

Article

A Note on Ordering Probability Distributions by Skewness

V. J. García ^{1,†}, M. Martel-Escobar ^{2,†} and F. J. Vázquez-Polo ^{2,*,†}

¹ Department of Statistics and Operations Research, University of Cádiz, 11002 Cádiz, Spain; victoriano.garcia@uca.es

² Department of Quantitative Methods & TiDES Institute, University of Las Palmas de Gran Canaria, 35017 Las Palmas de Gran Canaria, Las Palmas, Spain; maria.martel@ulpgc.es or francisco.vazquezpolo@ulpgc.es

* Correspondence: francisco.vazquezpolo@ulpgc.es; Tel.: +34-928-451-800

† These authors contributed equally to this work.

Received: 15 June 2018; Accepted: 10 July 2018; Published: 16 July 2018



Abstract: This paper describes a complementary tool for fitting probabilistic distributions in data analysis. First, we examine the well known bivariate index of skewness and the aggregate skewness function, and then introduce orderings of the skewness of probability distributions. Using an example, we highlight the advantages of this approach and then present results for these orderings in common uniparametric families of continuous distributions, showing that the orderings are well suited to the intuitive conception of skewness and, moreover, that the skewness can be controlled via the parameter values.

Keywords: positive and negative skewness; ordering; fitting distributions

MSC: 62E10; 62P99

1. Introduction

Detailed knowledge of the characteristics of probability models is desirable (if not essential) if data are to be modeled properly. In studying these properties, many authors have considered orderings within probability distribution families, according to diverse measuring criteria. The usual approach taken by researchers in this field is to evaluate or measure one or more theoretical characteristics of a given distribution and to study the effect produced by the value of its parameters on this measurement. In actuarial science, stochastic orders are widely used in order to make risk comparisons [1].

Some parametric distributions can be ordered according to the evaluation made of a given property, merely by comparing some of its parameters. Although most related orders are actually preorders, each one presents interesting applications. Many studies have been conducted in this area, and the following are particularly significant: Lehmann (1955) [2], which is of seminal importance; Arnold (1987) [3], who compared random variables according to stochastic ordering in a particular Lorenz order; Shaked and Shanthikumar (2006) [1], on stochastic orders; Nanda and Shaked (2001) [4], on reversed hazard rate orders; Ramos-Romero and Sordo-Díaz (2001) [5], on the likelihood ratio order; and Gupta and Aziz (2010) [6], on convex orders.

In this paper, we study the relationship between the skewness of some parametric distributions and the value of one of their parameters. The first question to be addressed is that of measuring the skewness. In this respect, Oja (1981) [7] introduced a set of axioms to be verified by any measurement of skewness considered. These axioms were established for indexes of skewness with one main constraint: that the skewness of a distribution should be evaluated by a single real number. This point is discussed below.

Many authors have proposed and obtained different descriptive elements to measure skewness (see, for instance, [8–13]). Ref [10] suggested a measurement of skewness corresponding to the (unique) mode, M , given by the following index:

$$\gamma_M(F) = 1 - 2F(M). \quad (1)$$

Ref [10] applied this index to ordering the gamma, log-logistic, lognormal and Weibull families of distributions by their skewness, taking into account the feasible values of their respective parameters. Index (1), which is proven to satisfy those axioms derived from Oja (1981) [7], is also recommended in [14] as a (very) good index of skewness. However, notice that (1) only compares the probability weight on the left side of a central point (the mode) with the value 1/2, but it does not account for how the weights are distributed to each side of the centre.

García et al. (2015) [15] introduced some further elements to be incorporated into the list of skewness measurements of a probability distribution. According to these authors, given a unimodal probability distribution $F(x)$, its skewness is considered to be a local function of a given distance, z , from the mode, M . For such a distance, and given the interval $[M - z, M + z]$, the aggregate skewness function, $\nu_F(z)$, compares the probability weight of F at either side of the interval:

$$\nu_F(z) = \Pr(X > M + z) - \Pr(X < M - z), \quad (2)$$

where $z \geq 0$. Thus, the (maximum) right skewness of the distribution F and its (minimum) left skewness are respectively given by

$$S^+(F) = \max_{z \geq 0} \nu_F(z), \quad S^-(F) = \min_{z \geq 0} \nu_F(z). \quad (3)$$

The distances, z_p and z_n , where these extreme values are achieved, are termed the critical distances to the mode. As the skewness function is bounded inside the interval $[-1, 1]$ and $\nu_F(\infty) = 0$, the bivariate index $(S^-(F), S^+(F))$ belongs to $[-1, 0] \times [0, 1]$. A given distribution function F such that $\nu_F(z) \geq 0$ for all $z \geq 0$ is said to be only skewed to the right; and if $\nu_F(z) \leq 0$ for all $z \geq 0$, it is said to be only skewed to the left.

The relationship $F <_c G$ (F c -precedes G) means that $G^{-1}[F(x)]$ is a convex function. For a continuous distribution F , the bivariate measurement of skewness $(S^-(F), S^+(F))$ verifies the following properties, where $aF + b$ and $-F$ mean the distributions of the corresponding transformation of a random variable that is F -distributed:

1. $(S^-(F), S^+(F)) = \mathbf{0}$, for any symmetric distribution F .
2. $(S^-(aF + b), S^+(aF + b)) = (S^-(F), S^+(F))$, for all $a > 0$, $-\infty < b < \infty$.
3. $(S^-(-F), S^+(-F)) = (-S^+(F), -S^-(F))$.
4. If $F <_c G$, then $(S^-(F), S^+(F)) \leq (S^-(G), S^+(G))$, understood as vector dominance.

These properties can be considered as a vectorial interpretation of the axioms given by Oja (1981) [7].

