Working Paper 216

On a New Measure of Skewness for Unimodal Distributions

By

Shubhabrata Das
&
Diptesh Ghosh

October 2003

Please address all your correspondence to:

Prof. Shubhabrata Das
Indian Institute of Management Bangalore
Bannerghatta Road
Bangalore – 560076, India
Email: shubho@iimb.ernet.in
Phone : 080 – 6993150
Fax : 080 - 6584050

Prof. Diptesh Ghosh
P & QM Area, IIM Ahmedabad
Vastrapur, Ahmedabad 380 015, India
Email : diptesh@iimahd.ernet.in
Phone : 079 – 6324917
Fax : 079 – 6306896

Abstract

In this paper we introduce a concept of skewness and suggest its measure among a class of unimodal distributions. The measure is built on the lack of symmetry of the density function around the the mode of the distribution. It is shown to satisfy all standard properties expected from a measure of skewness, including location and scale invariance. The extreme values of the measure are characterized. Although the measure is defined for a relatively narrow class of distributions, its utility is established by showing that it is applicable for most popularly used continuous distribution families. The introduced measure is compared with the other established measures like Pearson's skewness and standardized third moment and it is shown to be more strict. Two alternative ways of partial ordering among the distributions based on this skewness are also described. The utility of the proposed measure is examined in other cases including discrete distributions.

Keywords: Beta distribution, Bowley's measure of skewness, Gamma distribution, location and scale invariance, Pearson's mode skewness, Poisson distribution, skewness function, standardized third moment, unimodal distribution.

On a New Measure of Skewness for Unimodal Distributions

Shubhabrata Das   Diptesh Ghosh

1 Background and Introduction

For any symmetric distribution, the three most popular measures of central tendency, viz., the Mean μ, the Median m, and the Mode M coincide. The lack of symmetry for other distributions is expressed through measures of skewness. Although skewness has been studied extensively in the past hundred years, the job of deciding on a single measure is far from obvious, not only because of typical tradeoff between efficiency and robustness, but also since the concept of skewness is not universally understood or accepted. Skewness, by convention, may be defined as the lack of symmetry of the density function around of its measure of central tendency. However they may be classified according to the difference choices of central tendency as most measures are defined in terms of the relative values of the mean, median, mode and various percentiles and moments of the random variable. Possibly the earliest measure of skewness can be attributed to Pearson [14] who advocated $\frac{\mu-M}{\sigma}$, a measure which is typically referred to as the *Pearson's measure of skewness*. Pearson also proposed a second measure in terms of the relative values of the mean and median which is related to the earlier measure by the empirical mean-median-mode relationship. Although Galton proposed a measure in terms of symmetry of 80th percentile and 20th percentile around the median in 1896, a more popular measure of this type was proposed by Bowley [3] in the early 1900's. This measure $\frac{(F^{-1}(0.75)-m)-(m-F^{-1}(0.25))}{F^{-1}(0.75)-F^{-1}(0.25)}$ became known as Bowley's measure of skewness. Around the same time a moment based measure $\frac{\mu_3}{\sigma^3}$ is reported in Yule [16]; this is usually referred to as the *standardized third moment* and is the most commonly used skewness measure till date. In later years, several modifications of these classical measures were studied. For example, a measure which modifies the Pearson's measure through replacing the denominator by $E|X-M|$ was suggested in MacGillivray [11], and modifications of Bowley's measure in which the quartiles were replaced by other percentiles were studied (David [6]). Benjamini and Krieger [2], Doksum [7], among others, unified some of these approaches in a common framework. Recently, research has also focused on measures of skewness that are robust (see, e.g., Aucremanne et al. [1], and Brys et al. [4, 5]).

Irrespective of the choice of measures, it is generally agreed (viz. Benjamini [2]), that any measure of skewness should

- lie between -1 and 1 for any distribution;
- be 0 for any symmetric distribution; and
- be location and scale free.

Curiously, the literature on skewness measures for unimodal distributions does not contain a measure based on the asymmetry of the density function about the mode, although measures based on the asymmetry of the density function about the median have been reported. In this paper, we propose a measure based on the asymmetry of the density function about the mode. In the next section, we formally define this measure and present its properties. We also describe two alternative ways of partial ordering among the distributions based on this skewness. In Section 3 we show that several commonly occurring continuous unimodal asymmetric distributions are homogeneously skewed as per our skewness measure. In Section 4 we compare our skewness measure with the more established skewness measures like Pearson's measure of skewness and standardized third moment. Secion 5 deals with the extension of our skewness measure for distributions with flat modal regions (Section 5.1), for discrete distributions (Section 5.2), especially the Poisson distribution, and for continuous distributions with a unique antimode (Section 5.3). The last section summarizes the contributions of the paper, and proposes directions for future research.

2 Homogeneous Skewness

2.1 Proposed Measure

Our treatment of homogeneous skewness is restricted to unimodal distributions, i.e. distributions with densities having unique local maxima. Without any loss of generality, we will include all distributions with non-increasing/non-decreasing densities in this class. We primarily deal with unimodal continuous distributions, although extended coverage is discussed in Section 5.

