Skewness and Asymmetry: Measures and Orderings

Article in The Annals of Statistics · September 1986
DOI: 10.1214/aos/1176350046 · Source: OAI

H. L. Macgillivray
Queensland University of Technology

The Annals of Statistics
1986, Vol. 14, No. 3, 994–1011

SKEWNESS AND ASYMMETRY: MEASURES AND ORDERINGS[1]

BY H. L. MACGILLIVRAY

University of Queensland

Recent interest in skewness has tended to separate two aspects of the concept. Two distributions may be compared with respect to skewness, or a distribution may be self-compared, that is, the distributions of the random variables of X and $-X$ may be compared. This paper uses the unification of these two aspects to attempt to complete a skewness structure of orderings that identifies the roles of various skewness and scale measures and enables classification of the skewness properties of any distribution. The structure is also used to propose measures of asymmetry. Some skewness properties of the Weibull and Johnson systems are examined.

1. Introduction. Sections 1.1, 1.2, and 1.3 give a brief historical survey of work on skewness, including measures, partial orderings, and relationships between location parameters. With reference to this, Section 1.4 outlines the aims and content of this paper.

1.1. *Classical measures.* As for location, scale, and kurtosis, the concept of skewness was introduced with an apparently appropriate measure. Pearson (1895) proposed $(\mu - M)/\sigma$ as a measure of skewness for a univariate distribution with mean μ, mode M, and variance σ^2. Three other measures of skewness appear to have been introduced soon afterwards [Bowley (1901), pages 116 and 251, and Yule (1911), page 162]. These are $(\mu - m)/\sigma$, where m is the median, μ_3/σ^3, where μ_3 is the third central moment, and $(q_u + q_l - 2m)/(q_u - q_l)$, where q_u, q_l are the upper and lower quartiles, respectively. All are based on the criteria that a skewness measure should be scale-free and zero for symmetric distributions.

The initial roles of $(\mu - m)/\sigma$ and μ_3/σ^3 seem to lie in their relationships [empirical in the case of $(\mu - m)/\sigma$] to the Pearson skewness for distributions of the Pearson family. Gradually μ_3/σ^3 assumed more prominence as "the skewness," as is illustrated by the papers of Doodson (1917) and Haldane (1942), who, in examining Pearson's empirical relationship between $M - m$ and $M - \mu$, used $(\mu - M)/\sigma$ and μ_3/σ^3, respectively, to measure skewness. For $(\mu - m)/\sigma$, Hotelling and Solomons (1932) and Garver (1932) showed $|(\mu - m)/\sigma| \le 1$ with this result being refined by Majindar (1962).

The quartile measure of skewness was introduced through its sample version as a descriptive statistic and the bound of 1 on its absolute value was regarded as

Received January 1985; revised October 1985.

[1]Part of this research was done while the author was on leave at Heriot-Watt University, Edinburgh.

AMS 1980 *subject classifications.* Primary 62E10; secondary 60E05.

Key words and phrases. Asymmetry, skewness, scale, partial orderings of distributions, mean, median, moments, quantiles.

an advantage. It was used to test symmetry by David and Johnson (1954), who also suggested (1956) its generalisation to a quantile measure which is defined for a continuous random variable with distribution function F by

$$\gamma_u(F) \equiv \left[F^{-1}(1-u) + F^{-1}(u) - 2m_F\right]/\left[F^{-1}(1-u) - F^{-1}(u)\right], \quad u \in \left(0, \tfrac{1}{2}\right). \tag{1.1}$$

These measures have been used by Hinkley (1975), Hogg (1974) and Hogg et al. (1975).

Yule (1911, page 162) noted that the measures $(q_u + q_l - 2m)/(q_u - q_l)$ and $(\mu - M)/\sigma$ "are positive if the longer tail of the distribution lies toward the higher values of the variate." Fechner (1897) and Timerding (1915) examined asymmetric densities for which μ, m, and M occur in this or the reverse order, but their papers seemed to be little known. For different reasons, a number of authors [Groeneveld and Meeden (1977), Runnenburg (1978), and MacGillivray (1979)] almost simultaneously considered conditions giving the sign of $\mu - m$ and $m - M$, or μ_3. van Zwet (1979) showed the link to his partial ordering (see Section 1.2), and MacGillivray (1981) the link to the properties of the strictly totally positive kernel x^r [Karlin (1968), page 21]. It is curious that although asymmetric distributions with $\mu_3 = 0$ were familiar [see, for example, Ord (1968) and Johnson and Kotz (1970), page 253], conditions sufficient to determine the sign of μ_3 were not previously established.

1.2. *Partial orderings of distributions.* As noted by van Zwet (1964) and Oja (1981), the measures of Section 1.1 were equated with a concept without answers to the questions: when are the measures appropriate representatives of the concept, and, if not, what can be used? Each measure either assumes or imposes an ordering between distributions. van Zwet (1964) used convex transformations to examine and formalise the concepts of skewness and kurtosis, claiming that such concepts require meaningful orderings of distributions which measures of skewness or kurtosis must then preserve. For distribution functions F and G which have densities with an interval support, van Zwet defined G as having greater skewness to the right than F, if $G^{-1}(F(x))$ is convex on $I_F = \{x\colon 0 < F(x) < 1\}$, and showed that the standardised odd central moments preserve this ordering. That is, if $G^{-1}(F(x))$ is convex on I_F, then

$$\mu_{F,2k+1}/\sigma_F^{2k+1} \le \mu_{G,2k+1}/\sigma_G^{2k+1}, \qquad k = 1, 2, \dots .$$

This ordering with respect to the exponential distribution gives increasing failure rate (IFR) and decreasing failure rate (DFR) distributions and van Zwet's orderings, plus others based on star-shaped transformations, have been used in reliability theory [for example, Barlow and Proschan (1966)], power of rank tests [Doksum (1969)], and inequalities on order statistics [Lawrence (1975)].

Mann and Whitney (1947) had previously introduced the notion of "stochastic ordering," a partial ordering of distributions with respect to location. Bickel and Lehmann (1975) used this in their work on measures of location for asymmetric distributions, and also introduced (1976) a partial ordering for distributions, symmetric or asymmetric, with respect to spread or dispersion. Oja (1981) unified the work on partial orderings of distributions according to the different attributes of location, scale, skewness, showing that the definitions introduced by the authors above correspond to convexity of increasing order [Karlin (1968), page 23] of the same function, namely

$$\Delta_{F,G}(x) = G^{-1}(F(x)) - x \quad \text{on } I_F.$$

Oja introduced some weaker orderings for scale, skewness, and kurtosis, and discussed how to find measures that preserve the various orderings, concentrating on measures based on moments.

1.3. *Measures based on quantiles.* The function $\Delta(x)$ had been previously considered in the special case of $G(x) = 1 - F(-x) = \bar{F}(x)$ by Doksum (1975) in considering real-valued functionals satisfying the usual location axioms, and their values, giving the location set. The symmetry function $\theta_F(x) = \frac{1}{2}[x - \bar{F}^{-1}(F(x))]$ has the location set as the closure of its range, and is the only function such that $x - 2\theta_F(x)$ is nondecreasing a.e. (with respect to F), and $X - 2\theta_F(X)$ has the distribution of $-X$. Doksum used the difference between $\theta_F(x)$ and the median m_F as a measure of asymmetry, and defined F as being strongly skewed to the right if and only if $\theta_F(x)$ is nonincreasing for $x < m_F$ and nondecreasing for $x \geq m_F$, and skewed to the right if and only if $\theta_F(x) \geq m_F$, ($\theta_F(m_F) = m_F$). These definitions are weaker than that due to van Zwet (1979) and used by Oja (1981), that F is skewed to the right if there exists a symmetric distribution F_0 such that $F^{-1}(F_0(x))$ is convex on I_{F_0}. As stated by van Zwet (1979), this is equivalent to F more skewed to the right than $\bar{F}$, the distribution of $-X$.