As it is easily proven that $\nu_F(0) = \gamma_M(F)$, we can establish that (2) and (3) give considerably clearer and more complete information than (1) about the skewness of any distribution function.

Most families of continuous distributions are only skewed to the right (or only to the left), while double-sign skewness is abundant within the discrete families, as shown in [15]. Nevertheless, the joint use of the function (2) and the bivariate index (3) makes it possible to improve the ordering of the skewness-based distribution discussed in [10], as can be seen in the following example.

Example 1. Assume the following random variable $X \in [-2, \infty)$ with PDF given by:

$$f(x) = \begin{cases} \frac{1}{4}x + \frac{1}{2}, & -2 \leq x < 0; \\ \frac{1}{2} \exp(-x), & x \geq 0. \end{cases}$$

Assume also the PDF, $g(y)$ of $Y = -X$. Then, $\gamma_M(F) = 0 = \gamma_M(G)$. That is, according to coefficient $\gamma_M(\cdot)$, both distributions have the same null skewness, although they do not even have a symmetric support set. However, using expression (2), we find that

$$v_F(z) = \begin{cases} \frac{1}{2} \exp(-z) - \frac{1}{8}(z-2)^2, & 0 \leq z < 2; \\ \frac{1}{2} \exp(-z), & z \geq 2, \end{cases}$$

and $v_G(z) = -v_F(z)$, for all $z \geq 0$. These functions are plotted in Figure 1, where it can be seen that $v_F(z) \geq v_G(z)$ for all $z \geq 0$, $S^+(F) = -S^-(G) \neq 0$ and $S^-(F) = 0 = S^+(G)$. Clearly, the information about skewness obtained from the aggregate skewness function $v(z)$ and the indices $S^+(\cdot)$ and $S^-(\cdot)$ is considerably more comprehensive than that obtained from $\gamma_M(\cdot)$.

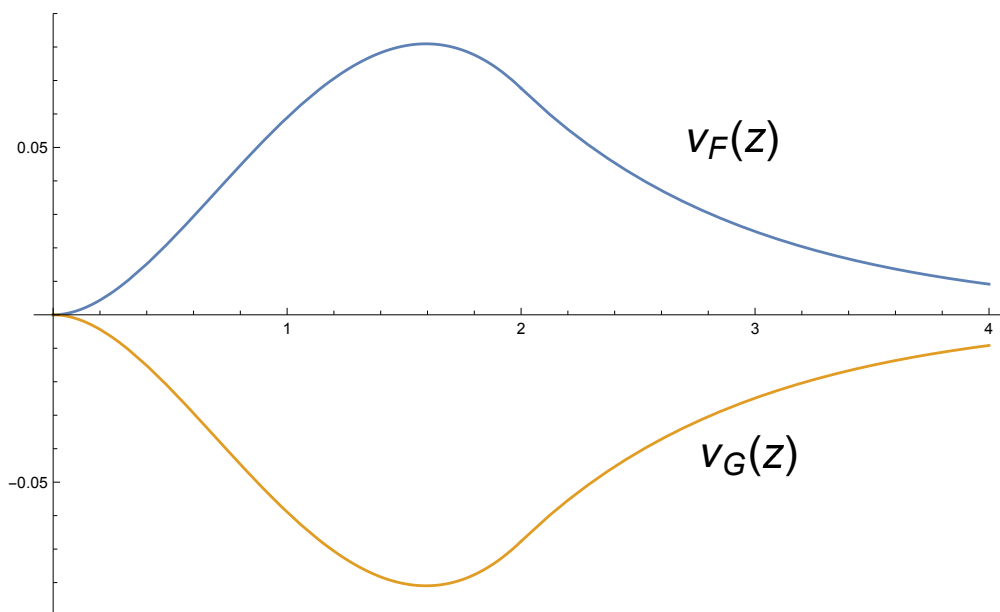


Figure 1. Skewness functions $v_F(z)$ and $v_G(z)$ in Example 1.

Outline

In applied statistical analysis, it is useful to have a large catalogue of plausible distributions with which to fit the data. According to García et al. (2015) [15], common measures of skewness can be complemented with a bivariate index of positive–negative skewness, and the authors show that the mode is the relevant central value to study both right and left skewness. In this paper, we extend the tool-box approach to fit data from probability distributions, introducing two orderings that are deduced from the skewness measures given in [15]. The first of those orderings is based on the positive part of the bivariate index of skewness, which in many instances coincides with the well known $\gamma_M(F)$. Nevertheless, the differences can be highly significant, as in the previous example. The second, more

noteworthy, order is based on the skewness function $v_F(z)$ and meets the first of the conditions, but not the reciprocal.

There are two reasons for ordering a family of distributions according to a given measurement of skewness. Firstly, as a property of the distribution, this ordering allows us to control its skewness by the appropriate selection of the parameter. When this is done (and the parameter is readily determined), the theoretical results have immediate applications in the data-fitting process. Secondly, when a given family of distributions is conceived as being more or less skewed according to the value of a parameter, and a measurement of skewness ratifies the ordering, it may be concluded that the functioning of this measurement provides a reasonably good fit with an intuitive conception of skewness.

The rest of this paper is organized as follows. In Section 2, we study the aggregate skewness function and the resultant skewness-based ordering of the gamma, log-logistic, lognormal, Weibull and asymmetric Laplace families of continuous probability distributions. In Section 3, we study the ordering of two of the most well-known distributions commonly used in PERT methods: the beta and the asymmetric triangular distributions. Finally, conclusions are presented in Section 4.

2. Families of Uniparametric Distributions Ordered by Skewness

Let F and G be unimodal distribution functions, with no centre or scale parameters, and modes M_F and M_G , respectively. We compare their respective skewness by two different criteria.