For the purpose of the present work, a unimodal distribution may be defined as follows: A distribution with density function $f(x)$ is called *unimodal* if there exists a unique M such that $f(x)$ is non-decreasing on $(-\infty, M)$, and non-increasing on (M, ∞). The value M is called the *mode* of the distribution. The mode can be either the left or right end point of the support of the distribution, as would be the case with distributions with increasing or decreasing density functions. The support of the distributions may or may not be finite.

Having defined unimodal distributions, we now introduce the class of *homogeneously skewed* distributions and a partial ordering of skewness within that class.

Definition 1 Homogeneously right skewed distribution: *A Unimodal distribution (equivalently, the random variable having a unimodal distribution) with Mode* M *and density function* $f(\cdot)$ *is said to be* homogeneously right-skewed, *or equivalently* homogeneously skewed to the right *if*

$$f(x+M) \geq f(x-M), \quad \forall x > 0. \tag{1}$$

It is said to be homogeneously right skewed in the strict sense *provided the strict inequality holds in (1) with positive measure, in addition. As usual, we may replace* $\forall x$ *by* a.e. x *(almost everywhere in x) in (1).*

Homogeneously left-skewed distributions are defined analogously, i.e. with the inequality sign in (1) reversed. A unimodal distribution is said to be *homogeneously skewed* provided it is either homogeneously right-skewed or homogeneously left-skewed.

Let us denote the class of homogeneously skewed distributions by $\mathcal{F}$. It is convenient to formally define the skewness function and a measure of skewness for distributions in $\mathcal{F}$.

Definition 2 The skewness function: *The skewness function of a distribution in* $\mathcal{F}$ *(with Mode* M *and p.d.f.* $f(\cdot)$*) is defined as*

$$\gamma_f(x) = f(M+x) - f(M-x), \quad x > 0. \tag{2}$$

Definition 3 The measure of homogeneous skewness: *The measure of homogeneous skewness of a distribution in* $\mathcal{F}$ *(with Mode* M *and p.d.f.* $f(\cdot)$*) is defined* *as*

$$\tau_f = \int_0^\infty \gamma_f(x)dx = \int_0^\infty \{f(M+x) - f(M-x)\}dx. \quad (3)$$

2.2 Properties of the Measure

It is easy to see that for any distribution belonging to $\mathcal{F}$, its measure of homogeneous skewness is positive (negative) if and only if the distribution is homogeneously skewed to the right (left).

The following theorem establishes the bound for the skewness measure and characterizes the extreme values and central value of the measure.

Theorem 1 *For any distribution $F \in \mathcal{F}$ with density f, its measure of homogeneous skewness τ_f*

1. *is equal to 0 if and only the distribution is symmetric;*
2. *lies between -1 and 1;*
3. *is equal to -1 if and only if the density is nondecreasing on its support; and*
4. *is equal to 1 if and only if the density is non-increasing on its support.*

Proof: The first statement follows from the fact that $\gamma_f(x)$ is of the same sign $\forall x$, and consequently $\tau_f = 0$ if and only if $\gamma_f(x) = 0, \ \forall x$, which is the case only for symmetric distributions. The key observation in proving the other parts is:

$$-f(M-x) \leq \gamma_f(x) \leq f(M+x), \quad \forall x, \quad (4)$$

integrating which leads to Statement 2. Statement 3 also follows by noting that the first inequality in (4) would hold for all x, if and only if $f(M+x) = 0, \ \forall x$, i.e. $F(M) = 1$, equivalently for $f(\cdot)$ nondecreasing on its entire support. The last statement follows from the parallel argument. □

It is worthwhile to mention that there exist asymmetric distributions outside the class $\mathcal{F}$, which are positively skewed according to other popular definitions of skewness and yet for which the integral in (3) is negative.

Theorem 2 *Suppose* X *is a random variable with distribution in* $\mathcal{F}$. *Then for any constants* a *and* b, *the measure of homogeneous skewness for* $Y = aX + b$ *and* X *are the same. In other words, the measure of homogeneous skewness* τ *is free of location and scale.*

Proof: The modes and densities of Y and X must be related by

$$M_Y = aM_X + b; \quad \text{and} \quad f_Y(y) = \frac{1}{a} f_X(\frac{y-b}{a}). \tag{5}$$

Hence the measure of homogeneous skewness for Y is

$$\begin{aligned}
\tau_Y &= \int_0^\infty \{f_Y(M_Y + y) - f_Y(M_y - y)\} dy \\
&= \int_0^\infty \frac{1}{a}\{f_X(M_X + \frac{y}{a}) - f_X(M_X - \frac{y}{a})\} dy \quad \text{by (5)} \\
&= \int_0^\infty \{f_x(M_X + x) - f_X(M_X - x)\} dx \ \ (\text{substituting } \frac{y}{a} = x\ ) \ = \tau_X
\end{aligned}$$

□

Theorems 1 and 2 ensure that the standard properties expected of a measure of skewness are satisfied by our proposed measure. In addition, the first theorem provides a nice characterization of the extreme values of the measure.