Doksum's (1975) weaker definition of skewness to the right is equivalent to van Zwet's (1979) condition, $F^{-1}(t) + F^{-1}(1 - t) \geq 2m_F$, for the mean, median, and mode to occur in reverse order, so that $(\mu_F - M_F)/\sigma_F$ and $(\mu_F - m_F)/\sigma_F$ are nonnegative. Doksum's symmetry function and index of skewness are related to the quantile measures of skewness, $\gamma_u(F)$ of (1.1). This is discussed in Section 2.

1.4. *Outline of paper.* Thus against a background of a variety of motivations and applications, a variety of skewness measures and/or orderings have been used. Some links have been given but the structure is not complete. There appears to be reasonably general agreement on two points. First, comparing the skewness of two distributions refers, whether implicitly or explicitly, to a partial ordering of distributions with respect to skewness, and a measure of skewness should preserve an acceptable ordering. Second, discussions of the skewness of a distribution to the left or right, refer to skewness comparisons of the distributions of X and $-X$, described here as self-comparisons. Therefore, for consistency, orderings and measures established in one context should apply also to the other. In particular, meaningful interpretation of a measure's numerical value and sign requires knowledge of the orderings it preserves.

Section 2 therefore aims to identify a skewness structure according to the following criteria:

(a) The same structure should apply to comparisons of different distributions and to self-comparisons for a single distribution.
(b) It should identify the roles of skewness measures and partial orderings previously introduced in either of the contexts in (a), and provide a hierarchy of such.
(c) No skewness measure should be used without identification of the ordering it preserves, and this ordering should aim to cover as large a class of distributions as possible.
(d) It should be possible to describe the skewness of any asymmetric distribution.
(e) The structure should not be more complicated than is necessary to meet (a)–(d).

In certain circumstances arbitrary distributions have been considered by some authors, but there are complications in the general structure, and for simplicity attention is confined here to the class, $\mathscr{F}$, of distribution functions $F(x)$ which have probability density functions $f(x)$ with an interval support.

Although not necessarily complete the resultant hierarchy of orderings is reasonably extensive. Some orderings may be more useful in practice than others, but the hierarchy is an important background structure underlying any description of skewness properties; like convergence, skewness has a range of "strengths."

On the basis of the skewness structure, and ideas in Doksum (1975), Section 2.4 discusses measures of asymmetry.

Section 3 examines the skewness properties of some distribution families, not only to illustrate various aspects of the skewness structure but also to increase the knowledge of the properties of these distributions. The Weibull family and Section 2.4 indicate that for a complete understanding of the skewness properties of some distribution families, it may be necessary to consider both self-comparisons and comparisons between the distributions.

2. Orderings and measures. van Zwet's (1964) partial ordering with respect to skewness on $\mathscr{F}$ is the strongest that has been considered and it is unlikely that anything stronger is useful. Using a slight modification of Oja's (1981) notation, $F \leq_2 G$ iff $G^{-1}(F(x))$ is convex on I_F, and F is said to be not more strongly skew to the right than G. "Convex" here includes the possibility of linearity. It may sometimes be convenient to define $F <_2 G$ iff $G^{-1}(F(x))$ is strictly convex, that is convex but not linear, on I_F, in which case G is more strongly skew to the right than F.

An equivalent definition of $F \leq_2 G$ is that $F(x)$ and $G(ax + b)$ cross each other at most twice for any a, b, or, using Karlin's (1968, pages 20 and 280–282) S notation, $S^{-}(F(x) - G(ax + b)) \leq 2$ for all a, b, with the sequence of signs being positive to negative to positive when equality holds. Thus van Zwet's skewness ordering has no reference to any measures of location and scale, and any weakening of the ordering in the sense of covering larger classes of distributions, involves reference to particular location and scale parameters.

2.1. *Orderings with respect to the mean.* Oja's (1981) ordering $\lesssim_2^*$ is a weakening of $\leq_2$ that is still preserved by the standardized odd moments μ_{2k+1}/σ^{2k+1}. The formal definition is $F \lesssim_2^* G$ iff there exists $x_1 \leq \mu_F \leq x_2$ such that

$$G^{-1}(F(x)) - x \lesseqgtr \frac{\sigma_G - \sigma_F}{\sigma_F} x + \left(\mu_G - \frac{\sigma_G}{\sigma_F}\mu_F\right) \tag{2.1}$$

$$\text{resp. for } x_1 \leq x < x_2, \quad x < x_1 \quad \text{or} \quad x \geq x_2.$$

Equivalently the standardised distributions $F(\mu_F + \sigma_F x)$ and $G(\mu_G + \sigma_G x)$ cross each other exactly once on each side of $x = 0$, with $F(\mu_F) \leq G(\mu_G)$. However, the restriction on the positions of the crossings, although implied if $F <_2 G$, is not required for an ordering preserved by the standardised odd moments.

From MacGillivray (1985), $M_{F,G} \equiv F(\mu_F + \sigma_F x) - G(\mu_G + \sigma_G x)$ is either identically zero or changes sign at least twice; if exactly twice, say from ≥ 0 to ≤ 0 to ≥ 0, then $\mu_{F,2k+1}/\sigma_F^{2k+1} < \mu_{G,2k+1}/\sigma_G^{2k+1}$, $k = 1, \dots$. Hence skewness with respect to the mean may be defined as follows:

DEFINITION 2.1. G is more skew to the right with respect to the mean than F, $F <_2^{\mu} G$, iff

$$S^-(M_{F,G}) = 2 \quad \text{from } \geq 0 \text{ to } \leq 0 \text{ to } \geq 0. \tag{2.2}$$

More generally, $F \leq_2^{\mu} G$ iff (2.2) holds or $M_{F,G} \equiv 0$.

THEOREM 2.1.

$$F \leq_2^{\mu} G \Rightarrow \mu_{F,2k+1}/\sigma_F^{2k+1} \leq \mu_{G,2k+1}/\sigma_G^{2k+1}, \qquad k = 1, \dots,$$

and if equality holds for any k, then $M_{F,G} \equiv 0$.

PROOF. This result is a special case of Theorem 1 of MacGillivray (1985). □

In the case of self-comparisons, $F \lesssim_2^* \overline{F} \Leftrightarrow F \leq_2^{\mu} \overline{F} \Leftrightarrow S^-(1 - F(\mu_F + x) - F(\mu_F - x)) \leq 1$, from ≥ 0 to ≤ 0, which is the established sufficient condition for $\mu_{F,2k+1} \leq 0$.

2.2. *Orderings and measures with respect to the median.* van Zwet (1964, page 16) gives an example using a discrete distribution that shows that a nondecreasing convex transformation of a random variable does not necessarily increase $(\mu - m)/\sigma$. An example using continuous distributions is provided by the transformation $Y = \exp X$ of an exponential random variable with mean $1/\lambda$. As $\lambda \to 2$, $(\mu_Y - m_Y)/\sigma_Y \to 0$, and hence there are values of λ for which $(\mu_Y - m_Y)/\sigma_Y < (\mu_X - m_X)/\sigma_X = 1 - \log 2$. So $(\mu - m)/\sigma$ does not preserve $\leq_2$ (nor therefore $\leq_2^{\mu}$), but because there exists a condition on a distribution sufficient for $\mu \lesseqgtr m$, the scale measure appears to be at fault.

LEMMA 2.1. *If*

$$(2.3)\qquad (G^{-1}(u) - m_G)/\eta_G \geq (F^{-1}(u) - m_F)/\eta_F, \qquad u \in (0,1),$$

where $\eta_F, \eta_G > 0$, *then* $(\mu_G - m_G)/\eta_G \geq (\mu_F - m_F)/\eta_F$.

