Stochastic Inequalities
IMS Lecture Notes – Monograph Series
Volume 22 (1993)

SKEWNESS AND KURTOSIS ORDERINGS: AN INTRODUCTION

By BARRY C. ARNOLD and RICHARD A. GROENEVELD

University of California, Riverside and Iowa State University

Competing skewness orderings are surveyed. It is argued that those based on natural skewness functionals are preferable to those related to convex orderings. Analogous kurtosis orderings are also discussed. Here the role of convex and Lorenz orderings appears more natural.

1. Introduction

What is skewness? An analogous question regarding inequality led Dalton eventually down the majorization path via the enunciation of clearly agreed upon inequality reducing transformations. Can a similar analysis be performed with skewness? In a sense the answer is easy; skewness is asymmetry, plain and simple. It is of course easy to recognize symmetric distributions but not so easy to decide whether one non symmetric distribution is more unsymmetric than another. Robin Hood (i.e. rich to poor) transfers are at the heart of the accepted inequality orderings. It is natural to search for analogous basic operations which will increase skewness. The present paper surveys suggested skewness orderings (although not in the detail provided by MacGillivray (1986)) but puts its major focus on promoting a particular group of skewness orderings. Clear parallels may be discerned between some of these orderings and the Lorenz inequality ordering generated by Robin Hood operations.

What is Kurtosis? This is a bit harder. To quote our dictionary (Webster's of course) it is the state or quality of peakedness or flatness of the graphic representation of a statistical distribution. Again a plethora of competing orderings have been proposed (see Balanda and MacGillivray (1988) for a recent survey). Again we champion a particular ordering related to Lorenz ordering.

AMS 1991 *subject classifications.* Primary 62E10, 60E99.

Key words and phrases. Asymmetry, peakedness, Lorenz order, star order, convex order, skewness functionals.

2. A Budget of Skewness Orderings

To simplify discussion, we restrict attention to the class $\mathcal{L}$ of distributions with median 0 and finite first absolute moment. The moment condition is not crucial for some of the orderings but several orderings will only be well-defined when first moments exist and it is convenient to restrict our focus to such distributions.

For any distribution $F \epsilon \mathcal{L}$, we define its quantile (or inverse distribution) function F^{-1} by

$$F^{-1}(u) = sup\{x : F(x) \leq u\}, \quad 0 < u < 1. \tag{2.1}$$

The zero median condition is equivalent to

$$F^{-1}(\frac{1}{2}) = 0 \tag{2.2}$$

and the first absolute moment is expressible as

$$E(|X|) = \int_0^{\frac{1}{2}} [F^{-1}(1-u) - F^{-1}(u)] du. \tag{2.3}$$

The class of median zero distributions with finite first moment can be identified conveniently with the class of all non-decreasing functions defined on $(0,1)$ satisfying (2.2) and (2.3). A symmetric distribution can be characterized by the requirement that

$$F^{-1}(1-u) + F^{-1}(u) = 0, \quad \forall u \epsilon (0, \frac{1}{2}). \tag{2.4}$$

Skewness corresponds to the violation of condition (2.4). Positive values of (2.4) for some, most or all values of u will be associated with positive skewness or skewness to the right. Negative skewness is associated with negative values of (2.4). We will define skewness orderings denoted by a symbol $\leq$ with a variety of subscripts and will in a cavalier fashion write them in terms of random variables or distribution functions i.e. $X \leq Y \Longleftrightarrow F_X \leq F_Y$.

It is generally conceded that measures of skewness and related skewness orderings should be scale invariant, i.e. for any *positive* constant c, X and cX exhibit the same degree of skewness. Accepting this viewpoint it is defensible to divide any random variable by its first absolute moment and effectively focus on skewness orderings defined on the class of random variables with

$$F_X^{-1}(\frac{1}{2}) = 0 \tag{2.5}$$

and

$$E|X| = 1. \tag{2.6}$$

Denote the class of distribution functions satisfying (2.5) and (2.6) by $\mathcal{L}_0$.