2.3 Ordering of Distributions in terms of Homogeneous Skewness

All the standard measures assume or impose an ordering of distributions with regard to their skewness. This has been elaborately explored in [11], [13], and [15] among others. With reference to our notion of homogeneous skewness, we have two alternative ways of ordering.

We advocate ordering the distributions in $\mathcal{F}$ according to their homogeneous skewness function.

Definition 4 *A distribution with density function* $f_1(\cdot)$) *in* $\mathcal{F}$ *is said to be at least as skewed as another distribution with density function* $f_2(\cdot)$ *in* $\mathcal{F}$ *if*

$$|\gamma_{f_1}(x)| \geq |\gamma_{f_2}(x)| \quad \forall x. \tag{6}$$

The above definition implies that if one distribution in $\mathcal{F}$ is more homogeneously skewed than another distribution in $\mathcal{F}$, then the γ of the former is also higher. However, this defines only a partial ordering of distributions in the sense that for two distributions in $\mathcal{F}$ neither may dominate the other, although clearly one may have higher measure of homogeneous skewness).

Another disadvantage of comparing through the skewness function is that its domain may change from a distribution to another. Alternatively one may define a total ordering of distributions in $\mathcal{F}$ through the measure of homogeneous skewness, i.e.

$$F_1 \leq_2 F_2 \Leftrightarrow \tau_{f_1} \leq \tau_{f_2}.$$

3 Homogeneous Skewness of Different Distributions

In Section 2.1, we have introduced a new measure of skewness for distributions in $\mathcal{F}$. Here we show that it is not overtly restrictive since many commonly used distributions belong to $\mathcal{F}$. We also make a comparison with other standard measures of skewness for these distributions. Note that for proving homogeneous skewness for distributions supported on $[a, b]$, it is sufficient to show that $\gamma_f(x)$ does not change sign for $0 \leq x \leq \min(M - a, b - M)$. In view of Theorem 2, it is sufficient to work with the standardized members from these distribution families rather than the general form.

3.1 Triangular distribution

In view of the discussion above, we consider a standardized form of the triangular distribution, i.e. one supported on $[0, 1]$. The density function of this distribution is

$$f(x) = \begin{cases} \frac{2x}{M} & \text{for } 0 \leq x \leq M \\ \frac{2(1-x)}{(1-M)} & \text{for } M \leq x \leq 1 \\ 0 & \text{otherwise.} \end{cases}$$

It is easy to see that, for $y > 0$,

$$\gamma(y) = \{f(M) - f(M - y)\} - \{f(M) - f(M + y)\} = 2y\left\{\frac{1}{M} - \frac{1}{1-M}\right\};$$

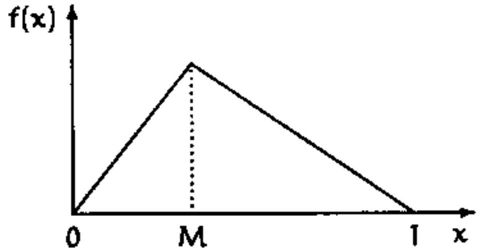

Figure 1: Density function for a triangular distribution

thus, the distribution is homogeneously skewed (i.e. $\in \mathcal{F}$) and is homogeneously right-skewed if and only if $M \leq 0.5$. It can be also shown that under identical condition, all other standard measures of skewness are nonnegative.

3.2 Gamma Distribution

The standard gamma distribution has a density function of the form:

$$f(x) = \begin{cases} 0 & \text{for } x \leq 0 \\ \frac{x^{\alpha-1}\exp(-x)}{\Gamma(\alpha)} & \text{for } x > 0. \end{cases}$$

The mode of this distribution is at $M = (\alpha - 1)$, provided $\alpha > 1$. We would ignore the case of $\alpha \leq 1$, since this leads to a decreasing density function in the entire range and consequently is trivially covered. The gamma distribution is a positively skewed distribution according to all traditional measures of skewness. In the following, we show that it is homogeneously skewed to the right.

Let us define:

$$h(y) = \frac{f(M+y)}{f(M-y)}, \tag{7}$$

for $0 < y < M$. Then its derivative is given by

$$h'(y) = 2y^2 (M+y)^{(M-1)} (M-y)^{-(M+1)} \exp(-2y) \tag{8}$$

This implies $h(y) > h(0) = 1$, $\forall 0 < y < M$, since for $y > M$, $f(M-y) = 0$, i.e. the distribution is homogeneously skewed to the right.

3.3 Beta Distribution

The standard beta distribution is defined on a support of $[0, 1]$. It has two parameters α and β, and its density function is given by:

$$f(x) = \begin{cases} \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)} & \text{for } 0 \leq x \leq 1 \\ 0 & \text{otherwise.} \end{cases}$$

The mode of this distribution is at $M = \frac{\alpha-1}{\alpha+\beta-2}$, provided $\alpha > 1, \beta > 1$; this is the scenario assumed in the remainder. For $\alpha \leq 1, \beta > 1$ and $\alpha > 1, \beta \leq 1$, the density functions are respectively decreasing and increasing and hence homogeneously skewed.