PROOF. Follows from $\mu_F - m_F = \int_0^1 (F^{-1}(u) - m_F)\, du$. □

LEMMA 2.2.

$$F \leq_2 G \Rightarrow (2.3) \quad \textit{iff } \eta_G/\eta_F = f(m_F)/g(m_G),$$

where f, g *are the probability density functions of* F, G.

PROOF. $G^{-1}(F(x))$ convex on I_F is equivalent to

(2.4) $[G^{-1}(F(x)) - G^{-1}(F(y))]/(x - y)$ is nondecreasing for $x \in I_F$ for any y.

Hence, taking $y = m_F$, $F \leq_2 G \Rightarrow$ (2.3) iff

$$\begin{aligned}\eta_G/\eta_F &= \lim_{u \to 1/2} (G^{-1}(u) - m_G)/(F^{-1}(u) - m_F)\\ &= f(m_F)/g(m_G),\end{aligned}$$

since f, g are nonzero in I_F, I_G for $F, G \in \mathscr{F}$. □

The above lemmas provide an ordering that weakens $\leq_2$ and is preserved by $(\mu - m)/$(a measure of scale), which therefore also preserves $\leq_2$. The weakening procedure is different to that of Section 2.1, and starts from a different characterisation of $\leq_2$, given by (2.4). There are intermediate steps between (2.4) and (2.3) and each has been used in a skewness context.

Oja's (1981) star-ordering, $\lesssim_{\text{star}}$ [which generalises the star-shaped orderings of, for example, Barlow and Proschan (1966), Doksum (1969), and Lawrence (1975)] is (2.4) for some value of y. However, the value of y is important, since (2.4) for a particular y is a weakening of $\leq_2$ to considering skewness with respect to y. This point then becomes the particular location parameter around which the skewness is taken. Taking $y = m_F$, the following definitions of skewness progressively weaken $\leq_2$ with respect to the median.

DEFINITION 2.2. $F \underset{\text{star}}{\leq_2^m} G$ iff

$$(2.5)\qquad [G^{-1}(F(x)) - m_G]/(x - m_F) \quad \text{is nondecreasing in } I_F,$$

or, equivalently,

$$(2.6)\qquad (G^{-1}(u) - m_G)/(F^{-1}(u) - m_F) \quad \text{is nondecreasing in } (0,1).$$

DEFINITION 2.3. $F \underset{D}{\leq_2^m} G$ iff

$$(2.7)\quad G^{-1}(F(x)) - [f(m_F)/g(m_G)]x \text{ is } \begin{Bmatrix}\text{nonincreasing}\\ \text{nondecreasing}\end{Bmatrix} \text{ for } x \begin{Bmatrix} \leq \\ \geq \end{Bmatrix} m_F \text{ in } I_F,$$

or, equivalently,

$$G^{-1}(u) - [f(m_F)/g(m_G)]F^{-1}(u) \text{ is } \left\{\begin{matrix}\text{nonincreasing}\\ \text{nondecreasing}\end{matrix}\right\} \quad \text{for } u\left\{\begin{matrix}\le\\ \ge\end{matrix}\right\}\tfrac{1}{2} \text{ in } (0,1). \tag{2.8}$$

DEFINITION 2.4. $F \le_2^m G$ iff

$$[G^{-1}(F(x)) - m_G]g(m_G) \ge (x - m_F)f(m_F) \quad \text{for } x \in I_F, \tag{2.9}$$

or, equivalently, for $u \in (0,1)$,

DEFINITION 2.5. $F \underset{\gamma}{\le_2^m} G$ iff

$$\begin{aligned}&[G^{-1}(1 - F(x)) - m_G]/[F^{-1}(1 - F(x)) - m_F]\\ &\qquad \ge [G^{-1}(F(x)) - m_G]/[x - m_F] \quad \text{for } x \le m_F \text{ in } I_F\end{aligned} \tag{2.11}$$

or, equivalently, for $u \in (0, \frac{1}{2})$,

$$\begin{aligned}&[G^{-1}(1 - u) - m_G]/[F^{-1}(1 - u) - m_F]\\ &\qquad \ge [G^{-1}(u) - m_G]/[F^{-1}(u) - m_F].\end{aligned} \tag{2.12}$$

THEOREM 2.2.

$$F \le_2 G \Rightarrow F \underset{\text{star}}{\le_2^m} G \Rightarrow F \underset{D}{\le_2^m} G \Rightarrow F \le_2^m G \Rightarrow F \underset{\gamma}{\le_2^m} G.$$

PROOF. (2.5) is (2.4) for $y = m_F$.
Differentiating (2.5) gives

$$\{f(x)/g[G^{-1}(F(x))] - [G^{-1}(F(x)) - m_G]/(x - m_F)\}/(x - m_F) \ge 0 \quad \text{in } I_F,$$

which gives $f(x)/g[G^{-1}(F(x))]\left\{\begin{matrix}\ge\\ \le\end{matrix}\right\}f(m_F)/g(m_G)$ for $x\{\gtrless\}m_F \Rightarrow$ (2.7).

Clearly (2.7) $\Rightarrow$ (2.9) $\Leftrightarrow$ (2.10) $\Rightarrow$ (2.12) $\Leftrightarrow$ (2.11). $\square$

Definition 2.3 gives the ordering generalised from Doksum's (1975) definition of strong skewness to the right, namely,

$$F^{-1}(\tfrac{1}{2} + u) + F^{-1}(\tfrac{1}{2} - u) \quad \text{is nondecreasing for } u \in [0, \tfrac{1}{2}]. \tag{2.13}$$

This is equivalent to $\bar{F} \underset{D}{\le_2^m} F$. The strength of skewness given by (2.13) has also been mentioned by van Zwet (1979) and used by David and Groeneveld (1982).

Definition 2.4 generalises the two conditions $1 - F(m + x) - F(m - x) \ge 0$ and $F^{-1}(1 - u) + F^{-1}(u) - 2m_F \ge 0$, which were given by van Zwet (1979) as sufficient for $\mu_F \ge m_F$, and which were Doksum's (1975) definitions of F skew to the right. Definition 2.4 plays the role of $\le_2^\mu$ but with respect to the median instead of the mean.

It should be noted again that in self-comparisons, that is, comparing F and $\bar{F}$, the measure of scale is not relevant and only the *sign* of a skewness measure is relevant. If a skewness measure is quoted quantitatively it immediately implies comparisons with other distributions and the measure of scale then plays an essential role.

Groeneveld and Meeden (1984) show that $\gamma_u(\cdot)$ and $(\mu - m)/E|X - m|$ preserve $\leq_2$; Lemmas 2.1 and 2.2 introduce $(\mu - m)f(m_F)$. Theorem 2.3 below identifies the roles of various skewness measures in the heirarchy of skewness orderings given above. If a measure preserves an ordering it also preserves the preceding orderings in the hierarchy of Theorem 2.2. Apart from preserving an acceptable skewness ordering, Oja (1981) also suggested that skewness measures should change only by multiplication by sign(a) under a linear transformation $ax + b$ of the random variable; all the skewness measures below satisfy this requirement. In some of the measures in Theorem 2.3, the mean and the quantile average $F^{-1}(1 - u) + F^{-1}(u) - 2m_F$ are generalised to the symmetrically weighted quantile averages defined by

$$\mu_K(F) \equiv \int_0^1 F^{-1}(u)\,dK(u) = \int_0^{1/2} \left[F^{-1}(1 - u) + F^{-1}(u)\right] dK(u), \tag{2.14}$$

where $K(u)$ is a distribution function on $(0,1)$ symmetric around $\frac{1}{2}$.

THEOREM 2.3.