A strong or uniform skewness ordering on $\mathcal{L}_0$ is provided by

$$\begin{aligned} X \leq_s Y \text{ iff } F_X &\leq_s F_Y \text{ iff } F_X^{-1}(1-u) + F_X^{-1}(u) \\ &\leq F_Y^{-1}(1-u) + F_Y^{-1}(u) \ \forall u \epsilon (0, \tfrac{1}{2}). \end{aligned} \tag{2.7}$$

A natural skewness functional is implicit in (2.7), namely

$$s_X(u) = F_X^{-1}(1-u) + F_X^{-1}(u), \quad 0 < u < \frac{1}{2}. \tag{2.8}$$

Rather than insist on a uniform domination of $s_X(u)$ by $s_Y(u)$, we might be willing to be forgiving of minor local violations. Two alternative skewness functionals are proposed namely

$$\begin{aligned} \nu_X(u) &= \int_0^u s_X(v)dv \\ &= \int_0^u [F_X^{-1}(1-v) + F_X^{-1}(v)]dv, \quad 0 < u < \frac{1}{2} \end{aligned} \tag{2.9}$$

and

$$\lambda_X(u) = \int_0^u [F_X^{-1}(\frac{1}{2}+v) + F_X^{-1}(\frac{1}{2}-v)]dv, \quad 0 < u < \frac{1}{2}. \tag{2.10}$$

Using these functionals we define analogous skewness orderings

$$X \leq_\nu Y \text{ iff } \nu_X(u) \leq \nu_Y(u), \ \forall u \epsilon (0, \frac{1}{2}) \tag{2.11}$$

and

$$X \leq_\lambda Y \text{ iff } \lambda_X(u) \leq \lambda_Y(u), \ \forall u \epsilon (0, \frac{1}{2}). \tag{2.12}$$

None of the above three skewness orderings seems to have been given close attention in the literature. Though certainly the uniform ordering has been implicitly considered.

Much of the literature on skewness orderings focusses on a variety of weakenings of Van Zwet's (1964) convex ordering. Our presentation is a somewhat simplified version of MacGillivray's (1986) section 2. Simplification is possible because of our focus on $\mathcal{L}_0$, the class of distributions with median 0 and unit first absolute moment.

The strong convex ordering of Van Zwet is defined by

$$\begin{aligned} X \leq_1 Y \quad &\text{iff} \quad F_Y^{-1}(F_X(x)) \text{ is convex on the support of } F \\ &\text{iff} \quad F_X^{-1}(u) \text{ and } [\text{a } F_Y^{-1}(u) + b] \text{ cross at most twice on} \\ &\qquad (0,1) \text{ for any } a > 0, b \epsilon \mathbb{R} \text{ with sign sequence of} \\ &\qquad F_X^{-1}(u) - aF_Y^{-1}(u) - b \text{ being } -,+,-. \end{aligned} \tag{2.13}$$

A star ordering (related to Oja's (1981) ordering) is defined by

$$X \leq_2 Y \text{ iff } \frac{F_Y^{-1}(u)}{F_X^{-1}(u)} \uparrow \text{ on } (0,1) - \{\frac{1}{2}\} \tag{2.14}$$

where the symbol $\uparrow$ is to be read here and henceforth as non-decreasing. Progressively further weakening yields the following orders.

(2.15) $X \leq_3 Y$ iff

$$F_Y^{-1}(u) - \frac{f_X(0)}{f_Y(0)} F_X^{-1}(u) \text{ is } \downarrow \text{ on } (0, \frac{1}{2}) \text{ and } \uparrow \text{ on } (\frac{1}{2}, 1)$$

$$X \leq_4 Y \text{ iff } F_Y^{-1}(u) f_Y(0) \leq F_X^{-1}(u) f_X(0) \text{ on } (0,1) - \{\frac{1}{2}\} \tag{2.16}$$

and

$$X \leq_5 Y \text{ iff } \frac{F_X^{-1}(1-u)}{F_X^{-1}(u)} \geq \frac{F_Y^{-1}(1-u)}{F_Y^{-1}(u)}, \quad 0 < u < \frac{1}{2}. \tag{2.17}$$

In the above definitions f_X and f_Y are the densities corresponding to F_X and F_Y.