According to the existing measures, the distribution is positively (negatively) skewed for $\beta > (<) \alpha$. In the following, we show that the distribution is homogeneously skewed to the right(left) under the same condition.

Now for convenience, let $p = \alpha - 1$ and $q = \beta - 1$ and let $h(\cdot)$ be as in (7). The derivative of $h(\cdot)$ in this case can be reduced to

$$h'(y) = 2y^2(\beta - \alpha)(M - y)^{-p-1}(M + y)^{p-1}(1 - M - y)^{q-1}(1 - M + y)^{-q-1}, \quad (9)$$

which is positive or negative for $0 < y < \min(M, 1 - M)$ depending on whether β is larger (smaller) than α. This proves homogeneous skewness for the beta distribution by noting that M is less (greater) than $1 - M$ under this condition.

3.4 Lognormal distribution

The standard Lognormal distribution is defined using two parameters μ and σ. Its density function is given by:

$$f(x) = \begin{cases} \frac{\exp\left(-\frac{(\ln(x)-\mu)^2}{2\sigma^2}\right)}{\sqrt{2\pi}\sigma x} & \text{for } x \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

The mode of this distribution is at $M = \exp(\mu - \sigma^2)$. Skewness computations using this distribution are tedious, but the distribution can be shown to be homogeneously right-skewed. A typical plot of the skewness function for this distribution is shown in Figure 2.

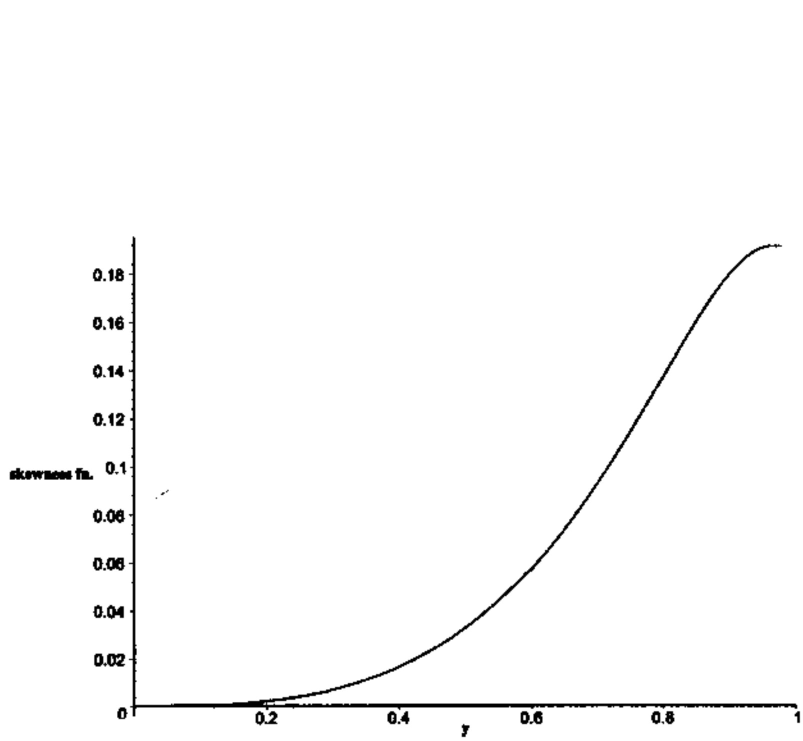

Figure 2: The skewness function $\gamma_f(y)$ for the lognormal distribution

3.5 Weibull Distribution

The standard Weibull distribution is defined using a single parameter c. Its density function is given by:

$$f(x) = \begin{cases} cx^{c-1}\exp(-x^c) & \text{for } x \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

The mode of this distribution is at $M = \left(\frac{c-1}{c}\right)^{\frac{1}{c}}$. This distribution is not homogeneously skewed as per our definition. Figure 3 shows the skewness function for the Weibull distribution when c varies between 3 and 4. Empirically, we have observed that the distribution is homogeneously skewed to the right for $c \leq 3$.

4 Comparison with Other Measures of Skewness

In the previous section we have shown that many commonly occurring asymmetric distributions are homogeneously skewed. We also observed that the the proposed