Definition 1. *If*

$$v_F(z) - v_G(z) \geq 0, \quad \forall z \geq 0, \quad (4)$$

then we say that F has equal or more aggregate skewness to the right at any point than G . We denote this by $F \geq_v G$.

Definition 2. *If F and G are both skewed only to the right, we say that F has equal or more maximum aggregate skewness to the right than G when*

$$S^+(F) - S^+(G) \geq 0, \quad (5)$$

and we denote this by $F \geq_+ G$.

With these definitions, it immediately follows that:

Proposition 1. *If $F \geq_v G$, then $F \geq_+ G$.*

The reverse implication is not true in general.

Proof. The proof follows immediately from the definitions given in (4) and (5). \square

In the next section, we consider some well known uniparametric families of continuous distributions, with no centre or scale parameters but depending on a skewness parameter, and examine whether they are ordered by aggregate skewness, or by maximum aggregate skewness. The gamma family is a very broad one, which includes many other well known distributions as particular cases. A study of the log-logistic, lognormal, Weibull and asymmetric Laplace families, one by one and in turn, when not included inside the previous one, will produce widely varying results.

2.1. Uniparametric Gamma Distributions

Let X be a uniparametric gamma distributed random variable, $G(\alpha)$. That is, its CDF $G(x; \alpha)$ is given by

$$G(x; \alpha) = \frac{1}{\Gamma(1 + \alpha)} \int_0^x t^\alpha e^{-t} dt, \quad (6)$$

for $x > 0$, where $-1 < \alpha < \infty$, and the mode is given by $M = \max \{\alpha, 0\}$. Then, for $-1 < \alpha \leq 0$, the density function decreases on x along the positive real line and we obtain that

$$v_G(z; \alpha) = 1 - G(z; \alpha) = \frac{1}{\Gamma(1 + \alpha)} \int_z^\infty t^\alpha e^{-t} dt.$$

In these cases, $v_G(z; \alpha)$ is a decreasing function on z , and $S^+(G(\alpha_1)) = v_G(0; \alpha_1) = 1$.

Proposition 2. Let $G(\alpha_1)$ and $G(\alpha_2)$ be gamma distributions with CDF as in (6). Then:

1. If $-1 < \alpha_1 < \alpha_2 < 0$, then $G(\alpha_2) \geq_v G(\alpha_1)$.
2. If $0 < \alpha_1 < \alpha_2$, then $G(\alpha_1) \geq_+ G(\alpha_2)$.

Proof. Part 1. We can write

$$v_G(z; \alpha_1) - v_G(z; \alpha_2) = G(z; \alpha_2) - G(z; \alpha_1).$$

By denoting $\alpha_2 = \alpha_1 + \varepsilon$, $\varepsilon > 0$, and then considering $u(z) = \frac{d}{dz} [v_G(z; \alpha_1) - v_G(z; \alpha_2)]$, we obtain

$$u(z) = \frac{z^{\alpha_2} e^{-z}}{\Gamma(1 + \alpha_2)} - \frac{z^{\alpha_1} e^{-z}}{\Gamma(1 + \alpha_1)} = z^{\alpha_1} e^{-z} \left[\frac{z^\varepsilon \Gamma(1 + \alpha_1) - \Gamma(1 + \alpha_2)}{\Gamma(1 + \alpha_2) \Gamma(1 + \alpha_1)} \right].$$

Therefore, $u(z) = 0$ when

$$z = z_0 = \left[\frac{\Gamma(1 + \alpha_2)}{\Gamma(1 + \alpha_1)} \right]^{1/\varepsilon},$$

$u(z)$ is negative for $0 < z < z_0$, and positive for $z > z_0$. Also, $v_G(0^+; \alpha_1) - v_G(0^+; \alpha_2) = 0$, $v_G(\infty; \alpha_1) - v_G(\infty; \alpha_2) = 0$. Then,

$$v_G(z_0; \alpha_1) - v_G(z_0; \alpha_2) = \frac{1}{\Gamma(1 + \alpha_1) \Gamma(1 + \alpha_2)} \int_0^{z_0} t^{\alpha_1} e^{-t} [\Gamma(1 + \alpha_1) t^\varepsilon - \Gamma(1 + \alpha_2)] dt,$$

is the integral of a negative function, so it is negative, and the proof is complete.

Part 2. For $0 < \alpha < \infty$, we have that

$$v_G(z; \alpha) = \begin{cases} \frac{1}{\Gamma(1+\alpha)} \left(\int_{z+\alpha}^\infty t^\alpha e^{-t} dt - \int_0^{\alpha-z} t^\alpha e^{-t} dt \right), & 0 \leq z < \alpha, \\ \frac{1}{\Gamma(1+\alpha)} \int_{z+\alpha}^\infty t^\alpha e^{-t} dt, & z \geq \alpha. \end{cases}$$

and,

$$\frac{dv_G}{dz} = \begin{cases} \frac{1}{\Gamma(1+\alpha)} \left[(\alpha - z)^\alpha e^{-(\alpha-z)} - (\alpha + z)^\alpha e^{-(\alpha+z)} \right], & 0 \leq z < \alpha, \\ \frac{-1}{\Gamma(1+\alpha)} z^\alpha e^{-z}, & z \geq \alpha. \end{cases}$$