An ordering based on stochastic ordering of the positive and negative parts of X and Y has been proposed:

$$X \leq_6 Y \text{ iff } X^+ \leq_{st} Y^+ \text{ and } Y^- \leq_{st} X^- \tag{2.18}$$

Instead of stochastic ordering in (2.18) we might invoke Lorenz ordering. Thus we have

$$X \leq_7 Y \text{ iff } X^+ \leq_L Y^+ \text{ and } Y^- \leq_L X^-. \tag{2.19}$$

A convenient reference for the definition of $\leq_L$ is Arnold (1987). Second and higher order stochastic dominance could of course be used as a basis of a skewness order definition but we will not pursue that possibility.

Finally we mention the David and Johnson (1956) skewness functional related to but distinct from $s_X(u)$ defined in (2.8). It takes the form

$$\tilde{s}_X(u) = \frac{F_X^{-1}(1-u) + F_X^{-1}(u)}{F_X^{-1}(1-u) - F_X^{-1}(u)}, \quad 0 < u < \frac{1}{2} \tag{2.20}$$

Analogs to (2.9) and (2.10) are available

$$\tilde{\nu}_X(u) = \int_0^u \left[\frac{F_X^{-1}(1-v) + F_X^{-1}(v)}{F_X^{-1}(1-v) - F_X^{-1}(v)} \right] dv, \quad 0 < u < \frac{1}{2} \tag{2.21}$$

and

$$\tilde{\lambda}_X(u) = \int_0^u \frac{F_X^{-1}(\frac{1}{2}+v) + F_X^{-1}(\frac{1}{2}-v)}{F_X^{-1}(\frac{1}{2}+v) - F_X^{-1}(\frac{1}{2}-v)} dv, \quad 0 < u < \frac{1}{2}. \tag{2.22}$$

An advantage of the David-Johnson functionals is that they are well defined without any assumption regarding the existence of $E|X|$. All three functions are uniformly bounded in absolute value by 1. Skewness orderings $\leq_{\tilde{s}}, \leq_{\tilde{\nu}}$ and $\leq_{\tilde{\lambda}}$ are defined in the natural way using these functionals. Note that $\leq_{\tilde{s}}$ is equivalent to $\leq_5$.

We remark finally without further comment on the possibility of introducing a non-negative weight function $\psi(v)$ inside the integral in definitions (2.9), (2.10), (2.21) and (2.22).

3. Inter–Relationships Among the Orderings

Thirteen skewness orderings were introduced in Section 2. How are they related? As MacGillivray (1986) noted, convex ordering is very strong and implies many reasonable definitions of skewness. Specifically from her paper we have the relations

$$\leq_1 \Longrightarrow \leq_2 \Longrightarrow \leq_3 \Longrightarrow \leq_4 \Longrightarrow \leq_5 . \tag{3.1}$$

It is not difficult to verify that we also have

$$\leq_6 \Longrightarrow \leq_s \begin{cases} \Longrightarrow \leq_\nu \\ \Longrightarrow \leq_\lambda \end{cases} \tag{3.2}$$

$$\leq_5 \Longleftrightarrow \leq_{\tilde{s}} \begin{cases} \Longrightarrow \leq_{\tilde{\nu}} \\ \Longrightarrow \leq_{\tilde{\lambda}} \end{cases} \tag{3.3}$$

and

$$\leq_2 \Longrightarrow \leq_7 \tag{3.4}$$

The key observations justifying (3.4) are that $X \leq_2 Y$ implies $X^+ \leq^* Y^+$ and $Y^- \leq^* X^-$ and star-ordering implies Lorenz ordering (see for example Arnold (1987, p. 78)).