(a) $$F \underset{\gamma}{\leq_2^m} G \Leftrightarrow \gamma_u(F) \leq \gamma_u(G) \quad \textit{for } u \in \left(0, \tfrac{1}{2}\right);$$

(b)(i) $$F \leq_2^m G \Rightarrow (\mu_K(F) - m_F)f(m_F) \leq (\mu_K(G) - m_G)g(m_G),$$

(ii) $$F \leq_2^m G \Rightarrow E(X - m_F)^{2k+1} f(m_F) \leq E(Y - m_G)^{2k+1} g(m_G), \qquad k = 1, 2, \ldots,$$

(iii) $$F \leq_2^m G \Rightarrow (\mu_F - m_F)/E|X - m_F| \leq (\mu_G - m_G)/E|Y - m_G|,$$

where $X, Y \sim F, G$.

(c) $$\begin{aligned} F \underset{\text{star}}{\leq_2^m} G \Rightarrow & \left[F^{-1}(1 - u) + F^{-1}(u) - F^{-1}(1 - \alpha) - F^{-1}(\alpha)\right] \\ & /\left[F^{-1}(1 - \alpha) + F^{-1}(\alpha)\right] \\ & \leq \left[G^{-1}(1 - u) + G^{-1}(u) - G^{-1}(1 - \alpha) - G^{-1}(\alpha)\right] \\ & /\left[G^{-1}(1 - \alpha) - G^{-1}(\alpha)\right] \end{aligned}$$

for $0 \leq u \leq \alpha \leq \frac{1}{2}$.

PROOF. (a) [This result is stated in Groeneveld and Meeden (1984)]. Without loss of generality m_F and m_G may be assumed to be 0. For $u \in (0, \frac{1}{2})$, (2.12) $\Rightarrow$

$$\begin{aligned} & G^{-1}(u)F^{-1}(1 - u) - G^{-1}(1 - u)F^{-1}(u) \\ & \quad \geq G^{-1}(1 - u)F^{-1}(u) - G^{-1}(u)F^{-1}(1 - u). \end{aligned} \tag{2.15}$$

Adding $G^{-1}(1-u)F^{-1}(1-u) - G^{-1}(u)F^{-1}(u)$ to each side of (2.15) gives

$$\begin{aligned}&[G^{-1}(1-u) + G^{-1}(u)][F^{-1}(1-u) - F^{-1}(u)]\\ &\quad \geq [F^{-1}(1-u) + F^{-1}(u)][G^{-1}(1-u) - G^{-1}(u)],\end{aligned}$$

and conversely, as required.

(b)(i) follows directly from (2.10) and definition of $\mu_K(\cdot)$.

(ii) follows from (2.9) written as $F(m_F + x/f(m_F)) \geq G(m_G + x/g(m_G))$.

(iii) follows from Groeneveld and Meeden (1984) since

$$\begin{aligned}F \leq_2^m G \Leftrightarrow &[G^{-1}(1-u_1) - m_G]/[F^{-1}(1-u_1) - m_F] \geq f(m_F)/g(m_G)\\ &\geq [G^{-1}(u_2) - m_G]/[F^{-1}(u_2) - m_F] \quad \text{for } 0 < u_1, u_2 \leq \tfrac{1}{2}.\end{aligned}$$

(c)
$$\begin{aligned}F \underset{\text{star}}{\leq_2^m} G \Rightarrow &f(F^{-1}(1-v))/g(G^{-1}(1-v))\\ &\geq [G^{-1}(1-\alpha) - m_G]/[F^{-1}(1-\alpha) - m_F]\\ &\geq [G^{-1}(\alpha) - m_G]/[F^{-1}(\alpha) - m_F]\\ &\geq f(F^{-1}(v))/g(G^{-1}(v)) \quad \text{for } 0 < v \leq \alpha \leq \tfrac{1}{2}.\end{aligned}$$

Hence,

$$\begin{aligned}&[1/g(G^{-1}(1-v)) - 1/g(G^{-1}(v))][F^{-1}(1-\alpha) - F^{-1}(\alpha)]\\ &\quad \geq [1/f(F^{-1}(1-v)) - 1/f(F^{-1}(v))][G^{-1}(1-\alpha) - G^{-1}(\alpha)].\end{aligned} \tag{2.16}$$

Integrating (2.16) with respect to v from u to α gives required result. □

REMARK 2.1. Three different measures of scale are incorporated in the skewness measures in Theorem 2.3, namely $1/f(m_F)$, $F^{-1}(1-u) - F^{-1}(u)$, and $E|X - m_F|$. Each of these preserves the spread-ordering due to Bickel and Lehmann (1976) which is used by Oja (1981) and is defined by

$$\begin{aligned}&F \leq_1 G \quad \text{iff } G^{-1}(F(x)) - x \text{ is nondecreasing,}\\ &\Leftrightarrow G^{-1}(v) - G^{-1}(u) \geq F^{-1}(v) - F^{-1}(u), \qquad 0 < u < v < 1.\end{aligned} \tag{2.17}$$

They also satisfy the other requirement of these authors for a scale measure in that under a linear transformation $ax + b$ of a random variable, the scale measure is changed only by multiplication by $|a|$. Oja (1981) weakens $\leq_1$ with respect to the mean to give a scale ordering $\leq_1^*$ which is still preserved by the variance; this is the scale analogy to Section 2.1 above. Similarly to Definition 2.4 above, $\leq_1$ may be weakened with respect to the median to give scale with respect to the median, defined by

DEFINITION 2.6. $F \leq_1^m G$ iff

$$G^{-1}(F(x)) - x\{\lesseqgtr\}m_G - m_F \quad \text{for } x\{\lesseqgtr\}m_F \text{ in } I_F, \tag{2.18}$$

or, equivalently,

$$G^{-1}(u) - F^{-1}(u)\{\lesseqgtr\}m_G - m_F \quad \text{for } u\{\lesseqgtr\}\tfrac{1}{2}. \tag{2.19}$$

The three scale measures appearing in Theorem 2.3 preserve $\leq_1^m$.

REMARK 2.2. Although $1/f(m_F)$ arises naturally in this context as a scale measure and has been used as such in central-density scaling [see, for example, Rogers and Tukey (1972) and Rosenberger and Gasko (1983)], it is perhaps not as appealing as $F^{-1}(1-u) - F^{-1}(u)$ or $E|X-m|$. For example, both $|(\mu - m)/E|X-m||$ and $|\gamma_u(\cdot)|$ are bounded by 1, with $|\gamma_u(F)| \to_{u \to 0} 1$ at least for semi-infinite I_F. However $1/f(m_F)$ is essentially the scale measure that arises in the orderings as $\leq_2$ is gradually weakened. This is illustrated by considering two distribution functions F and G with medians zero and the same expected absolute deviation from the median. If $F \leq_2 G$, $S^-(F-G) = 2$ so that an ordering on F and G using this scale measure would give inconsistencies with the criteria of Section 1.4.

REMARK 2.3. Oja's (1981) $\lesssim_2^{**}$ ordering is a weakening of $\leq_2$ as in Section 2.1 but $G^{-1}(F(x))$ is compared with some arbitrary line $ax+b$, thus involving particular but arbitrary measures of location and scale. Similarly, $\leq_2$ may be weakened as in this section but with respect to some quantile other than the median. To date neither of these concepts appear to have been discussed in a skewness context; however the second concept is of importance in the kurtosis context [Balanda and MacGillivray (1986)].

2.3. *Central and tail skewness.* There is still the problem of what to say about the skewness of F when it does not satisfy any of the orderings given so far with respect to $\bar{F}$. One solution is to identify a portion of the distribution centred on the median, that is, some proportion $(1-2\alpha)$ of the central part of the distribution, where F and $\bar{F}$ may be compared according to one or more of the orderings of Theorem 2.2. Doksum (1975) also suggested restricting attention to such a central portion for some fixed α, from the point of view of parameter robustness. The structure of Theorem 2.2 and measures such as $\gamma_u(\cdot)$ may all be applied to some central portion only, with notation such as $\leq_{2\,\mathrm{star}}^{m,c}$, $\leq_{2\,D}^{m,c}$, and $\leq_2^{m,c}$ and the skewness called central skewness.