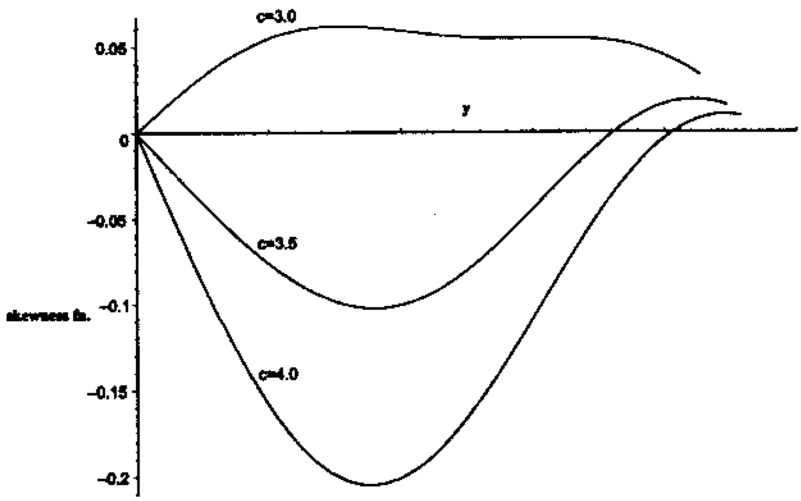

Figure 3: The skewness function $\gamma_f(y)$ for the Weibull distribution

measure is in harmony with other measures (in terms of sign, although not necessarily in terms of its value) for these distribution families. Therefore, it is of natural interest to examine the relationship between this notion of skewness and other common definitions of skewness in the general framework.

4.1 Comparison with Pearson's Measure of Skewness

To start with, we compare the proposed measure with Pearson's measure of skewness; this is the only natural comparison since among all the popular measures, Pearson's measure is the only one which is based on mode of the distribution. Theorem 3 and the subsequent discussion will establish that the proposed measure is more stringent.

Theorem 3 *If a random variable or its distribution is homogeneously skewed to the right (left), then its Pearson's measure of skewness is necessarily positive (negative).*

Proof:

$$\mu - M = \int_{-\infty}^{\infty} (x - M) f(x) dx \tag{10}$$

$$
\begin{aligned}
&= \int_{-\infty}^{M} (x - M) f(x) dx + \int_{M}^{\infty} (x - M) f(x) dx \\
&= -\int_{0}^{\infty} y f(M - y) dy + \int_{0}^{\infty} y f(M + y) dx \\
&= \int_{0}^{\infty} y \gamma_f(y) dy. \qquad (11)
\end{aligned}
$$

Now for distributions homogeneously skewed to the right (left), $\gamma_f(x) \geq (\leq)\ 0,\ \forall x$, and consequently $\mu \geq (\leq) M$, implying that the distribution is positively (negatively) skewed as per Pearson's definition. □

It is easy to visualize asymmetric distributions (having positive/negative skewness as per Pearson's measure) which are not homogeneously skewed. Consider, for example, a distribution with the density function $f(x)$ having the quadrangular shape as given in Figure 4. The mode of this distribution is at $x = M$; $M - p = q - M$ and $f(p) > f(q)$. This distribution is positively skewed as per Pearson's notion of

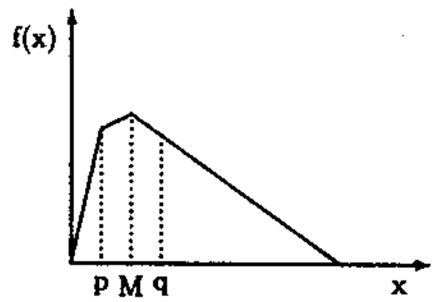

Figure 4: A positively skewed distribution according to Pearson's skewness measure

skewness. However, $\gamma_f(x)$ is negative in the range $(0, q)$, but positive for $x > q$, and consequently the distribution is not homogeneously skewed.

The example above in conjunction with Theorem 3, shows that the concept of homogeneous skewness is stronger than that of the Pearson's skewness.

4.2 Comparison with Other Measures of Skewness

It is not possible to draw a direct comparison with other measures of skewness, which are typically based on asymmetry around mean or median. In order to draw a general comparison we need to consider suitable modifications of the standard measures. We therefore modify the standardized third central moment by considering the the third moment around the mode.

Theorem 4 *For distributions homogeneously skewed to the right (left), the skewness measure based on third moment around mode is also positive (negative).*

Proof: Following steps similar to (10) — (11), one can show that

$$E(X - M)^3] = \int_0^\infty y^3 \gamma_f(y) dy;$$

this proves the theorem on similar lines to Theorem 3. □

In other words, the notion of homogeneous skewness is more stringent in this comparison also. The counter-example, as outlined in the previous subsection may also work here to show that there exists distributions having the third moment based skewness positive or negative which are not homogeneously skewed.

Although our measure appears to be more stringent than Bowley's measure in spirit, a direct comparison is not possible, even if one were to consider modified versions of Bowley's measure by replacing the median by mode in these measures. It may be tempting to redefine our measure in terms of symmetry around median, i.e. take

$$\tau_f^m = \int_0^\infty \gamma_f^m(x) dx = \int_0^\infty \{f(m + x) - f(m - x)\} dx.$$

to facilitate such comparison. However, it is easy to see that $\gamma_f^m(\cdot)$ so defined, must change sign at least once for any distribution, making such a modification not worth considering. In passing it is worthwhile to recall that the condition that $\gamma_f^m(\cdot)$ changes sign exactly once is known to be sufficient condition for the mean, median, mode inequality/ordering (as proposed by Pearson) to hold good and hence relevant to a study of skewness (see, e.g., [12]).