The convex ordering $\leq_1$ is much stronger than necessary. It is our contention that the candidate orderings most worthy of consideration are $\leq_5$ (equivalently $\leq_{\tilde{s}}$), $\leq_6$ and $\leq_s$. Additionally it is felt that the concept of skewness does not necessarily involve comparison of positive and negative parts of random variables. On this basis $\leq_6$ is not as appealing. We are left with $\leq_s$ and $\leq_{\tilde{s}}$ together with the more forgiving integrated versions provided by $\leq_\nu, \leq_{\tilde{\nu}}, \leq_\lambda$ and $\leq_{\tilde{\lambda}}$.

At this juncture, if we were forced to select a single skewness ordering to recommend, it might well be $\leq_\nu$. Note that $\nu_X(u) \leq \nu_Y(u)\ \forall\ u \epsilon (0, \frac{1}{2})$ is equivalent to

$$E(X|X \leq F_X^{-1}(u)) + E(X|X \geq F_X^{-1}(1-u))$$

$$\leq E(Y|Y \leq F_Y^{-1}(u)) + E(Y|Y \geq F_Y^{-1}(1-u))$$

i.e. the average of the right and left tails of Y exceeds the average of the right and left tails of X (at each percentile). This is not an unreasonable definition of Y being more skewed to the right than X.

Details regarding the relationships (3.2)–(3.4) together with counterexamples to any other inter-relationships among the orderings will appear in a more extensive report. The example in the following section highlights the potential clash between $\leq_1$ and $\leq_s$.

4. Examples of the Skewness Functionals

Consider a random variable X whose distribution is given by

$$F_X(x) = 1 - (x + 2^b)^{-(1/b)}, \quad x > 1 - 2^b \tag{4.1}$$

where $b \epsilon (0,1)$. The resulting inverse distribution function is

$$F_X^{-1}(u) = (1-u)^{-b} - 2^b, \quad 0 < u < 1$$

and hence

$$\begin{aligned} E(|X|) &= \int_0^{\frac{1}{2}} [F_X^{-1}(1-u) - F_X^{-1}(u)]du \\ &= (2^b - 1)/(1-b) \end{aligned}$$

so that our skewness functionals assume the form

$$s_X(u) = \frac{(1-u)^{-b} + u^{-b} - 2^{b+1}}{(2^b-1)/(1-b)},$$

$$\nu_X(u) = \frac{1 + u^{1-b} - (1-u)^{1-b} - (1-b)2^{b+1}u}{2^b - 1},$$

and

$$\lambda_X(u) = \frac{(\frac{1}{2}+u)^{1-b} - (\frac{1}{2}-u)^{1-b} - (1-b)2^{b+1}u}{2^b-1}.$$

The functional ν_X is monotonically increasing in b. The larger the values of b, the more positively skewed is the resulting distribution as measured by ν_X.

Note that if X_1 and X_2 have distribution (4.1) with corresponding parameters $b_1 < b_2$, then $X_{b_1} \leq_1 X_{b_2}$ and as observed above $X_{b_1} \leq_\nu X_{b_2}$. However numerical calculations indicate that, for a given b_1, there exists $b_2 > b_1$ such that $X_{b_1} \not\leq_s X_{b_2}$ and $X_{b_1} \not\leq_\lambda X_{b_2}$.

5. Skewness Accentuating Transformations

We may reasonably seek to characterize all transformations $g : \mathbb{R} \to \mathbb{R}$ which have the property that for any $X \epsilon \mathcal{L}_0$ we have X less skew than $g(X)$. It is evidently true that if we choose a function g such that $g(0) = 0$ and both $g(x)$ and $g(x)/x \uparrow$ on $\mathbb{R} - \{0\}$ then $X \leq_1 g(X)$ for any $X \epsilon \mathcal{L}_0$. Thus transformations of this kind accentuate skewness using the strongest (Van Zwet) skewness ordering. They thus accentuate skewness using any of the other orderings implied by the Van Zwet orderings. These transformations however do not necessarily accentuate skewness as measured by $\leq_6$.