For many distributions, the central skewness may cover a sufficiently wide proportion $1-2\alpha$ of the distribution for practical purposes. When this is not the case, a further description of the skewness properties may be required, in terms of changes between skewness to the left or right according to some ordering. Doksum (1975) suggested examining changes between increasing and decreasing of the symmetry function $\theta_F(x)$, thus referring to the $\leq_{2\,D}^{m}$ ordering. However in portions of the distribution that do not include the median, the orderings of Theorem 2.2 do not necessarily form a hierarchy. Since a weakening of this structure is required, it may be more appropriate and easier to consider changes of sign of $g(m_G)[G^{-1}(F(x)) - m_G] - f(m_F)(x - m_F)$, or, equivalently, and more conveniently $g(m_G)(G^{-1}(u) - m_G) - f(m_F)(F^{-1}(u) - m_F)$.

DEFINITION 2.7. G is at least as skew to the right with respect to the median as F in the proportion interval $(1-2\alpha_2, 1-2\alpha_1)$ if (2.10) holds for u and $1-u$ with $u \in [\alpha_1, \alpha_2]$.

COROLLARY 2.1. *If Definition* 2.7 *holds, then for* $u \in [\alpha_1, \alpha_2]$,

$$\gamma_u(F) \le \gamma_u(G) \quad \textit{and} \quad \nu_u(F) \le \nu_u(G), \tag{2.20}$$

where $\nu_u(F) \equiv (F^{-1}(1-u) + F^{-1}(u) - 2m_F)f(m_F)$.

In the special case $G = \overline{F}$, (2.20) is necessary and sufficient.

The definition has been given in terms of $\le_2^m$ rather than $\le_{2\,\gamma}^m$ to allow for central-density scaling.

Central skewness refers to a proportion interval $[0, 1 - 2\alpha]$; when the proportion interval is $[1 - 2\alpha, 1]$, F and G are being compared with respect to skewness in the tails. Central and tail skewness tend to be more easily established than skewness in other proportion intervals. For example, all distributions with $I_F = [0, \infty]$ are at least tail-skew to the right. The asymmetric Tukey lambda family [Ramberg et al. (1979)] is defined by

$$F^{-1}(u) = \lambda_1 + \left[u^{\lambda_3} - (1-u)^{\lambda_4}\right]/\lambda_2, \qquad 0 \le u \le 1, \tag{2.21}$$

and exhibits a variety of skewness behaviour, but all members with $\lambda_4 > (<)\lambda_3$ are at least tail-skew to the left (right) [MacGillivray (1982)].

DEFINITION 2.8. $F \le_2^{m,c} G$ if Definition 2.7 holds for the proportion interval $[0, 1 - 2\alpha]$ for some α.

$F \le_2^{m,t} G$ if Definition 2.7 holds for the proportion interval $[1 - 2\alpha, 1]$ for some α.

When $\gamma_u(F)$ has only one change of sign for $u \in [0, \frac{1}{2}]$, central skewness and tail skewness are all that is required to describe the skewness properties of F. The relative importance of the central and tail skewness depends on the value of u at which the change-over occurs as well as the context of consideration. Figure 1 presents the hierarchy of skewness orderings.

2.4. *Measures of asymmetry.* Doksum (1975) defines an index of asymmetry for the central $100(1 - 2\alpha)\%$ of the distribution by

$$\left[\sup_{\alpha \le u \le 1/2} \theta_F(F^{-1}(u)) - \inf_{\alpha \le u \le 1/2} \theta_F(F^{-1}(u))\right] \Big/ \sigma_F, \tag{2.22}$$

where σ_F is the standard deviation or some other measure of spread, and where $\theta_F(F^{-1}(u)) = \frac{1}{2}[F^{-1}(u) + F^{-1}(1-u)]$.

From Section 2.2, it follows that it is more appropriate to consider either central-density or inter-quantile scaling, thus using instead of (2.22), either

$$\sup_{\alpha \le u \le 1/2} \nu_u(F) - \inf_{\alpha \le u \le 1/2} \nu_u(F) \tag{2.23}$$

or

$$\sup_{\alpha \le u \le 1/2} \gamma_u(F) - \inf_{\alpha \le u \le 1/2} \gamma_u(F). \tag{2.24}$$

The class of distributions that are skew to the right according to $\le_2^m$ is now denoted by $\mathscr{F}_R$ and the class of distributions skew to the left according to $\le_2^m$ by $\mathscr{F}_L$.

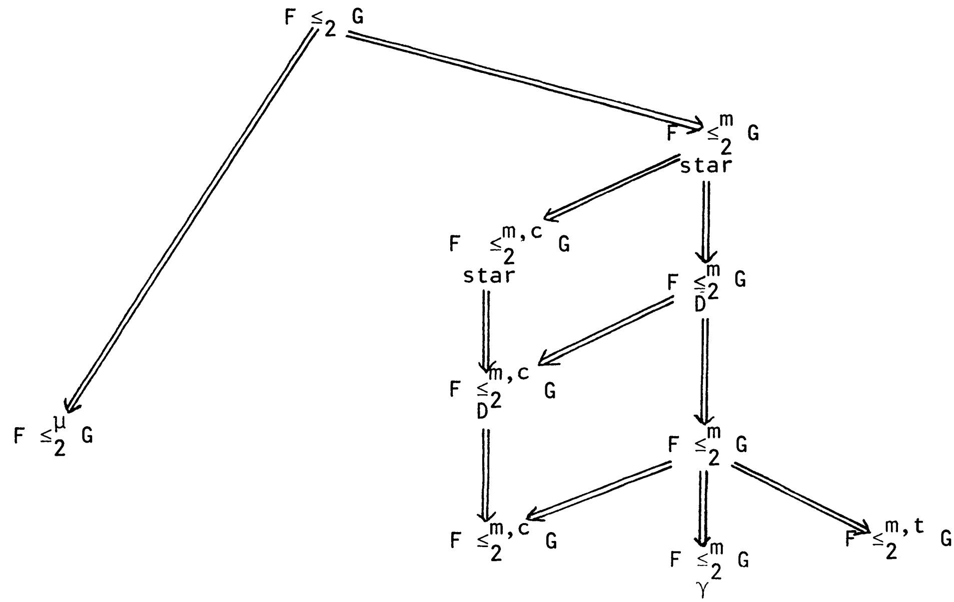

FIG. 1. *Partial orderings of F and G with respect to skewness.*

Within $\mathscr{F}_R$, $\inf_{\alpha \le u \le 1/2} \nu_u(F) = 0 = \inf_{\alpha \le u \le 1/2} \gamma_u(F)$, and hence (2.23) and (2.24) preserve $\le_2^m$ for any $\alpha \le \frac{1}{2}$ (and hence all the preceding orderings in the hierarchy of Theorem 2.2). For the class $\mathscr{F}_R$, Doksum calls the index of asymmetry an index of skewness for the central $100(1 - 2\alpha)\%$. Within $\mathscr{F}_L$, $\sup_{\alpha \le u \le 1/2} \nu_u(F) = 0 = \sup_{\alpha \le u \le 1/2} \gamma_u(F)$, so that $-(2.23)$ and $-(2.24)$ preserve $\le_2^m$ for any $\alpha \le \frac{1}{2}$. Of course, these comments may be generalised to the wider classes of distributions that are centrally skew to the right or left for some fixed α.