Then, clearly we have that $dv_G/dz < 0$ for all $z \geq \alpha$. For $0 \leq z < \alpha$, if we denote

$$w(z) = (\alpha - z)^\alpha e^z - (\alpha + z)^\alpha e^{-z},$$

then the sign of dv_G/dz is the sign of $w(x)$. As $w(0) = 0$, $w(\alpha) = -(2\alpha)^\alpha e^{-\alpha} < 0$, and

$$\frac{dw}{dz} = z(z + \alpha)^{\alpha-1} e^{-z} - z(\alpha - z)^{\alpha-1} e^z \leq 0,$$

we conclude that ν_G is a decreasing function on $z \geq 0$ and $S^+(G(\alpha)) = \nu_G(0; \alpha)$.

$$S^+(G(\alpha)) = \frac{\Gamma(1 + \alpha, \alpha)}{\Gamma(1 + \alpha)},$$

where $\Gamma(1 + \alpha, \alpha)$ is the incomplete Gamma function, and then $S^+(G(\alpha))$ is a decreasing function on α , when $\alpha \rightarrow \infty$. Nevertheless, a simple plotting of the functionals $\nu_G(z; \alpha_i)$ for any $0 < \alpha_1 < \alpha_2$ shows that both functionals cross each other and that they are not ordered by " \geq_ν ". Thus, the proof is completed. \square

2.2. Log-Logistic Distributions

The CDF of a uniparametric log-logistic distributed random variable X is given by

$$F_{LL}(x; \theta) = (1 + x^{-\theta})^{-1}, \tag{7}$$

for $x > 0$, with $\theta > 0$. The mode of these distributions depends on θ . If $0 < \theta \leq 1$, then $M = 0$, and

$$\nu_{LL}(z; \theta) = \frac{1}{1 + z^\theta},$$

and $S^+(F_{LL}(\theta)) = 1$. The functionals $\nu_{LL}(z; \theta)$ for different values of θ inside the rank cross each other at $z = 1$, and these distributions are ordered neither by skewness function nor by skewness indexes. Nevertheless, for $\theta > 1$, the mode is

$$0 < M = \left(\frac{\theta - 1}{\theta + 1}\right)^{1/\theta} < 1. \tag{8}$$

Notice that M is an increasing function of θ when $\theta > 1$, because

$$\frac{dM}{d\theta} = M \cdot \left[\frac{2}{(\theta^2 - 1)\theta} - \frac{1}{\theta^2} \ln \frac{\theta - 1}{\theta + 1} \right] > 0. \tag{9}$$

When $\theta > 1$, it is also known from Arnold and Groeneveld (1995) that

$$\nu_{LL}(0; \theta) = \frac{1}{\theta}.$$

As $\nu_{LL}(z; \theta)$ is a decreasing function, it is then stated that $1 < \theta_1 < \theta_2$ implies $F_{LL}(\theta_1) \geq_+ F_{LL}(\theta_2)$. Furthermore, the skewness functions are ordered, as we prove below.

Proposition 3. Let be $F_{LL}(\theta_1)$ and $F_{LL}(\theta_2)$ log-logistic distributions with CDF as in (7), where $1 < \theta_1 < \theta_2$. Then,

$$F_{LL}(\theta_1) \geq_\nu F_{LL}(\theta_2). \tag{10}$$

Proof. Let $\theta > 1$. Then,

$$\nu_{LL}(z; \theta) = \begin{cases} \frac{1 - (M^2 - z^2)^\theta}{1 + (M + z)^\theta + (M - z)^\theta + (M^2 - z^2)^\theta}, & 0 \leq z \leq M, \\ \frac{1}{1 + (M + z)^\theta}, & z > M, \end{cases}$$

If we consider $1 < \theta_1 < \theta_2$, such that the respective modes verify $0 < M_1 < M_2 < 1$, we can then denote

$$a = (M_1 - z)^{\theta_1} < b = (M_1 + z)^{\theta_1},$$

$$c = (M_2 - z)^{\theta_2} < d = (M_2 + z)^{\theta_2}.$$

and consider the function h given by

$$h(\theta) = (M \pm z)^\theta, \quad 0 \leq z \leq M,$$

with M as in (8). Then,

$$\frac{dh}{d\theta} = \frac{(M \pm z)^{\theta-1}}{\theta(\theta+1)^2} \left(\frac{2\theta}{M^{\theta-1}} + \theta(1+\theta)^2(M \pm z) \ln(M \pm z) + (1+\theta)^2 M \ln \frac{\theta+1}{\theta-1} \right),$$

For $z < M$, this implies that $a < c, b < d$. With this notation, we can write $v_{LL}(z; \theta_1) - v_{LL}(z; \theta_2)$ as follows.

Firstly, for $0 \leq z \leq M_1$,

$$v_{LL}(z; \theta_1) - v_{LL}(z; \theta_2) = \frac{1-ab}{(1+a)(1+b)} - \frac{1-cd}{(1+c)(1+d)}$$

$$= \frac{(c-a) + (d-b) + ac(d-b) + bd(c-a) + 2(cd-ab)}{(a+1)(b+1)(c+1)(d+1)} > 0.$$

Secondly, for $M_1 < z \leq M_2$, we only need to compare $d - b$, because

$$v_{LL}(z; \theta_1) - v_{LL}(z; \theta_2) = \frac{1}{1+b} - \frac{1-cd}{(1+c)(1+d)}$$

$$= \frac{c + (d-b) + 2cd + bcd}{(1+b)(1+c)(1+d)} > 0.$$

Finally, when $z > M_2$,

$$v_{LL}(z; \theta_1) - v_{LL}(z; \theta_2) = \frac{1}{1+b} - \frac{1}{1+d} = \frac{d-b}{(1+b)(1+d)} > 0.$$

