International Journal of Advanced Statistics and Probability, 2 (2) (2014) 42-47
©Science Publishing Corporation
www.sciencepubco.com/index.php/IJASP
doi: 10.14419/ijasp.v2i2.2563
Research Paper

Is there an absolutely continuous random variable with equal probability density and cumulative distribution functions in its support? Is it unique? What about in the discrete case?

Ernesto J. Veres Ferrer, Jose M. Pavía *

Department of Applied Economics, Universitat De Valencia, 46022-Valencia, Spain
**Corresponding author E-mail: pavia@uv.es*

Copyright © 2014 Ernesto J. Veres Ferrer, Jose M. Pavía. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

This paper inquires about the existence and uniqueness of a univariate continuous random variable for which both cumulative distribution and density functions are equal and asks about the conditions under which a possible extrapolation of the solution to the discrete case is possible. The issue is presented and solved as a problem and allows to obtain a new family of probability distributions. The different approaches followed to reach the solution could also serve to warn about some properties of density and cumulative functions that usually go unnoticed, helping to deepen the understanding of some of the weapons of the mathematical statistician's arsenal.

Keywords: *Cumulative Distribution Function; Density Function; Elasticity; Mathematical Statistics; Reversed Hazard Rate*.

1. Introduction

This paper studies the existence and uniqueness of a univariate continuous random variable for which both cumulative distribution and density functions are equal and inquires about the possible extrapolation of the solution to the discrete case. The topic is presented as a problem and serves to warn about some properties of the density and cumulative functions that usually go unnoticed. This research therefore helps to deepen the understanding of some of the weapons of the mathematical statistician's arsenal. Indeed, the solution to the question can be addressed from a strictly mathematical point of view, as a solution of a differential equation, and from a probabilistic standpoint, either directly or via the characteristic function. As a result of the inquiries a new family of density functions is found. We delve into the properties satisfied by this random variable as well as relate this to other well-known random variables, helping to place this new distribution.
This study could be also used for teachers of statistics to enrich a mathematical statistical program by using an example that can be discussed at different points of the course, allowing them to direct the class to some exclusive areas of the program and even to introduce students in the field of statistical research by relating the problem to some statistical tools less known such as the reverse hazard rate and the elasticity of a random variable.

2. Continuous case

The answer in the continuous case is (roughly) affirmative. There is a location parametric family of continuous random variables with equal density and cumulative distribution. The random distribution is unique conditional to the location parameter.
The unitary exponential random variable Δ upwardly bounded by τ has equal cumulative F_Δ and density f_Δ functions:

$$f_\Delta = F_\Delta = \begin{cases} e^{(x-\tau)} & x \le \tau \\ 0 & x > \tau \end{cases}, \tag{1}$$

And its characteristic function is given by:

$$\varphi_{\Delta} = e^{it\tau}\frac{1}{1+it} \tag{2}$$

As an example, Figure 1 shows the density (and cumulative) function [1] for $\tau = 3$.

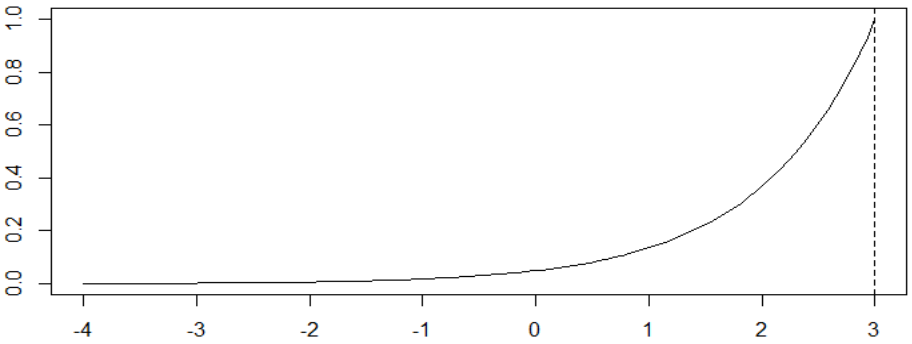


Fig. 1: Density Function of the Unitary Exponential Random Variable Upwardly Bounded By $\tau = 3$

The parameter τ sets the upper limit for the support of Δ. When $\tau = 0$ we will call the resulting variable the negative unitary exponential random variable, Ξ.