5 Discussion

In this section, we investigate the applicability of the concept of homogeneous skewness proposed here to distribution functions other than continuous unimodal distributions. In Section 5.1 we consider distributions with a flat modal region and see that a minor adjustment in the skewness function (2) is enough to deal with them. In

Section 5.2 we touch upon discrete distributions to observe that while there is only a minor challenge from theoretical angle, the concept of homogeneous skewness is much less applicable in terms of standard discrete distributions. In this regard, special attention is devoted to the Poisson distribution. Section 5.3 deals with distributions having a unique antimode.

5.1 Distributions with Flat Modal Regions

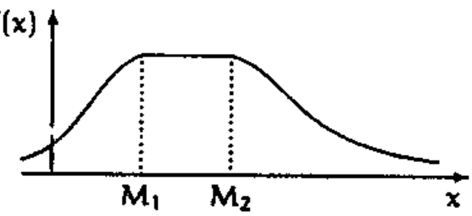

Figure 5: Density function of a distribution with flat modal region

It is easy to extend our definition to distributions with a flat modal region instead of a unique mode (see Figure 5). For these distributions, the density function $f(\cdot)$ is nondecreasing on $(-\infty, M_1)$, non-increasing on (M_2, ∞) and constant on (M_1, M_2) with $f(x) < F(M_1)$, $\forall x \notin (M_1, M_2)$. Then the skewness function (2) may be modified to:

$$\gamma_f(x) = f(M_2 + x) - f(M_1 - x),$$

and subsequently the measure of homogeneous skewness may be defined as before.

5.2 Discrete Distributions

The concept of homogeneously skewed distributions introduced in this article carries over naturally to discrete distributions *per se*, whose density function with respect to an appropriate counting measure, $f(\cdot)$, also known as the probability mass function, is assumed to have a peak at M. For such distributions, the skewness function $\gamma(\cdot)$ and the measure τ remains unaltered. The only measure difference is in terms of interpreting the values since, the extreme values ± 1 are no longer feasible. It is possible to take care of this aspect by defining the measure for discrete unimodal distributions as follows:

Definition 5 *For a discrete distribution in $\mathcal{F}$ with unique mode M and mass function $f(\cdot)$, the measure of homogeneous skewness may be defined as*

$$\tau_f = \sum_{x>0}\{f(M+x) - f(M-x)\} + f(M) \times \text{Sign}\Big(\sum_{x>0}\{f(M+x) - f(M-x)\}\Big). \quad (12)$$

Note that the second term in (12) has no impact on the sign of the skewness measure and is brought in only to ensure that the skewness of a decreasing/increasing density is 1 or -1. However, the measure will satisfy the other properties proved in Section 2.2 with or without this additional term.

The other cause of concern is from the point of view of scope since, unlike the continuous case, the standard discrete distributions are not necessarily homogeneously skewed.

Consider, for example, the case of a Poisson distribution which is positively skewed as per most standard measures (although the Pearson's measure of skewness would be zero if the mean is integral). However, for a Poisson distribution with mean 1.25 (which has mode at 1), $f(2) - f(0) < 0$, and thus this is not homogeneously skewed as per our measure. In fact the following theorem characterizes the homogeneous skewness for the Poisson distribution.

Theorem 5 *Consider a Poisson distribution with mean $\mu = \lfloor\mu\rfloor + \theta$ with $0 < \theta \leq 1$. Then there exists a cutoff $\xi \leq 0.5$ such that the Poisson distribution is homogeneously skewed to the right if and only if $\theta > \xi$. ξ is a function of $\lfloor\mu\rfloor$. Also, $\xi \uparrow 0.5$ as $\lfloor\mu\rfloor \uparrow \infty$.*

Proof: The density function of a Poisson distribution with mean μ (and mode $M = \lfloor\mu\rfloor$) is given by

$$f(x) = \begin{cases} \frac{\mu^x \exp(-\mu)}{x!} & x = 0, 1, 2, \ldots \\ 0 & \text{otherwise.} \end{cases}$$

Clearly, to prove homogeneous skewness or otherwise, we only need to prove that $f(\lfloor\mu\rfloor + y) - f(\lfloor\mu\rfloor - y) \geq 0$ when $y \leq \lfloor\mu\rfloor$. In the remainder of the proof, let us assume therefore that $y \le \lfloor\mu\rfloor$. Then

$$
\begin{aligned}
& f(M+y) - f(M-y) \\
&= \frac{\mu^{\lfloor\mu\rfloor+y}\exp(-\mu)}{(\lfloor\mu\rfloor+y)!} - \frac{\mu^{\lfloor\mu\rfloor-y}\exp(-\mu)}{(\lfloor\mu\rfloor-y)!} \qquad (13)\\
&= \frac{\mu^{\lfloor\mu\rfloor-y}\exp(-\mu)}{(\lfloor\mu\rfloor-y)!}\Big(\frac{\mu^{2y}}{(\lfloor\mu\rfloor-y+1)\dots(\lfloor\mu\rfloor+y)} - 1\Big) \\
&= \frac{\mu^{\lfloor\mu\rfloor-y}\exp(-\mu)}{(\lfloor\mu\rfloor-y)!}\Big(\prod_{k=1}^{y}\frac{(\lfloor\mu\rfloor+\theta)^2}{(\lfloor\mu\rfloor-k+1)(\lfloor\mu\rfloor+k)} - 1\Big) \qquad (14)\\
&= \Xi(\lfloor\mu\rfloor, y)\Big(\prod_{k=1}^{y}\psi_k(\mu) - 1\Big) \quad \text{say.}
\end{aligned}
$$