6. Kurtosis

An analogous variety of kurtosis orderings exist. Most restrict attention to symmetric distributions. A reason for this is the difficulty of interpretation of the concept of kurtosis in the absence of symmetry. Setting aside such niceties for the moment, it is possible to provide kurtosis orderings analogous to several of the skewness orderings described in this paper. Some of these reduce to already known kurtosis orderings if symmetry is imposed. As in our skewness discussion we standardize all variables to have median 0 and first absolute moment equal to 1; i.e. we restrict attention to $\mathcal{L}_0$. To distinguish our kurtosis ordering from the corresponding parallel skewness orderings we place a superscript k above the inequality sign. Thus $\leq_2^k$ will be a kurtosis ordering analogous to the skewness ordering $\leq_2$. Here is the list (as usual $X = X^+ - X^-$ where $X^+ \geq 0$ and $X^- \geq 0$).

$$X \leq_1^k Y \quad \text{iff} \quad F_{Y^+}^{-1}(F_{X^+}(x)) \text{ is convex on the support of } X^+ \text{ and } F_{Y^-}^{-1}(F_{X^-}(x)) \text{ is convex on the support of } X^-. \tag{6.1}$$

$$X \leq_2^k Y \quad \text{iff} \quad F_{Y^+}^{-1}(u)/F_{X^+}^{-1}(u) \uparrow \text{ on } (0,1) \text{ and } F_{Y^-}^{-1}(u)/F_{X^-}^{-1}(u) \uparrow \text{ on } (0,1) \tag{6.2}$$

$$X \leq_7^k Y \quad \text{iff} \quad X^+ \leq_L Y^+ \text{ and } X^- \leq_L Y^- \tag{6.3}$$

$$X \leq_\lambda^k Y \quad \text{iff} \quad \int_0^u [F_X^{-1}(\frac{1}{2}+v) - F_X^{-1}(\frac{1}{2}-v)]dv \geq \int_0^u [F_Y^{-1}(\frac{1}{2}+v) - F_Y^{-1}(\frac{1}{2}-v)]dv \ \forall u \epsilon (0, \frac{1}{2}) \tag{6.4}$$

In the case in which X and Y are symmetric random variables, this last ordering is equivalent to $|X| \leq_L |Y|$. In the absence of symmetry, the Lorenz order of absolute values may be considered to be candidate variant kurtosis order. We may define

$$X \leq_8^k Y \text{ iff } |X| \leq_L |Y|. \tag{6.5}$$

This ordering has an attractive simplicity. It certainly captures some of the idea of kurtosis when the random variables are symmetric. Interpretation in the asymmetric case is potentially more problematic. It is not difficult to construct an asymmetric example in which $X \leq_\lambda^k Y$ but $X \not\leq_8^k Y$ and an example in which $X \leq_8^k$ but $X \not\leq_\lambda^k Y$. One advantage of the absolute Lorenz ordering ($\leq_8^k$) is its potential for straightforward extension to higher dimensions. For m dimensional random vectors $\underline{X}$ and $\underline{Y}$ centered to have medians $\underline{0}$, we can define $\underline{X} \leq_8^k \underline{Y}$ if and only if $d(\underline{X}, \underline{0}) \leq_L d(\underline{Y}, \underline{0})$ where d is a metric in $\mathbb{R}^m$.

More details on these kurtosis orderings and related summary measures of kurtosis will appear in a separate report.

References

Arnold, B. C. (1987) *Majorization and the Lorenz Order: A Brief Introduction.* Lecture Notes in Statistics **43**, Springer-Verlag, Berlin.

Balanda, K. P. and MacGillivray, H. L. (1988) Kurtosis: a critical review. *Amer. Statist.* **42** 111–119.

David, F. N. and Johnson, N. L. (1956) Some tests of significance with ordered variables. *J. Royal Stat. Soc.* **B18** 1–20.

MacGillivray, H. L. (1986) Skewness and asymmetry: measures and orderings. *Ann. Statist.* **14** 994–1011.

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

van Zwet, W. R. (1964) *Convex Transformations of Random Variables.* Mathematisch Centrum, Amsterdam.

Department of Statistics
University of California, Riverside
Riverside, CA 92502

Department of Statistics
Iowa State University
Ames, IA 50010