Hence for distributions that are skew to the right or to the left for their whole domain, the proposed measures of asymmetry (2.23) and (2.24) are measures of the amount of skewness regardless of whether it is to the right or left. This seems to be a reasonable interpretation of the concept of asymmetry and it can be extended to comparing distributions that do not belong to either $\mathscr{F}_R$ or $\mathscr{F}_L$. For example, if $F \in \mathscr{F}_R$ and $G \in \mathscr{F}_L$ satisfy $F \le_2^m \overline{G}$, then, $\sup|\nu_u(F)| \le \sup|\nu_u(G)|$ and $\sup|\gamma_u(F)| \le \sup|\gamma_u(G)|$. G is more skew to the left than F is to the right, and it is reasonable to say that G is more asymmetric than F. Continuing this concept to cover distributions whose skewness may change direction within their domain, leads to a definition of an ordering between distributions with respect to asymmetry.

DEFINITION 2.9. F is skew to the right (left) in the proportion interval $(1 - 2\alpha_2, 1 - 2\alpha_1)$ if F is as least as skew to the right (left) with respect to the median as $\overline{F}$ in this interval.

Definition 2.10. Suppose that $[0,1]$ is partitioned into intervals so that in each interval F is skew to the right or left and G is skew to the right or left.

Let

$$F_1 = \begin{cases} F & \text{where } F \text{ is skew to the right,} \\ \bar{F} & \text{where } F \text{ is skew to the left.} \end{cases}$$

Similarly for G_1.

Then G is at least as asymmetric as F, $F \leq_A G$, if, in each interval of the partition, G_1 is at least as skew to the right as F_1.

Similarly G is at least as asymmetric as F in the central $100(1-2\alpha)\%$ of the distribution, $F \leq_{A,\alpha} G$, if instead of $[0,1]$, only the proportion interval $[0, 1-2\alpha]$ is considered.

Theorem 2.4. *$|\nu_u(\cdot)|$, $|\gamma_u(\cdot)|$, and their suprema over $u \in (0, \frac{1}{2})$ preserve $\leq_A$. Similarly $|\nu_u(\cdot)|$, $|\gamma_u(\cdot)|$, and their suprema over $u \in (\alpha, \frac{1}{2})$ preserve $\leq_{A,\alpha}$.*

Proof. Follows immediately. □

Hence $\sup_{\alpha \leq u \leq 1/2}|\nu_u(\cdot)|$ and $\sup_{\alpha \leq u \leq 1/2}|\gamma_u(\cdot)|$ are suggested as measures of asymmetry instead of (2.23) and (2.24), coinciding with them for distributions belonging to $\mathscr{F}_R$ or $\mathscr{F}_L$.

The quantities $\sup_u|\nu_u(\cdot)|$ and $\sup_u|\gamma_u(\cdot)|$ also have further justification as measures of asymmetry. Of all the distributions symmetric about m_F, $H(x)$, defined by $H^{-1}(u) = \frac{1}{2}[F^{-1}(1-u) - F^{-1}(u)] + m_F$, best approximates $F(x)$ in the sense that $\sup_u|F^{-1}(u) - H^{-1}(u)|$ is a minimum [Doksum (1975)]. Since $|F^{-1}(u) - H^{-1}(u)| = \frac{1}{2}|F^{-1}(u) + F^{-1}(1-u) - 2m_F|$, this is the distance the uth quantile has to move to become the uth quantile of the "closest" symmetric distribution. $H(x)$ is also the only distribution symmetric about m_F with the same inter-quantile scale measure, namely $F^{-1}(1-u) - F^{-1}(u)$ for all u. Thus, as well as preserving a reasonable ordering with respect to asymmetry, $\sup|\nu_u(F)|$ and $\sup|\gamma_u(F)|$ give a minimum standardised distance between F and a distribution symmetric about m_F.

Figure 2 shows $\gamma_u(F)$ for some members of the Ramberg et al. (1979) asymmetric Tukey lambda family with $\lambda_3 = 2\lambda_4$. Let F correspond to $\lambda_4 = 2.5$ and G to $\lambda_4 = 3$. Then for the interval $0.1 \leq u \leq 0.5$, $\sup|\gamma_u(F)| \geq \sup|\gamma_u(G)|$ but $\sup\gamma_u(F) - \inf\gamma_u(F) \leq \sup\gamma_u(G) - \inf\gamma_u(G)$. Essentially a change in the *direction* of the skewness within the domain of a distribution does not necessarily increase the overall asymmetry—it may help to decrease it. The graph for $\lambda_4 = 0.2$ is also included to show that the family, at least for $\lambda_3 = 2\lambda_4$, is not ordered according to any of the orderings of Theorem 2.2.

3. Skewness properties of some distribution families. The relationship of F and $\bar{F}$ is examined in this section for some important families of distributions, also illustrating some points of interest from Section 2. All the results have been obtained analytically but because the study of sign changes can sometimes be laborious, no proofs are given here. Functions such as $\bar{F}^{-1}(F(x))$,

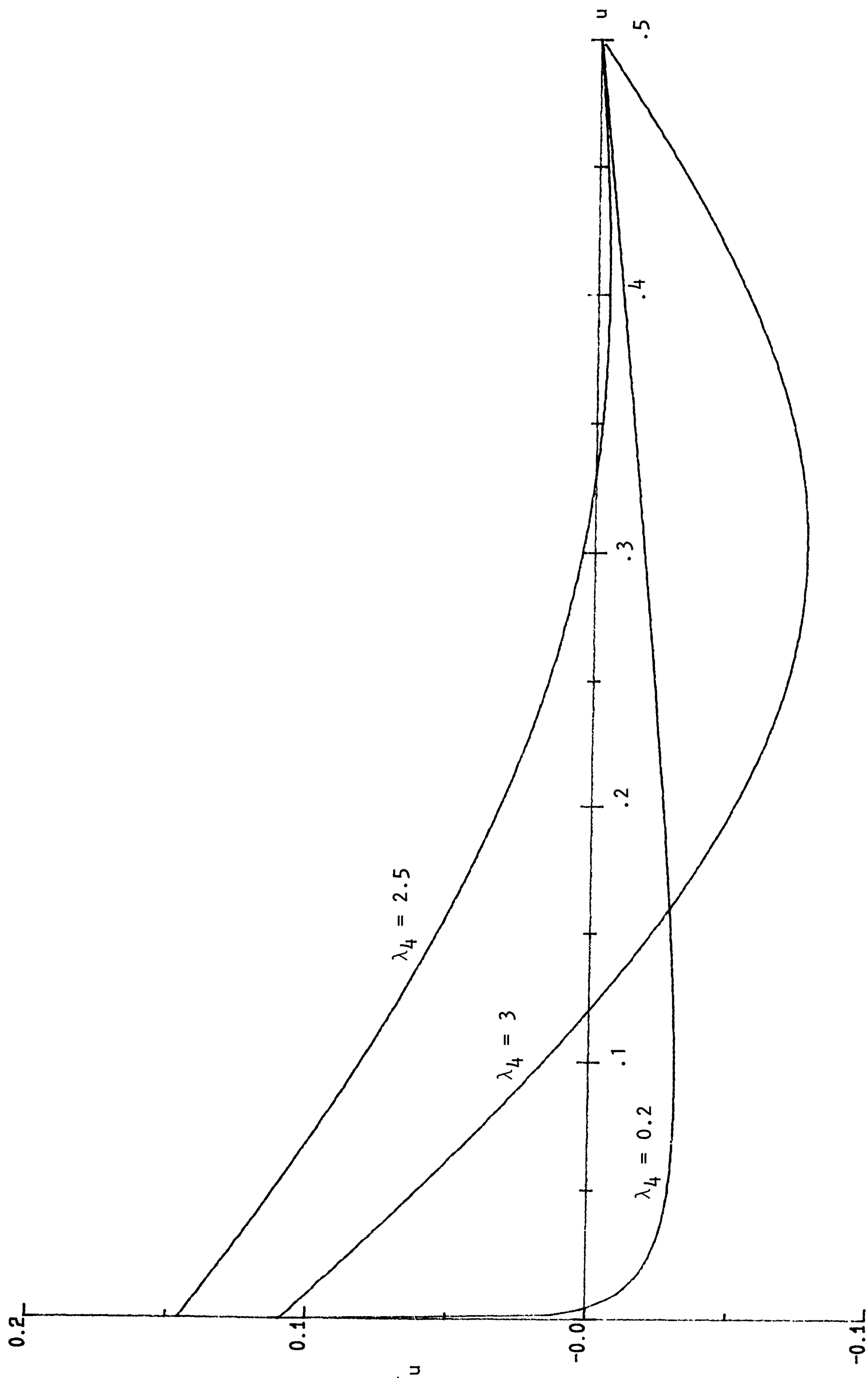

FIG. 2. γ_u *for members of an asymmetric Tukey lambda family.*

$F^{-1}(1-u) + F^{-1}(u) - 2m_F$, or $F(\zeta + x) + F(\zeta - ax) - 1$ may be considered, and sometimes may all need to be considered.