Hence, the proof is completed. \square

2.3. Lognormal Variance Distributions

$$LN(x; \sigma) = \Phi\left(\frac{\ln x}{\sigma}\right), \tag{11}$$

for $x, \sigma > 0$, where $\Phi(\cdot)$ is the standard normal distribution function. The mode is given by $M_\sigma = \exp(-\sigma^2)$ and

$$v_{LN}(z; \sigma) = 1 - \Phi\left(\frac{\ln[z + \exp(-\sigma^2)]}{\sigma}\right) - \Phi\left(\frac{\ln[\exp(-\sigma^2) - z]}{\sigma}\right).$$

Proposition 4. Let $LN(\sigma_1)$ and $LN(\sigma_2)$ be lognormal distributions with CDF as in (11), where $0 < \sigma_1 < \sigma_2$. Then,

$$LN(\sigma_2) \geq_v LN(\sigma_1).$$

Proof. For $0 < \sigma_1 < \sigma_2$, the corresponding modes are $M_1 > M_2$, and

$$\Phi\left(\frac{\ln(z + M_1)}{\sigma_1}\right) > \Phi\left(\frac{\ln(z + M_2)}{\sigma_2}\right),$$

$$\Phi\left(\frac{\ln(-z + M_1)}{\sigma_1}\right) > \Phi\left(\frac{\ln(-z + M_2)}{\sigma_2}\right),$$

because Φ is a strictly increasing function. Thus, we obtain that $\nu_{LN}(z; \sigma_1) > \nu_{LN}(z; \sigma_2)$ for all $z > 0$ and the proof is completed. \square

2.4. Uniparametric Weibull Distributions

Consider the uniparametric Weibull distributions family given by the CDF

$$W(x; c) = 1 - \exp(-x^c), \quad x > 0, c > 0. \tag{12}$$

The mode is known to be at 0, for $c \leq 1$ (as a limit, when $c < 1$) and at

$$0 < M_c = \left(\frac{c-1}{c}\right)^{1/c} < 1,$$

for $c > 1$. The expression for ν_W is given by

$$\nu_W(z; c) = \begin{cases} \exp[-(M_c + z)^c] + \exp[-(M_c - z)^c] - 1, & 0 < z < M_c \\ \exp[-(M_c + z)^c], & z \geq M_c \end{cases}$$

On the one hand, when $c < 1$, note that $\nu_{W(c)}(1) = e^{-1}$, so all these functions intersect at this point. Graphically, it can be seen that there is no ordering by “ \geq_ν ”, and also that $S^+(W(c)) = 1$, when $c < 1$. On the other hand, for $1 \leq c_1 < c_2$, the following result is obtained.

Proposition 5. Let $W(c_1)$ and $W(c_2)$ be Weibull distributions with $1 \leq c_1 < c_2$ and CDF as in (12). Then,

$$W(c_1) \geq_\nu W(c_2).$$

Proof. For $1 \leq c_1 < c_2$, the corresponding modes are $M_1 < M_2$. Then, for $0 < z < M_1$,

$$\nu_W(z; c_1) - \nu_W(z; c_2) = \{ \exp[-(M_1 + z)^{c_1}] - \exp[-(M_2 + z)^{c_2}] \}$$

$$+ \{ \exp[-(M_1 - z)^{c_1}] - \exp[-(M_2 - z)^{c_2}] \} > 0,$$

because each part of the expression inside brackets $\{ \cdot \}$ is positive. If we take $M_1 \leq z < M_2$, then

$$\nu_W(z; c_1) - \nu_W(z; c_2) = \{ \exp[-(M_1 + z)^{c_1}] - \exp[-(M_2 + z)^{c_2}] \}$$

$$+ \{ 1 - \exp[-(M_2 - z)^{c_2}] \} > 0,$$

for a similar reason. Finally, if we take $z > M_2$, then

$$\nu_W(z; c_1) - \nu_W(z; c_2) = \exp[-(M_1 + z)^{c_1}] - \exp[-(M_2 + z)^{c_2}] > 0,$$

and the proof is completed. \square

2.5. Asymmetric Laplace Distributions

The asymmetric Laplace distribution has been introduced in the literature by different ways ([16,17]). In this paper we will use Kozubowski and Podgórski (2002) [18] (later refined in [19]) to refer it. This distribution is obtained by using the scheme introduced by Fernández and Steel (1998) [20] to produce skewness on a symmetric distribution. In this way, the pdf of a skewed or asymmetric Laplace distribution can be written in the form

$$f(x; \mu, \sigma, \kappa) = \begin{cases} \frac{\sqrt{2}}{\sigma} \frac{\kappa}{1+\kappa^2} \exp\left[-\frac{\sqrt{2}}{\kappa\sigma}(\mu-x)\right], & x < \mu, \\ \frac{\sqrt{2}}{\sigma} \frac{\kappa}{1+\kappa^2} \exp\left[-\frac{\kappa\sqrt{2}}{\sigma}(x-\mu)\right], & x \geq \mu, \end{cases}$$

where $\sigma, \kappa > 0$, and $-\infty < \mu < \infty$. Then, we assign values $(0, 1)$ to the centre and scale parameters (μ and σ , respectively) in order to study the aggregate skewness function, and the extreme right and left skewness indices then depend only on the skewness parameter $\kappa > 0$. Thus, it is easily proven that:

1. The aggregate skewness function of an $AL(\kappa)$ distribution can be written as

$$v_{AL}(z; \kappa) = \frac{1}{1+\kappa^2} \left[\exp(-\sqrt{2}\kappa z) - \kappa^2 \exp\left(-\frac{\sqrt{2}}{\kappa}z\right) \right].$$