Furthermore, this family belongs to the broader family of (two-parameter) exponential random variables of parameter λ (> 0) upwardly bounded by τ. These random variables, Ψ, have density, cumulative and characteristic functions given, respectively, by:

$$f_{\Psi} = \begin{cases} \lambda \cdot e^{\lambda(x-\tau)} & x \le \tau \\ 0 & x > \tau \end{cases} \qquad F_{\Psi} = \begin{cases} e^{\lambda(x-\tau)} & x \le \tau \\ 1 & x > \tau \end{cases} \qquad \varphi_{\Psi} = e^{it\tau}\frac{\lambda}{\lambda+it}, \tag{3}$$

Where the scale parameter λ modulates the growth rate of the density function of Ψ. For instance, in a survival context or in reliability engineering (see, e.g., Khorashadizadeh et al. [1]), where a controlled event has been detected at moment τ and therefore it is known that it has occurred at an instant $t \le \tau$, the parameter λ graduates the risk of (actual) occurrence of the controlled event at previous instants. Larger values of λ express higher risks of occurrence close to τ, the detection instant.

2.1. Proof

The solution will be proved following three alternative strategies. We will need to lay down two previous results before, however.

Lemma 1 *Let f(x) is a density function of a continuous random variable with support D. If f(x) is non-decreasing in D then D is necessarily upwardly bounded.*

Proof: By reductio ad absurdum let us assume that D is not upwardly bounded.

Then *f(x)* must be upwardly bounded, because otherwise given *M (>0)*, x_M will exist such as $f(x) \ge M \ \forall x \ge x_M$ and therefore:

$$1 = \int_{-\infty}^{+\infty} f \cdot dx = \int_{-\infty}^{x_M} f \cdot dx + \int_{x_M}^{+\infty} f \cdot dx \ge \int_{-\infty}^{x_M} f \cdot dx + \int_{x_M}^{+\infty} M \cdot dx = +\infty$$

An issue that is impossible. Hence, $\exists k \mid f(x) \le k \ \forall x \in D$.

Now if x_k existed such as $\forall x \ge x_k \ \ f(x) = k$, it would follow:

$$1 \ge \int_{x_k}^{+\infty} f \cdot dx = \int_{x_k}^{+\infty} k \cdot dx = +\infty$$

This, again, is impossible. Consequently, *f(x)* must tend asymptotically to *k*, $\lim_{x \to +\infty} f(x) = k$. Therefore, given $\varepsilon > 0 \ \exists x_{\varepsilon} \mid k - f(x) < \varepsilon \ \forall x > x_{\varepsilon} \to f(x) > k - \varepsilon \ \forall x > x_{\varepsilon}$ and follows:

$$1 \ge \int_{x_{\varepsilon}}^{+\infty} f \cdot dx \ge \int_{x_{\varepsilon}}^{+\infty} (k - \varepsilon) \cdot dx = +\infty$$

This, once again, is impossible. ■

Proposition 1: *If a random variable has equal density and cumulative functions then it is upwardly bounded*.

Proof: It follows directly from Lemma 1 given that (i) density and cumulative functions are equal and (ii) cumulative functions are always non-decreasing.
So, we will denote by τ the upper limit of *D*, *D* being the support of a random variable with equal density and cumulative functions. ■

Proof 1: By hypothesis

$$f_\Delta(x) = F_\Delta(x) \rightarrow f_\Delta(x) = \frac{d^r}{dx^r} F_\Delta(x) \quad x \leq \tau \ \forall r \in \aleph \tag{4}$$

Hence following, for instance, Elsgoltz [2] we have that the differential equation given by (4), subject to the condition that $F_\Delta(\tau) = 1$, is given by (1). This solution is moreover unique. Thus, the family of unitary exponential upwardly bounded by $\tau\ (\in \Re)$ random variables is the only one whose elements have the same density and cumulative functions.■

Proof 2: As it is well-known (see, e.g., Veres-Ferrer and Pavía [3]), the reversed hazard function univocally characterises a random variable. In the continuous case, the reversed hazard rate, *r(x)*, is defined as the quotient between density and cumulative functions (see, e.g., Chechile [4]). In our case, therefore, *r(x) = 1* in *D* and moreover the variable has τ as upper limit. Hence, using equation (3) in Veres-Ferrer and Pavía [3], where cumulative and reversed hazard functions are related, it follows that:

$$F(x) = \exp\left\{-\int_x^\tau r(u)\cdot du = -\int_x^\tau du\right\} = e^{x-\tau}$$

This confirms the previous result and its uniqueness. ■

Proof 3: Proposition 2. *The characteristic function of an absolutely continuous random variable θ with identical density f_θ and cumulative F_θ functions is*:

$$\phi_\vartheta = e^{it\tau}(1+it)^{-1} \tag{5}$$

Proof By Proposition 1, θ is bounded by $\tau < +\infty$ and, as it is well-known for any random variable, $\lim_{x \to -\infty} F_\theta(x) = 0$.