3.1. *Weibull distribution.* This has $F(x) = 1 - e^{-\theta x^c}$, $x > 0$; $c > 0$, $\theta > 0$. Hence $F^{-1}(u) = [\theta^{-1}\log(1-u)^{-1}]^{1/c}$, and it can easily be shown that $F_1 <_2 F_2$ for $c_1 > c_2$. That is, as c increases, the distribution becomes less skew to the right. Now it is known that μ_3 changes sign from positive to negative values at $c \approx 3.6$ [Johnson and Kotz (1970), page 253] and examination of $m - M$ shows that this changes from positive to negative values at $c = (1 - \log 2)^{-1} \approx 3.26$. Since $I_F = (0, \infty)$, the distribution is always at least tail-skew to the right, but the changes in μ_3 and $m - M$ show that $F \notin \mathscr{F}_R$. Detailed examination shows that for $c \leq 1/(1 - \log 2)$, F is strongly skew to the right, and for $c > 1/(1 - \log 2)$, it is centrally skew to the left and tail-skew to the right. It can also be shown that for $c > 1/(1 - \log 2)$, $F^{-1}(1-u) + F^{-1}(u) - 2m_F$ not only has one change of sign but also only one turning point.

The strong result may be shown by considering the second derivative of $\bar{F}^{-1}(F(x))$ at $x = F^{-1}(u)$. The weaker results may be shown by considering derivatives of $F^{-1}(u) + F^{-1}(1-u) - 2m_F$, and, by considering $\log f(\zeta + x) - \log f(\zeta - x)$, it may also be shown that $S^-[F^{-1}(u) + F^{-1}(1-u) - 2m_F] = 1$, from positive to negative values for $u \in (0, \frac{1}{2})$.

Figures 3 and 4 show that for most of I_F, the family virtually behaves as a strongly ordered family that includes a symmetric distribution. The tail skewness has practically no influence except on global measures such as μ_3 and $\mu - m$ for c between 3 and 4.

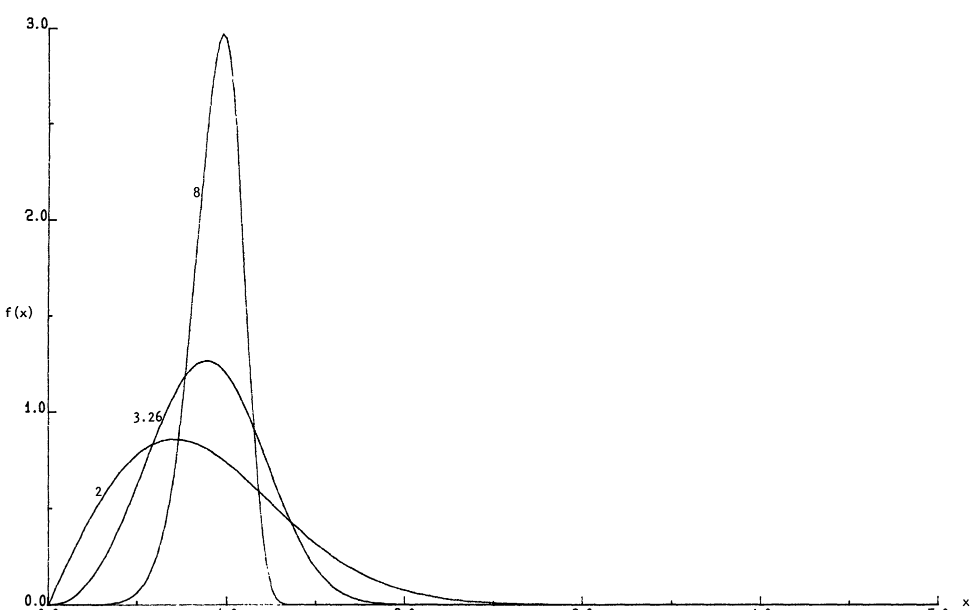

FIG. 3. *Weibull density functions for $c = 2, 3.26, 8$.*

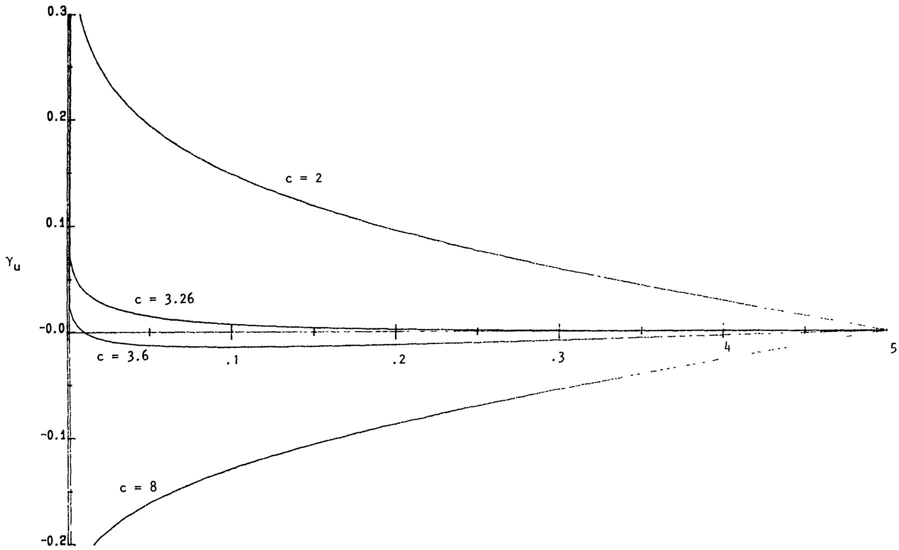

FIG. 4. γ_u *for members of the Weibull family.*

3.2. *Johnson system* [Johnson (1949)]. For this system $F(x) = \Phi(b(x))$ where Φ is the distribution function of the standard normal and $b(x)$ is an increasing function. For skewness properties, $\Phi(z)$ may in fact be any distribution symmetric about zero; the properties are determined by the transformation $b(x)$. Johnson identified three main systems called S_L, S_U, and S_B, corresponding to three general forms of $b(x)$.

For the S_L system, $b(x) = \delta \log x + \gamma$, $x > 0$. For $\delta > 0$, this is concave and F is strongly skew to the right; conversely for $\delta < 0$ and similarly for any concave or convex $b(x)$.