2. $v_{AL}(z; \kappa)$ is an increasing negative function of z when $\kappa > 1$, and it is a decreasing positive function of z when $0 < \kappa < 1$. $v_{AL,1}(z; 1) = 0$, for all $z \geq 0$. That is, any AL distribution is skewed only to the right or to the left, depending on κ . In any case, the function verifies $\lim_{z \rightarrow \infty} v_{AL}(z; \kappa) = 0$ but, when $\kappa \neq 1$, the function never reaches that limit value. To prove these results, it is sufficient to note that

$$\frac{dv_{AL}(z; \kappa)}{dz} = \frac{\sqrt{2}\kappa}{\kappa^2 + 1} \left[\exp\left(-\frac{\sqrt{2}}{\kappa}z\right) - \exp(-\sqrt{2}\kappa z) \right].$$

3. At $z = 0$, the skewness function takes the following value:

$$v_{AL}(0; \kappa) = \frac{1 - \kappa^2}{1 + \kappa^2}.$$

Then, $v_{AL}(0; \kappa)$ is the value for $S^+(F_{AL}(\kappa))$ or $S^-(F_{AL}(\kappa))$, depending on its sign.

4. $v_{AL}(z; \kappa)$ is a strictly decreasing function on κ . This is easily shown by means of

$$\frac{dv_{AL}(z; \kappa)}{d\kappa} = -\frac{\sqrt{2}z\kappa^2 + 2\kappa + \sqrt{2}z}{(\kappa^2 + 1)^2} \left[\exp\left(-\sqrt{2}\frac{z}{\kappa}\right) + \exp(-\sqrt{2}\kappa z) \right] < 0,$$

for all $z > 0$, and all $\kappa > 0$.

As a conclusion, we can enunciate the following Proposition, whose proof is straightforward and hence omitted.

Proposition 6. Assume $0 < \kappa_1 < \kappa_2 < \infty$, and let $F_{AL}(\kappa_1)$ and $F_{AL}(\kappa_2)$ be the respective asymmetric Laplace distributions. Then:

1. $F_{AL}(\kappa_1) \geq_v F_{AL}(\kappa_2)$.
2. If $0 < \kappa_1 < 1$, then $F_{AL}(\kappa_1)$ is skewed only to the right.
3. If $\kappa_1 > 1$, then $F_{AL}(\kappa_1)$ is skewed only to the left.

3. The Beta and the AST Distributions

The methods for Project Management and Review Technique (PERT) are well known and widely applied when the needed activities for a given project must be ordered according to precedence in

time. Some of these methods require modelling the time length of each activity as a random variable, following an expert's opinion. The beta and the asymmetric triangular distributions are commonly used by engineers to describe these time lengths. In any case, the indications of the experts can be related to a maximum and a minimum values and a mode, often completed with further considerations about the shape and skewness of the PDF of the time random variable. Then, a deep study of the skewness of both families of probability distributions would be welcome to improve the model fit.

On the one hand, the asymmetric standard triangular distribution (ASTD), free of center and scale parameters, depends on only one parameter $0 \leq \theta \leq 1$, and has the pdf:

$$f(x|\theta) = \begin{cases} 2x\theta^{-1}, & 0 \leq x \leq \theta, \\ 2(1-x)(1-\theta)^{-1}, & \theta \leq x \leq 1, \\ 0, & \text{elsewhere.} \end{cases}$$

There is a large body of literature that shows the use of the ASTD in PERT methods (see [21] and [19] and cites therein). Note that cases $\theta = 0, 1$ are members of the beta family of distributions.

For $0 < \theta < 1$, the $ASTD(\theta)$ CDF can be written as follows:

$$F(x|\theta) = \begin{cases} x^2\theta^{-1}, & 0 \leq x \leq \theta, \\ (2x - x^2 - \theta)(1-\theta)^{-1}, & \theta \leq x \leq 1. \end{cases}$$

As the mode is found to be at $x = \theta$, its skewness function is found to be

$$\nu_{ASTD}(z;\theta) = (1-2\theta) - \frac{(1-2\theta)}{\theta(1-\theta)}z^2,$$

for $0 \leq z \leq \min\{\theta, 1-\theta\}$. In the case $\theta = 0.5$, the skewness function is null. Then, for $0 < \theta < 0.5$ and $\theta < z \leq 1-\theta$,

$$\nu_{ASTD}(z;\theta) = \frac{(z-1+\theta)^2}{1-\theta}.$$

In the case $0.5 < \theta < 1$, for $1-\theta < z \leq \theta$,

$$\nu_{ASTD}(z;\theta) = -\frac{(\theta-z)^2}{\theta},$$

and it is easily found that

$$\nu_{ASTD}(z;\theta) = -\nu_{ASTD}(z;1-\theta), \quad (13)$$

for $0 \leq z < \infty$.

Some algebra allows to prove that, being $0 < \theta_1 < \theta_2 < 1$,

1. $ASTD(\theta_1) \geq_v ASTD(\theta_2)$.
2. If $0 < \theta_i < 0.5$, then $S_{ASTD}^+(\theta_i) = \nu_{ASTD}(0;\theta_i) = 1 - 2\theta_i$, and $S_{ASTD}^-(\theta_i) = 0$.
3. If $0.5 < \theta_i < 1$, then $S_{ASTD}^+(\theta_i) = 0$ and $S_{ASTD}^-(\theta_i) = \nu_{ASTD}(0;\theta_i) = 1 - 2\theta_i$.

Therefore, the skewness of the ASTD distributions is completely controlled by the parameter θ .

On the other hand, the pdf of a beta distribution is given by

$$f_B(x;\alpha,\beta) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)}, \quad 0 \leq x \leq 1,$$

where $\alpha, \beta > 0$, and $B(\cdot, \cdot)$ is the beta function. Given that its CDF $F(x;\alpha,\beta)$ verifies that $F(x;\alpha,\beta) = 1 - F(x;\beta,\alpha)$ and the sign of its skewness depends only on the condition $\beta \geq \alpha$ or $\beta \leq \alpha$, we can study only the case $\beta > \alpha$.