Thus:

$$\phi_\theta = \int_{-\infty}^\tau e^{itx}\cdot f(x)\cdot dx = \int_{-\infty}^\tau e^{itx}\cdot f'(x)\cdot dx = \left| \begin{array}{l} e^{itx} = u \rightarrow it\cdot e^{itx}\cdot dx = du \\ f'(x)\cdot dx = dv \rightarrow f(x) = v = F(x) \end{array} \right| =$$

$$= \left(e^{itx}\cdot F(x)\right)_{-\infty}^{\tau} - \int_{-\infty}^\tau it\cdot e^{itx}\cdot f(x)\cdot dx = e^{it\tau} - it\cdot\phi_\theta$$

, and solving for ϕ_ϑ the result follows. ■

Comparing (2) and (5), it is again confirmed that the unitary exponential upwardly bounded by $\tau\ (\in \Re)$ random variable is the only one with equal density and cumulative functions.
The equality satisfied between density and distribution functions of Δ entails a perfect memory of the random variable. At each point of its domain, the probability density accounts for the cumulative probability up to it. This involves a continual recall of its past behaviour.

2.2. Further results

1) The (ordinary) moments of Δ could be obviously obtained by successive derivations of its moment generating function, $m_T(s) = e^{s\tau}(1+s)^{-1}$. However, in this case, they could also be reached by recurrence directly from the mathematical expressions of the integration by parts. In particular, denoting by α_j the j*th* ordinary moment, we have that $\alpha_j = \tau^j - j\cdot\alpha_{j-1}$ and that therefore its two first ordinary moments are $\alpha_1 = \mu_\Delta = \tau - 1$ and $\alpha_2 = (\tau-1)^2 + 1$.
Likewise, regarding central moments, we have that $\delta_j = 1 - j\cdot\delta_{j-1}$ (being δ_j the j*th* central moment), from which it follows that regardless of the parameter τ the variance of Δ is constant and equal to one, $\sigma_\Delta^2 = 1$. Indeed, given the above recurrence equation it is straightforward to see that no central moment depends on τ. This is particularly true for skewness and kurtosis: which are respectively -2 and 9, confirming the strong negative asymmetry and the leptokurtic nature of the unitary exponential upwardly bounded random distribution.

2) Obtaining the median is also straightforward. This is $Med_\Delta = \tau - ln\,2$. Furthermore, the main coefficients for the negative unitary exponential random variable, Ξ, are $\mu_\Xi = -1$, $\alpha_{2\Xi} = 2$, $\sigma_\Xi^2 = 1$ and $Med_\Xi = -ln\,2$.

3) If Δ follows a *unitary exponential random variable upwardly bounded by* τ and it is known that Δ is upper bounded by $\vartheta(<\tau)$, then ϑ is the new upper bound of the conditional variable. This new parameter plays the same role as the former parameter in the non-conditional distribution of Δ. In other words, given $\vartheta < \tau$ and Δ distributed as unitary exponential upwardly bounded by τ, it is true that $f(x|\Delta \le \vartheta) = e^{x-\vartheta}, x \le \vartheta$.

Proof. For $x \le \vartheta$, $P(\Delta \le x|\Delta \le \vartheta) = P(\Delta \le x)/P(\Delta \le \vartheta) = e^{x-\tau}(e^{\vartheta-\tau})^{-1} = e^{x-\vartheta}$. ■

4) If from a Bayesian standpoint we consider τ as a random variable, we have that when Δ and τ are distributed as unitary exponential random variables upwardly bounded by τ and κ, respectively, then the posterior distribution of τ conditional to $\Delta=\vartheta$ follows a uniform distribution on the interval $[\vartheta, \kappa]$.

