle successioni ereditarie in Italia, Francia e Inghilterra), *Giornale degli Economisti, Serie II*, 16, 222-231.
- Cordeiro, G.M., Silva, R.B. and Nascimento, A.D.C. (2020). Recent advances in lifetime and reliability models, Bentham books, DOI:10.2174/97816810834521200101.
- David, H.A. and Nagaraja, H. (2003). *Order Statistics*, 3rd Edn. John Wiley and Sons, New York.
- Decker, D.L. (1975). Computer evaluation of the complementary error function, *American Journal of Physics*, 43, 833-834.
- Ferrari, S. and Cribari-Neto, F. (2004). Beta regression for modelling rates and proportions, *Journal of Applied Statistics*, 31, 799-815.
- Ghitany, M.E., Mazucheli, J., Menezes, A.F.B. and Alqallaf, F. (2019). The unit-inverse Gaussian distribution: A new alternative to two-parameter distributions on the unit interval, *Communications in Statistics - Theory and Methods*, 48:14, 3423-3438.
- Grassia, A. (1977). On a family of distributions with argument between 0 and 1 obtained by transformation of the Gamma distribution and derived compound distributions, *Australian journal of statistics*, 19, 108-114.
- Haq, M.A., Hashmi, S., Aidi, K., Ramos, P.L. and Louzada, F. (2020). Unit modified Burr-III distribution: estimation, characterizations and validation test, *Annals of Data Science*, <https://doi.org/10.1007/s40745-020-00298-6>
- Johnson, N.L. (1949). Systems of frequency curves generated by methods of translation, *Biometrika*, 36, 149-176.
- Kieschnick, R. and McCullough, B.D. (2003). Regression analysis of variates observed on $(0, 1)$: Percentages, proportions and fractions, *Statistical Modelling*, 3, 193-213.
- Klein, J.P. and Moeschberger, M.L. (2006). *Survival analysis: techniques for censored and truncated data*, Springer, Berlin.
- Korkmaz, M.C. (2020). The unit generalized half normal distribution: a new bounded distribution with inference and applications, *UPB Scientific Bulletin, Series A: Applied Mathematics and Physics*, 82, 2, 133-140.
- Kumaraswamy, P. (1980). A generalized probability density function for double-bounded random processes, *Journal of Hydrology*, 46, 79-88.
- Lemonte, A.J., Barreto-Souza, W. and Cordeiro, G.M. (2013). The exponentiated Kumaraswamy distribution and its log-transform, *Brazilian Journal of Probability and Statistics*, 27, 1, 31-53.
- Linhart, H. and Zucchini, W. (1986). *Model Selection*, Wiley, New York, United States.

- Marinho, P.R.D., Silva, R.B., Bourguignon, M., Cordeiro, G.M. and Nadarajah, S. (2019). AdequacyModel: An R package for probability distributions and general purpose optimization, PLoS ONE, 14, 8, e0221487.
- Mazucheli, J., Menezes, A.F. and Dey, S. (2019) Unit-Gompertz distribution with applications, *Statistica*, 79, 1, 25-43.
- Mazucheli, J., Menezes, A.F.B. and Ghitany, M.E. (2018). The unit-Weibull distribution and associated inference, *Journal of Applied Probability and Statistics*, 13, 1-22.
- Mazucheli, J., Menezes, A.F.B., Fernandes, L.B., de Oliveira, R.P. and Ghitany, M.E. (2020). 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, 954-974.
- Mazucheli, J., Menezes, A.F.B. and Chakraborty, S. (2019). On the one parameter unit-Lindley distribution and its associated regression model for proportion data, *Journal of Applied Statistics*, 46, 700-714.
- Mazucheli, J., Menezes, A.F. and Dey, S. (2018). The unit-Birnbaum-Saunders distribution with applications, *Chilean Journal of Statistics*, 9, 47-57.
- Nelsen, R.B. (2006). An introduction to copulas, 2nd edition, Springer-Verlag, Berlin.
- Pourdarvish, A., Mirmostafae, S.M.T.K. and Naderi, K. (2015). The exponentiated Topp-Leone distribution: Properties and application, *Journal of Applied Environmental and Biological Sciences*, 5, 78, 251-256.
- R Development Core Team (2005). R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL: <http://www.R-project.org>.
- Rodrigues, J., Bazán, J.L. and Suzuki, A.K. (2019). A flexible procedure for formulating probability distributions on the unit interval with applications, *Communications in Statistics - Theory and Methods*, DOI: 10.1080/03610926.2018.1549254
- Shaw, W.T. and Buckley, I.R. (2009). The alchemy of probability distributions: beyond Gram-Charlier expansions, and a skew-kurtotic-normal distribution from a rank transmutation Map. arXiv preprint arXiv:0901.0434.
- Surles, J.G. and Padgett, W.J. (2001). Inference for reliability and stress-strength for a scaled Burr-type X distribution, *Lifetime Data Analysis*, 7, 187-200.
- Tadikamalla, P.R. (1981). On a family of distributions obtained by the transformation of the gamma distribution, *Journal of Statistical Computation and Simulation*, 13, 209-214.
- Tadikamalla, P.R. and Johnson, N.L. (1982). Systems of frequency curves generated by transformations of logistic variables, *Biometrika*, 69, 2, 461-465.
- Tahir, M.H. and Cordeiro, G.M. (2016). Compounding of distributions: a survey and new generalized classes, *Journal of Statistical Distributions and Applications*, 3, 13, 35 p.
- Topp, C.W. and Leone, F.C. (1955). A Family of J-shaped frequency functions, *Journal of the American Statistical Association* 50, 209-219.

Weisstein, E.W. (2020). Rayleigh Distribution, From MathWorld–A Wolfram Web Resource.
<https://mathworld.wolfram.com/RayleighDistribution.html>