We are interested on the cases $\alpha, \beta > 1$, where there is a unique mode M ,

$$M = \frac{\alpha - 1}{\alpha + \beta - 2} \doteq b(\alpha, \beta).$$

Hence, we only consider cases where $1 < \alpha < \beta$, where there exists a right skewness; the cases $1 < \beta < \alpha$, with left skewness, can be immediately deduced by taking the parameters in reverse.

Notice that $\Pr(X > M + z) > 0$ requires $0 \leq z \leq b(\beta, \alpha)$, and that $\Pr(X < M - z) > 0$ requires $0 \leq z \leq b(\alpha, \beta)$. Then,

$$v_B(z; \alpha, \beta) = \begin{cases} 1 - I_{M+z}(\alpha, \beta) - I_{M-z}(\alpha, \beta) > 0, & 0 \leq z \leq b(\alpha, \beta) \\ 1 - I_{M+z}(\alpha, \beta) > 0, & b(\alpha, \beta) < z \leq b(\beta, \alpha) \\ 0, & z > b(\beta, \alpha), \end{cases} \quad (14)$$

where,

$$I_z(\alpha, \beta) = \int_0^z \frac{t^{\alpha-1} (1-t)^{\beta-1} dt}{B(\alpha, \beta)}$$

is the well known Beta Regularized function.

Firstly, observe that

$$v_B(0; \alpha, \beta) = 1 - \frac{2}{B(\alpha, \beta)} \int_0^M x^{\alpha-1} (1-x)^{\beta-1} dx,$$

and

$$\frac{1}{B(\alpha, \beta)} \int_0^M x^{\alpha-1} (1-x)^{\beta-1} dx < \frac{1}{B(\alpha, \beta)} \int_0^m x^{\alpha-1} (1-x)^{\beta-1} dx \approx \frac{1}{2},$$

where

$$m = \frac{\alpha - \frac{1}{3}}{\alpha + \beta - \frac{2}{3}}$$

is the approximate median of the distribution.

Secondly, if $0 \leq z \leq b(\alpha, \beta) < b(\beta, \alpha)$, then

$$B(\alpha, \beta) \cdot \frac{dv_B(z; \alpha, \beta)}{dz} = -[b(\alpha, \beta) + z]^{\alpha-1} [b(\beta, \alpha) - z]^{\beta-1} - [b(\alpha, \beta) - z]^{\alpha-1} [b(\beta, \alpha) + z]^{\beta-1},$$

which is negative within the rank of z . For $b(\alpha, \beta) < z \leq b(\beta, \alpha)$,

$$B(\alpha, \beta) \cdot \frac{dv_B(z; \alpha, \beta)}{dz} = -[b(\alpha, \beta) + z]^{\alpha-1} [b(\beta, \alpha) - z]^{\beta-1} < 0.$$

Hence, for $0 \leq z \leq b(\beta, \alpha)$, $v_B(z; \alpha, \beta)$ is a strictly decreasing continuous function with $v_B(0; \alpha, \beta) > 0$ and $v_B(b(\beta, \alpha); \alpha, \beta) = 0$.

Now we focus on the family of Beta distributions with given mode, M . That is, we consider the subfamily of Beta distributions:

$$B\left(\alpha + 1, 1 + \frac{1 - M}{M} \alpha\right),$$

with $\alpha > 0$. Then, with the aid of a proper software (we have used Wolfram Mathematica 10), one can obtain the derivative

$$\frac{\partial}{\partial \alpha} v_B\left(z; \alpha + 1, 1 + \frac{1 - M}{M} \alpha\right),$$

and maximize this function, in two cases:

First case, the constrains are $\alpha \geq 1$, $0 < m < 1/2$, $0 \leq z \leq b(\alpha + 1, 1 + (1 - M)\alpha/M)$. The maximum value of the function is 0, and it is achieved when $M = 0.5$, $\alpha \simeq 3.54147$, $z \simeq 0.309936$.

Second case, the constrain are $\alpha \geq 1$, $0 < m < 1/2$, $b(\alpha + 1, 1 + (1 - M)\alpha/M) < z \leq b(1 + (1 - M)\alpha/M, \alpha + 1)$. The maximum value of the function is -5.07056×10^{-6} , and it is achieved when $M = 0.123564$, $\alpha \simeq 1.62726$, $z \simeq 0.632457$.

With these results, we can conclude that $v_B(z; \alpha + 1, 1 + (1 - M)\alpha/M)$ decreases with the feasible values of α . That way, the subsets of Beta distributions with fixed mode are ordered on skewness (see Figure 2). As the parameter values increase, these Beta distributions become less skewed.

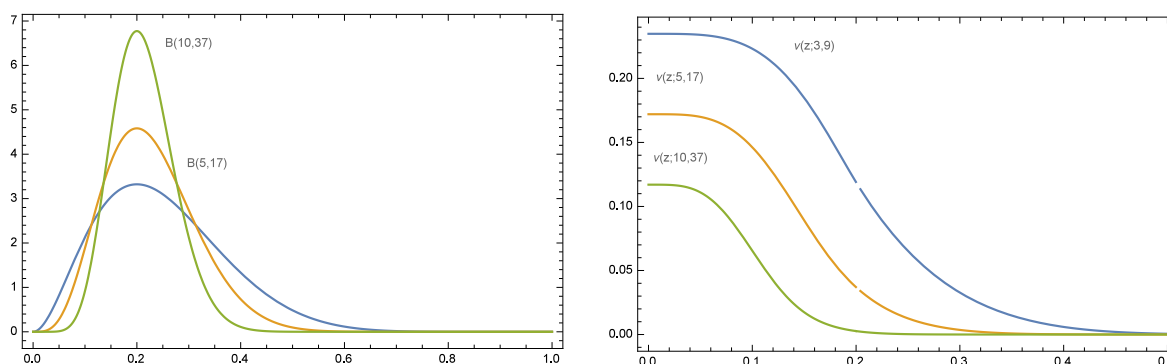


Figure 2. Beta distributions with common given mode $M = 0.2$ (left panel) for $\alpha = 2, 4$ and 9 and their skewness functions v_B (right panel).