Proof: Given that $f(x|\tau) = e^{x-\tau}, x \le \tau$ and $f(\tau) = e^{\tau-\kappa}, \tau \le \kappa$; it follows that $f(x,\tau) = e^{x-\kappa}, x \le \tau \le \kappa$, and $f(x) = \int_x^\kappa e^{x-\kappa} d\tau = (\kappa - x)e^{x-\kappa}$; from which, we obtain $f(\tau|\vartheta) = f(\tau,\vartheta)/f(\vartheta) = \frac{1}{\kappa-\vartheta}$ *with* $\vartheta \le \tau \le \kappa$. ■

5) If *W* follows a uniform distribution on [0, e^τ], then the bi-univocal transformation *T=ln W* follows a unitary exponential random variable upwardly bounded byτ.

Proof: $F_T(t) = P(T \le t) = P(\ln W \le t) = P(W \le e^t) = \int_0^{e^t} \frac{1}{e^\tau} \cdot dw = e^{t-\tau}$ with support $[-\infty, \tau]$. ■

Consequently, if F_V is the cumulative distribution function of a random variable *V*, the transformed variable $T=ln\ F_V$ is negative unitary exponential random variable Ξ, given that the probability integral transform random variable $Y = F_V$ is uniform on [0, 1] (see, e.g., DeGroot [5], p. 148).

6) If *Y* follows an exponential distribution of parameter λ (> 0) left-truncated by τ, then -*Y* follows an exponential random variable of parameter λ upwardly bounded by -τ.

Proof: -*Y* has $]-\infty, -\tau]$ as support and its cumulative function is $P(-Y \le y) = P(Y \ge -y) = e^{-\lambda(-y-\tau)} = e^{\lambda(y+\tau)}$. ■

7) If Δ follows a unitary exponential random variables upwardly bounded byτ, then $\Gamma = \tau - \Delta$ follows an exponential distribution with (rate) parameter $\lambda = 1$.

Proof: It is clear that the support of Γ is $[0, +\infty[$ and that as well $P(\Gamma \le y) = P(\tau - \Delta \le y) = P(\Delta \ge \tau - y) = 1 - e^{\tau - y - \tau} = 1 - e^{-y}$. ■

Furthermore, given 6), this property could be re-stated as bidirectional: Δ is unitary exponential upwardly bounded by $\tau \Leftrightarrow -\Delta$ is exponential distribution with parameter $\lambda = 1$ and left-truncated by -τ.

8) The characteristic function (3), $\varphi_\Psi = e^{it\tau}\lambda(\lambda + it)^{-1}$, derives directly from 6).

Proof: This follows on taking into account that (i) the characteristic function of *Y*, an exponential distribution of parameter λ (> 0) left-truncated by γ, is $\phi_Y = e^{it\tau}\lambda(\lambda - it)^{-1}$ and (ii) Y and Ψ are linearly related by $\Psi = aY + b$, with a=-1 y b=0. ■

9) If ξ follows an exponential distribution with parameter λ=1, and Ξ is an independent negative unitary exponential random variable, then the density function of ξ - Ξ is $f_{\xi-\Xi}(s) = s \cdot e^{-s}, s \ge 0$.

Proof: Example 2 in DeGroot [5], p. 160. ■

Furthermore, the density of ξ + Ξ is $f_{\xi-\Xi}(s) = \frac{1}{2}e^{-|s|}, -\infty < s < +\infty$.

Proof: Exercise 6 in DeGroot [5] (pp. 162 and 669), given that by 7) -Ξ is an exponential of rate λ=1. ■

For the same reason, if $\Xi_1, \Xi_2, \ldots, \Xi_n$ are independent negative unitary exponential random variables, then (-1)[$\Xi_1 + \Xi_2 + \ldots + \Xi_n$] follows a gamma distribution of shape *n* and rate one: $(-1)\sum_{i=1}^n \Xi_i \sim \Gamma(n,1)$.

Proof: this follows from the fact that (i) an exponential with rate λ is a gamma with shape 1 and rate λ and (ii) the sum of independent gammas with equal rate parameters is a gamma whose shape is the sum of the shapes (see Theorem 3 in DeGroot [5], p. 275). ■

Likewise, the random variable η = max [$\Xi_1, \Xi_2, \ldots, \Xi_n$] is distributed as an exponential random variable of parameter n upwardly bounded by 0.