Differentiating expression (13) w.r.t. μ keeping $\lfloor\mu\rfloor$ and y constant, we see that $f(M+y) - f(M-y)$ increases with θ for a fixed $\lfloor\mu\rfloor$. We also see that $\Xi(\lfloor\mu\rfloor, y)$ is always positive. When $0.5 \le \theta < 1$, and $0 < k \le \lfloor\mu\rfloor$

$$\psi_k(\mu) = \frac{\mu^2}{(\lfloor\mu\rfloor-k+1)(\lfloor\mu\rfloor+k)} \ge \frac{(\lfloor\mu\rfloor+\frac{1}{2})^2}{(\lfloor\mu\rfloor+\frac{1}{2}-\frac{2k-1}{2})(\lfloor\mu\rfloor+\frac{1}{2}+\frac{2k-1}{2})} \ge 1.$$

The last two observations implies that when $0.5 \le \theta < 1$, expression (14) is always non-negative and the Poisson distribution is homogeneously right-skewed. This implication taken together with the first observation implies that a cutoff $\xi \le 0.5$ exists for each value of $\lfloor\mu\rfloor$, above which the Poisson distribution is homogeneously right skewed, thus proving the first part of the theorem.

In order to prove the second part of the theorem, notice that $\psi_k(\mu)$ is decreasing in k for a fixed μ. It follows from expression (14) that $f(M+y) < f(M-y)$, for some integral y in $(0, M]$ if and only if $\psi_1(\mu) < 1$ which is equivalent to $f(\lfloor\mu\rfloor+1)$ being less than $f(\lfloor\mu\rfloor-1)$. Now observe that when $\theta = 0.5$, $\psi_1(\mu) \downarrow 1$ as $\lfloor\mu\rfloor \to \infty$; indeed with any fixed $\theta \in (0, 0.5)$, $\psi_1(\mu)$ is a decreasing sequence in $\lfloor\mu\rfloor$ at least as long as $\psi_1(\mu) \ge 1$. This concludes the proof of the theorem. □

Since a Binomial distribution with large n and small p can be well approximated by a Poisson distribution, it is easy to find similar examples of either kind in the case of Binomial distribution. However, a complete characterization, similar to Theorem 5 seems elusive.

5.3 Distributions with Unique Antimodes

A distribution with p.d.f. $f(\cdot)$ is said to have a unique antimode $\bar{M}$ if $f(\cdot)$ is decreasing (non-increasing) on $(a, \bar{M})$ and increasing (non-decreasing) on $(\bar{M}, b)$ (see Figure 6).

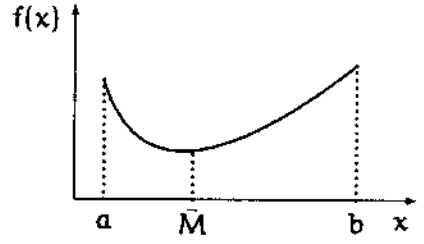

Figure 6: A distribution with a unique antimode

Such distributions seldom arise in practice, and the concept of skewness for these distributions is not clear. However the following modification of the skewness function (2) is intuitive:

$$\gamma_f(x) = f(\bar{M} - x) - f(\bar{M} + x), \quad x > 0. \tag{15}$$

With this modification, τ may be defined through $\gamma_f(\cdot)$ as before. A possible distribution family in standardized form is depicted in Figure 7. This family may be referred to as inverted Triangular distributions, parameterized by slopes S_1 and S_2 and the antimode $\bar{M}$. In Figure 7a we have a homogeneously skewed distribution to the right and in Figure 7b we have a homogeneously skewed distribution to the left. For $S_1 > S_2, \bar{M} < 0.5$, or $S_1 < S_2, \bar{M} > 0.5$, the distribution is not homogeneously skewed.

The Beta distribution with α and β both less than 1 is one well known distribution family with unique antimode. It turns out not to be homogeneously skewed although, the skewness function does not change sign in $(0, \min(\bar{M}, 1 - \bar{M}))$. Indeed this aspect of change of sign of the skewness function at the critical point of support of the distribution, may require further exploration, as do a more complete treatment of skewness of distributions with unique antimode.