4. Conclusions

In this paper two main objectives are achieved: on the one hand, the given examples show that the skewness function orders the mesh in good accordance with the intuitive conception of skewness. Moreover, these examples show that the skewness of a distribution obtained from certain parametric families can be controlled by reference to their parameters.

As we show, the function $v_F(z)$ facilitates the description of a random variable by means of a probability distribution, by making any skewness in the model easily observable and should be undertaken to examine the use of these properties in data fitting.

In practice, much can be learned from this model, but there remains the risk that it may be wrongly specified in real applications. Thus, in practice we must be willing to assume that the underlying distribution has a unique mode and belongs to a uniparametric family of distributions.

In many practical situations, the maximum skewness index coincides with the well known $\gamma_M(F)$, but this second index only takes into account the difference of probability weights at each side of the mode, while the first takes a value from the point where this difference is maximum. Moreover, the aggregate skewness function gives more accurate information about how the probability weight is distributed along both sides of the mode. Accordingly, the condition $F \geq_v G$ provides highly valuable information.

Author Contributions: All authors have contributed equally to this paper.

Funding: This research received no external funding.

Acknowledgments: This research was partially funded by MINECO (Spain) grant number EC02017-85577-P. The authors are grateful for helpful suggestions made by two reviewers.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Shaked, M.; Shanthikumar, J.G. Stochastic Orders. In *Springer Series in Statistics 43*; Springer: New York, NY, USA, 2007.
2. Lehmann, E.L. Ordered families of distributions. *Ann. Math. Statist.* **1955**, *26*, 399–419. [[CrossRef](#)]

3. Arnold, B.C. Majorization and the Lorenz Order: A brief introduction. In *Lecture Notes in Statistics 43*; Springer: New York, NY, USA, 1987.
4. Nanda, A.K.; Shaked, M. The hazard rate and reverse hazard rate orders, with applications to order statistics. *Ann. Inst. Statist. Math.* **2001**, *53*, 853–864. [[CrossRef](#)]
5. Ramos–Romero, H.M.; Sordo–Díaz, M.A. The proportional likelihood ratio order and applications. *Questio* **2001**, *25*, 211–223.
6. Gupta, A.K.; Aziz, M.A.S. Convex Ordering of Random Variables and its Applications in Econometrics and Actuarial Science. *Eur. J. Pure Appl. Math.* **2010**, *3*, 79–85.
7. Oja, H. On location, scale, skewness and kurtosis of univariate distributions. *Scand. J. Stat.* **1981**, *8*, 154–168.
8. Van Zwet, W.R. Mean, Median, Mode II. *Stat. Neerl.* **1979**, *33*, 1–5. [[CrossRef](#)]
9. MacGillivray, H.L. Skewness and asymmetry: measures and orderings. *Ann. Stat.* **1986**, *14*, 994–1011. [[CrossRef](#)]
10. Arnold, B.C.; Groeneveld, R.A. Measuring Skewness with respect to the Mode. *Am. Stat.* **1995**, *49*, 34–38.
11. Sato, M. Some remarks on the mean, median, mode and skewness. *Aust. J. Stat.* **1997**, *39*, 219–224. [[CrossRef](#)]
12. Von Hippel, P.T. Mean, Median and Skew: Correcting a Textbook Rule. *J. Stat. Educ.* **2005**, *13*. [[CrossRef](#)]
13. Das, S.; Mandal, P.K.; Ghosh, D. On Homogeneous Skewness of Unimodal Distributions. *Indian J. Stat.* **2009**, *71-B*, 187–205.
14. Rubio, F.J.; Steel, M. On the Marshall–Olkin transformation as a skewing mechanism. *Comput. Stat. Data Anal.* **2012**, *56*, 2251–2257. [[CrossRef](#)]
15. García, V.J.; Martel–Escobar, M.; Vázquez–Polo, F.J. Complementary information for skewness measures. *Stat. Neerl.* **2015**, *69*, 442–459. [[CrossRef](#)]
16. Mc Gill, W.J. Random fluctuations of response rate. *Psychometrika* **1962**, *27*, 3–17. [[CrossRef](#)]
17. Holla, M.S.; Bhattacharya, S.K. On a compound Gaussian distribution. *Ann. Instit. Stat. Math.* **1968**, *20*, 331–336. [[CrossRef](#)]
18. Kozubowski, T.J.; Podgórski, K. Maximum likelihood estimation of asymmetric Laplace parameters. *Ann. Inst. Stat. Math.* **2002** *54*, 816–826.
19. Kotz, S.; van Dorp, J.R. Uneven two-sided power distribution: with applications in econometric models. *Stat. Methods Appl.* **2004**, *13*, 285–313. [[CrossRef](#)]
20. Fernández, C.; Steel, M.F.J. On Bayesian Modeling of Fat Tails and Skewness. *J. Am. Stat. Assoc.* **1998**, *93*, 359–371.
21. Johnson, D. The triangular distribution as a proxy for the beta distribution in risk analysis. *The Statistician* **1997**, *46*, 387–398. [[CrossRef](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).