Proof: This result comes from the fact that the minimum of independent exponentially distributed random variables is also exponentially distributed with rate being the sum of the rates (see exercise 7 in DeGroot [5], p. 278). ■

10) If Δ follows a unitary exponential Δ upwardly bounded by τ (>0) random variable, then its elasticity (Veres-Ferrer and Pavía [3]) is linear and decreasing for $x \leq 0$ and linear and increasing in $[0,\tau]$: $e_{\Delta}(x) = |x|$ for $x \leq \tau$.

Proof: This follows from the definition of the elasticity itself. ■

In particular, if $\tau \leq -1$, Δ is elastic in its complete support: any increase in the value of the variable results in a greater increase in their cumulative probability; except at $x = -1$, where elasticity is unitary. If $-1 < \tau \leq 0$, Δ is inelastic for $-1 < x < 0$, being perfectly inelastic in $x = 0$: any increase in the value of the variable results in a smaller increase in their accumulated probability. The same situation occurs for $0 < x < 1$, when $\tau > 0$, whereas Δ becoming elastic for $x > 1$, if $\tau > 1$. The elasticity of the random variable Ξ is always decreasing, evolving from perfect elasticity to perfect inelasticity.

3. Discrete case

In the same way as in the continuous case, in the discrete case there is a family of location parametric discrete random variables whose probability in each point of its support is equal to the probability accumulated up to that point[1], being the uniqueness conditional to the location parameter.

The discrete random variable η_τ with mass function defined as:

$$p_x = P(\eta_\tau = x) = 2^{x-\tau-1} \quad \forall x \leq \tau \quad x,\tau \in Z \tag{6}$$

Satisfies in the discrete case the equivalent condition of the continuous case, (4):

$$P(\eta_\tau = x) = p_x = P(\eta_\tau < x) = \sum_{j<x} p_j \tag{7}$$

Expressions in (6) and (7) obviously verify $p_X < p_{X+1} \ \forall x \leq \tau \ \ x,\tau \in Z$ and clearly (6) is a mass function given that $p_X > 0$ on its support and that:

$$\sum_{x\leq\tau} 2^{x-\tau-1} = \sum_{x\geq-\tau} 2^{-x-\tau-1} = 2^{-\tau-1} \sum_{x\geq-\tau} 2^{-x} = \sum_{x\geq 1} 2^{-x} = 1$$

Likewise, the characteristic function of η_τ is given by:

$$\varphi_{\eta_\tau}(s) = \sum_{\forall x\leq\tau} e^{isx} 2^{x-\tau-1} = \sum_{\forall x\geq-\tau} e^{-isx} 2^{-x-\tau-1} = 2^{-\tau-1} \sum_{\forall x\geq-\tau} (2e^{is})^{-x} = \frac{2^{-\tau-1} 2^{\tau} e^{is\tau}}{1-\frac{1}{2e^{is}}} = \frac{e^{is(\tau+1)}}{2e^{is}-1} \tag{8}$$

3.1. Proof

In the same way that in the continuous case, the condition (7) entails that the random variable must be upwardly bounded.

Lemma 2 *If ξ is a discrete random variable with increasing mass function, $p_x < p_{x+1} \ \forall x$, then exists $\tau \in Z$ such us $x \leq \tau \ \forall x$ (where without loss of generality it could be assumed that the support of any discrete random variable is included in Z).*

Proof: By reduction ad absurdum, Let assume that ξ is defined $\forall x \in Z$, then for $0 < \varepsilon < 1$ x_k would exist such as $p_k > \varepsilon$ $\forall k \geq x_\varepsilon$ and consequently:

$$1 = \sum_{\forall k} p_k = \sum_{k<x_\varepsilon} p_k + \sum_{k\geq x_\varepsilon} p_k > \sum_{k<x_\varepsilon} p_k + \sum_{k\geq x_\varepsilon} \varepsilon = +\infty$$

Something obviously impossible. ■

Once we have seen this previous finding, the solution will be stated through three alternative proofs.

Proof 1: Proposition 3. *Under (7), the characteristic function of ξ is* $\phi_\xi(s) = \frac{2p_\tau e^{is(\tau+1)}}{2e^{is}-1}$ (9)

[1] It should be noted that in the discrete case we cannot use the statement 'with equal mass and cumulative functions', given that for each x in the support of the random variable we have $P(X<x) \neq P(X\leq x) = F(x)$.