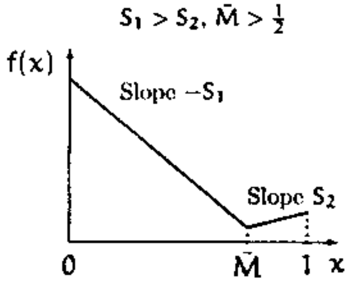

Figure 7a

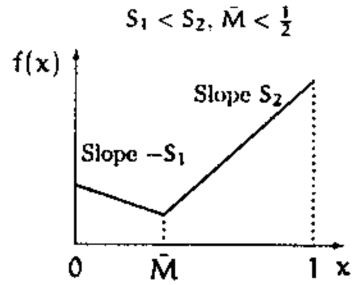

Figure 7b

Figure 7: Inverted triangular distributions

6 Summary and Directions for Future Research

In this paper, we proposed a new measure of skewness for unimodal distributions based on the asymmetry of the density function around the mode. In the first section, we briefly reviewed the literature on measures of skewness, and presented some properties of such measures. In the second section, we formally presented our skewness function (equation (2)) and measure of skewness (equation (3)), and derived the properties of this measure. Through Theorem 1, we showed that our measure lies between -1 and $+1$ for continuous unimodal distributions, with boundary values being achieved for increasing and decreasing densities and the measure equals 0 for symmetric distributions. The measure is shown to be location and scale invariant in Theorem 2. Based on our skewness function, we have also proposed partial/total ordering of homogeneously skewed distributions in this section. In Section 3 we show that our notion of homogeneous skewness covers the Triangular, Beta, Gamma, and Lognormal distributions, but the Weibull distribution is not always homogeneously skewed. In Section 4, our measure is compared with other standard measures of skewness like the Pearson's skewness measure (Theorem 3) and the third standardized moment about the mode (Theorem 4), and is found to be stricter than both. In the previous section, we explored few extensions and modifications to broaden the applicability of the proposed measure. However, many standard discrete distributions do not appear to be homogeneously skewed — Theorem 5 shows for example, that the Poisson distribution is homogeneously skewed roughly half the time.

The characterization of homogeneous skewness or lack of it for other asymmetric distributions, both discrete and continuous, is a direction in which this study can be carried forward. Among the ones we touched upon, this is specially the case with Binomial and Weibull distributions.

Some common distributions viz. Gamma distributions, Beta distributions, exhibit a stronger form of skewness than homogeneous skewness. This is in the sense that the skewness functions is not only of the same sign but also increasing (in absolute value) as shown through (8) and (9). This may be termed as homogeneous skewness of second degree/order and explored further.

References

[1] Aucremanne, L., Brys, G., Hubert, M., Rousseeuw, P.J., and Struyf, A., A study of Belgian inflation, relative prices and nominal rigidities using new robust measures of skewness and tail weight, in Theory and Applications of Recent Robust Methods, eds. M. Hubert, G. Pison, A. Struyf and S. van Aelst, Statistics for Industry and Technology, Birkhauser: Basel (2004).

[2] Benjamini, Y. and Krieger, A.M., Skewness — Concepts and measures, in Encyclopaedia of Statistical Sciences, eds. N.L.Johnson, and S.Kotz. John Wiley, 663–670, 1985.

[3] Bowley, A.L., Elements of Statistics, Charles Scriber's Sons, NY, 1901.

[4] Brys, G., Hubert, M., Struyf, A., A robust measure of skewness, to appear in Journal of Computational and Graphical Statistics (2003).

[5] Brys, G., Hubert, M., and Struyf, A., A comparison of some new measures of skewness, in Developments in Robust Statistics, ICORS 2001, eds. R. Dutter, P. Filzmoser, U. Gather and P.J. Rousseeuw, Springer-Verlag: Heidelberg, 98-113, 2003.

[6] David, F.N., and Johnson, N.L., Some tests of significance with ordered variables, J. R. Statist. Soc. B, 18, 1–20, 1956.

[7] Doksum, K.A., Measures of location and asymmetry, Scand. J. Statist., 2, 11–22, 1975.

[8] Groeneveld, R. A., Bowley's measures of skewness, in Encyclopaedia of Statistical Sciences, eds. N.L. Johnson, and S. Kotz, John Wiley, 619–621, 1985.

[9] Johnson, N.L., and Kotz, S., Unimodality, in Encyclopaedia of Statistical Sciences, eds. N.L. Johnson, and S. Kotz. John Wiley, 415–416, 1985.

[10] Kenney, J. F. and Keeping, E. S., Mathematics of Statistics, Pt. 1, 3rd ed. Princeton, NJ: Van Nostrand, 101–102, 1962.

[11] MacGillivray, H.L., Skewness and asymmetry: Measures and orderings, Annals of Statist., 14, 994–1011, 1986.

[12] MacGillivray, H.L., Mean, median, mode and skewness, in Encyclopaedia of Statistical Sciences, eds. N.L. Johnson, and S. Kotz. John Wiley, 364–367, 1985.

[13] Oja, H., On location, scale, skewness and kurtosis of univariate distributions, Scand. J. Statist. 8, 154–168, 1981.

[14] Pearson, K., Contributions to the mathematical theory of evolution: Skew variation in homogeneous material, Philos. Trans. R. Soc. London, 186, 343, 1895.

[15] van Zwet, W.R., Mean, median, mode II, Statist Neerl. 33, 1–5, 1979.

[16] Yule, G.U. Introduction to the Theory of Statistics, Griffin, London, 1911.