For the S_U system $b(x) = \delta \sinh^{-1}x + \gamma$, $-\infty < x < \infty$, and for the S_B system $b(x) = \delta \log[x/(1-x)] + \gamma$, $0 < x < 1$. Without loss of generality, consider $\delta > 0$. For $\gamma < 0$, S_U has $\bar{F} \underset{D}{\overset{m}{\leq}}{}_{2} F$ and $\bar{F} \overset{\mu}{\leq}{}_{2} F$ while S_B has $F \underset{D}{\overset{m}{\leq}}{}_{2} \bar{F}$ and $F \overset{\mu}{\leq}{}_{2} \bar{F}$; conversely for $\gamma > 0$. For these two systems, there is an interesting point when considering $F(\zeta + x) + F(\zeta - ax) - 1$. In both cases this has no more than two changes of sign for any ζ and a, but when there are two changes of sign there sequence of signs depends on ζ and a so that F and $\bar{F}$ cannot be ordered by $\leq_2$.

The Pearson system also has the property that

$$S^{-}(F(\zeta + x) + F(\zeta - ax) - 1) \leq 2,$$

but again, for individual distributions it is necessary to check that the sequence of sign changes in the case of equality does not change with ζ and a.

4. Conclusion. This paper has aimed to bring together into one structure the variety of views of skewness that have been previously considered, filling in any gaps. The structure so obtained is intended to provide a background for reference purposes as obviously not all the orderings and measures will be required for any one situation.

Yule (1911, page 162) described the quantile measure $\gamma_{0.25}(\cdot)$ as a "rather rough-and-ready" measure of skewness. Although the relative importance of the different orderings and measures depends on circumstances, and it is unlikely that any one could be described as most important, it appears that the general quantile measures $\gamma_u(\cdot)$ play a valuable role in discussing both skewness and asymmetry.

A particular point that has emerged in the course of the paper is that describing the skewness of individual distributions is not only a special case of comparing different distributions, but also may be a necessary inclusion in a full description of the comparative skewness properties of a family of distributions.

Acknowledgments. The author would like to thank Mr. John Zornig for his assistance in preparing the diagrams, and the referees for their helpful comments.

REFERENCES

BALANDA, K. P. and MACGILLIVRAY, H. L. (1986). *Kurtosis and Spread.* Submitted for publication.

BARLOW, R. E. and PROSCHAN, F. (1966). Inequalities for linear combinations of order statistics from restricted families. *Ann. Math. Statist.* **37** 1574–1592.

BICKEL, P. J. and LEHMANN, E. L. (1975). Descriptive statistics for non-parametric models. I. Introduction. II. Location. *Ann Statist.* **3** 1038–1069.

BICKEL, P. J. and LEHMANN, E. L. (1976). Descriptive statistics for non-parametric models. III. Dispersion. *Ann. Statist.* **4** 1139–1158.

BOWLEY, A. L. (1901). *Elements of Statistics.* Staples Press Ltd., London. (6th ed., 1937).

DAVID, F. N. and JOHNSON, N. L. (1954). A test for skewness with ordered variables. *Ann. Eugenics Lond.* **18** 351–353.

DAVID, F. N. and JOHNSON, N. L. (1956). Some tests of significance with ordered variables. *J. Roy Statist. Soc. Ser. B* **18** 1–20.

DAVID, H. A. and GROENEVELD, R. A. (1982). Measures of local variation in a distribution: Expected length of spacings and variances of order statistics. *Biometrika* **69** 227–232.

DOKSUM, K. A. (1969). Starshaped transformations and the power of rank tests. *Ann. Math. Statist.* **40**, 1167–1176.

DOKSUM, K. A. (1975). Measures of location and asymmetry. *Scand. J. Statist.* **2** 11–22.

DOODSON, A. T. (1917). Relation of the mode, median and mean in frequency curves. *Biometrika* **11** 425–429.

FECHNER, G. TH. (1897). *Kollektivmasslehre.* Engleman, Leipzig.

GARVER, R. (1932). Concerning the limits of a measure of skewness. *Ann. Math. Statist.* **3** 358–360.

GROENEVELD, R. A. and MEEDEN, G. (1977). The mode, median and mean inequaltiy. *Amer. Statist.* **31** 120–121.

GROENEVELD, R. A. and MEEDEN, G. (1984). Measuring skewness and kurtosis. *The Statistician* **33** 391–399.

HALDANE, J. B. S. (1942). The mode and median of a nearly normal distribution with given cumulants. *Biometrika* **32** 294–299.

HINKLEY, D. V. (1975). On power transformations to symmetry. *Biometrika* **62** 101–111.

HOGG, R. V. (1974). Adaptive robust procedures: A partial review and some suggestions for future applications and theory. *J. Amer. Statist. Assoc.* **69** 909–923.

HOGG, R. V., FISHER, D. M. and RANDLES, N. H. (1975). A two-sample adaptive distribution-free test. *J. Amer. Statist. Assoc.* **70** 656–661.

HOTELLING, H. and SOLOMONS, L. M. (1932). The limits of a measures of skewness. *Ann. Math. Statist.* **3** 141–142.

JOHNSON, N. L. (1949). Systems of frequency curves generated by methods of translation. *Biometrika* **36** 149–176.

JOHNSON, N. L. and KOTZ, S. (1970). *Continuous Univariate Distributions* **1**. Houghton-Mifflin, Boston.

KARLIN, S. (1968). *Total Positivity* **1**. Stanford Univ. Press, Stanford, Calif.

LAWRENCE, M. J. (1975). Inequalities of s-ordered distributions. *Ann. Statist.* **3** 413–428.

MACGILLIVRAY, H. L. (1979). Moment inequalities with applications to particle size distributions. Unpublished Ph.D. thesis, Univ. Queensland.

MACGILLIVRAY, H. L. (1981). The mean, median, mode inequality and skewness for a class of densities. *Austral. J. Statist.* **23** 247–250.

MACGILLIVRAY, H. L. (1982). Skewness properties of asymmetric forms of Tukey lambda distributions. *Comm. Statist. A—Theory Methods* **11** 2239–2248.

MACGILLIVRAY, H. L. (1985). A crossing theorem for distributions and their moments. *Bull. Austral. Math. Soc.* **31** 413–419.

MAJINDAR, K. N. (1962). Improved bounds on a measure of skewness. *Ann. Math. Statist.* **33** 1192–1194.

MANN, H. B. and WHITNEY, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *Ann. Math. Statist.* **18** 50–60.

OJA, H. (1981). On location, scale, skewness and kurtosis of univariate distributions. *Scand. J. Statist.* **8** 154–168.

ORD, J. K. (1968). The discrete student's t distribution. *Ann. Math. Statist.* **39** 1513–1516.

PEARSON, K. (1895). Contributions to the mathematical theory of evolution II. Skew variation in homogeneous material. *Philos. Trans. Roy. Soc. London Ser. A* **186** 343.

RAMBERG, J. S., TADIKAMALLA, P. R., DUDEWICZ, E. J. and MYKYTKA, E. F. (1979). A probability distribution and its uses in fitting data. *Technometrics* **21** 201–214.

ROGERS, W. H. and TUKEY, J. W. (1972). Understanding some long-tailed symmetrical distributions. *Statist. Neerlandica* **26** 211–226.

ROSENBERGER, J. L. and GASKO, M. (1983). Comparing location estimators: Trimmed means, medians and trimean. In *Understanding Robust and Exploratory Data Analysis* (D. C. Hoaglin, F. Mosteller and J. W. Tukey, eds.) 297–338. Wiley, New York.

RUNNENBURG, J. TH. (1978). Mean, median, mode. *Statist. Neerlandica* **32** 73–79.

TIMERDING, H. E. (1915). *Die Analyse der Zufalls.* Braunschweig. (In Leiden University Library.)

VAN ZWET, W. R. (1964). Convex transformations of random variables. *Math. Centrum.* Amsterdam.

VAN ZWET, W. R. (1979). Mean, median, mode II. *Statist. Neerlandica* **33** 1–5.

YULE, G. U. (1911). *Introduction to the Theory of Statistics.* Griffin, London. (13th ed., 1944).

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF QUEENSLAND
ST. LUCIA
QUEENSLAND 4067
AUSTRALIA