Proof Using Lemma 1 and the fact that mass and cumulative functions are equal we have that $\varphi_\xi(s)=\sum_{\forall x\le\tau}p_x e^{isx}=\sum_{\forall x\le\tau}\left(\sum_{\forall j\le x-1}p_j\right)e^{isx}=\sum_{\forall x\le\tau-1}p_x\left(\sum_{j=x+1}^{\tau}e^{isj}\right)$, and observing that the second summation is the sum of a geometric series of ratio e^{is}, we have that:

$$\sum_{\forall x\le\tau-1}p_x\frac{e^{is(\tau+1)}-e^{is(x+1)}}{e^{is}-1}=p_\tau\frac{e^{is(\tau+1)}}{e^{is}-1}-\frac{e^{is}}{e^{is}-1}\sum_{\forall x\le\tau-1}p_x e^{isx}=p_\tau\frac{e^{is(\tau+1)}}{e^{is}-1}-\frac{e^{is}}{e^{is}-1}\left(\varphi_\xi(s)-p_\tau e^{is\tau}\right)$$

From which we obtain $2p_\tau\frac{e^{is(\tau+1)}}{e^{is}-1}=\varphi_\xi(s)\left(1+\frac{e^{is}}{e^{is}-1}\right)$ that leads to the searched result solving for $\phi_\xi(s)$. ■

Corollary. *The random variable defined in (6) is the only one that satisfies (7).*

Proof Given that necessarily $p_\tau=2^{\tau-\tau-1}=\frac{1}{2}$, it is direct to state the coincidence between the characteristic functions of ξ (equation (9)) and η_τ (equation (8)). ■

Proof 2: The uniqueness can be checked directly from the fact that:

$$P(\eta_\tau<x)=\sum_{\forall j<x}p_j=\sum_{\forall j<x}2^{j-\tau-1}=\sum_{\forall j>-x}2^{-j-\tau-1}=2^{-\tau-1}\sum_{\forall j>-x}2^{-j}=\frac{2^{-\tau-1}2^{x-1}}{1-\frac{1}{2}}=2^{x-\tau-1}=P(\eta_\tau=x)=p_x\ \blacksquare$$

Proof 3: This result follows straightforward using the bi-univocal relationship that in discrete random variables links the quotient function and the cumulative distribution function (Section 4 in Veres-Ferrer and Pavía [6]). In particular, taking into account the upward limit and that the quotient function of a random variable satisfying (7) is unitary, (6) is obtained applying backward recursively the equation of page 48 in Veres-Ferrer and Pavía [6].

3.2. Further results

1) By successive derivations of the characteristic function of η_τ it is simple to obtain the different moments. In particular, $\alpha_1=\mu_{\eta_\tau}=\tau-1$, $\alpha_2=(\tau-1)^2+1$ and $\sigma^2_{\eta_\tau}=1$.

2) Likewise, similar to the continuous case, the elasticity of η_τ is $\forall x\le\tau\ \ x,\tau\in Z$:

$$e_i=\frac{\left(2^{i+1-\tau-1}-2^{i-\tau-1}\right)\Big/2^{i-\tau-1}}{1\Big/|x|}=|x|$$

References

[1] M. Khorashadizadeh, A.H. Rezaei Roknabadi, G.R. Mohtashami Borzadaran, Characterization of Life Distributions Using Log-odds Rate in Discrete Aging, *Communications in Statistics - Theory and Methods* 42 (2013) 76-87.

[2] L. Elsgoltz, *Ecuaciones diferenciales y cálculo variacional*, Mir, Moscow, 1969.

[3] E.J. Veres-Ferrer, J.M. Pavía, On the relationship between the reversed hazard rate and elasticity, *Statistical Papers* 55 (2014) 275-284.

[4] R.A. Chechile, Properties of reverse hazard functions, *Journal of Mathematical Psychology* 55 (2011) 203–222.

[5] M.H. DeGroot, *Probabilidad y Estadística*, Addison-Wesley Iberoamericana, Mexico, 1988.

[6] E.J. Veres-Ferrer, J.M. Pavía, La elasticidad: una nueva herramienta para caracterizar distribuciones de probabilidad, *Rect@* 13 (2012) 145-158.