Dipartimento di Economia Marco Biagi

DEMB Working Paper Series

N. 164

Assessing skewness in financial markets

Giovanni Campisi[1], Luca La Rocca[2], Silvia Muzzioli[3]

January 2020

[1] University of Modena and Reggio Emilia, Department of Economics Marco Biagi
Address: Viale Berengario 51, 41121, Modena, Italy
E-mail: giovanni.campisi@unimore.it

[2] University of Modena and Reggio Emilia, Department of Physics, Informatics and Mathematics
Address: Via Campi 213/B, 41125, Modena, Italy
E-mail: luca.larocca@unimore.it

[3] University of Modena and Reggio Emilia, Department of Economics Marco Biagi and CEFIN
Address: Viale Berengario 51, 41121, Modena, Italy
E-mail: silvia.muzzioli@unimore.it

ISSN: 2281-440X online

Assessing skewness in financial markets

Giovanni Campisi[1], Luca La Rocca[2], and Silvia Muzzioli[1]

[1]*Marco Biagi Department of Economics, University of Modena and Reggio Emilia*, Via Jacopo Berengario 51, 41121 Modena, Italy (giovanni.campisi@unimore.it, silvia.muzzioli@unimore.it)
[2]*Department of Physics, Informatics and Mathematics, University of Modena and Reggio Emilia*, Via Giuseppe Campi 213/b, 41125 Modena, Italy (luca.larocca@unimore.it)

January 31, 2020

Abstract

It is common knowledge that investors like large gains and dislike large losses. This translates into a preference for right-skewed return distributions, with right tails heavier than left tails. Skewness is thus interesting not only as a way to describe the shape of a distribution, but also for risk measurement. We review the statistical literature on skewness and provide a comprehensive framework for its assessment. We present a new measure of skewness, based on a relative comparison between above average and below average returns. We show that this measure represents a valid complement to the state of the art.

Keywords: Asymmetry; Convex ordering; Unimodality; Volatility.

1 Introduction

In the financial literature there is a large variety of works dealing with the construction and implementation of risk measures to allow investors to make informed trading decisions. In light of the recent financial crises, the importance of these measures is increased in order to prevent tail risk events.

One of the most important risk measures is the VIX index introduced by the Chicago Board Options Exchange (CBOE). It is a forward looking measure in the sense that it measures volatility that investors expect to see in the next month (Whaley, 2009). Volatility indices are deemed by market operators to capture *market fear*: high index values are associated with high uncertainty in the underlying market, low index values with stable conditions; see Whaley (2000) and Muzzioli (2013b). Traditionally, financial returns are assumed to follow a normal distribution. In this context, volatility is a good measure of risk based on the idea that investors dislike uncertainty. However, a huge amount of works highlight that financial returns are non-normally distributed; see e.g. Fama (1965), Peiró (1999), Lempérière *et al.* (2017) and Elyasiani *et al.* (2018). Specifically, financial returns are found to present an empirical distribution with heavy tails and a negative skew. In other words, extreme and negative events are more probable to arise than in the normal distribution. This has a consequence on the need to include higher order moments as indicators of market risk.

One example is the CBOE SKEW index. The CBOE SKEW index has been listed on the CBOE since February 2011 to measure the tail risk not fully captured by the VIX index. While VIX measures the overall risk in the 30-day $S\&P500$ log-returns without disentangling the probabilities attached to positive and negative returns, the skewness index (CBOE SKEW) is intended to measure the perceived tail risk, i.e., the probability that investors attach to extreme negative returns. The CBOE SKEW index relies on Pearson's (third order) moment coefficient of skewness. It is well known that Pearson's moment coefficient of skewness is not a robust measure of skewness. In the statistical literature there are several examples in which Pearson's moment coefficient of skewness leads to controversial conclusions. From a financial point of view, this could have serious problems. Indeed, the role of the CBOE SKEW index as an indicator of market fear has been questioned since it moves frequently in the same direction as returns.

We face the need to investigate other measures of skewness and apply them to measure risk in financial markets. A problem to be overcome concerns the robustness of higher order moments since they may even be not well-defined for distributions with polynomial tails (a significant fraction of heavy tailed distributions). An interesting direction is to study the properties of skewness measures and apply them to financial markets. We contribute to the ongoing debate in this area by studying a variety of skewness measures, grounding their relevance to risk measurement on the simple observation that investors like large gains and dislike large losses.

There is a rich statistical literature on skewness, which we review in this paper to present a comprehensive framework for its assessment. We find that only the most classical measures of skewness rely on higher-order moments, while a number of other more recent measures do not. We also propose a new measure of skewness: the Risk Asymmetry indeX (RAX) based on the same measure introduced by Elyasiani *et al.* (2018) in a model-free set up with financial options. We contribute to the investigation of alternative skewness measures from a statistical perspective, arguing that the notion of skewness can play an important role in risk measurement.

The paper proceeds as follows. In Section 2 we present our framework for the study of univariate skewness, with a review of the tools available for both qualitative and quantitative assessments, identifying the statistical properties that a valid measure of skewness should satisfy. In Section 3 we propose the new measure of skewness, focussing on its statistical properties and interpretation. Section 4 contains an analysis of the performance of the most relevant measures of skewness. Section 5 discusses our work in terms of limitations and directions for future research.

2 Symmetry and asymmetry

Skewness is defined as a relaxation of symmetry to allow for asymmetry in a specific direction. We therefore start our exposition by introducing the notion of symmetry, which is not at all controversial, at least in the univariate case. A univariate random variable X is *symmetric* about the real value m when $X - m$ and $m - X$ have the same distribution; see for instance Doksum (1975). This property of the distribution of X can be written as $\mathbb{P}\{X \leq m - t\} = \mathbb{P}\{X \geq m + t\}$ for all $t > 0$, which amounts to saying that all corresponding left and right tails of X with respect to m have the same weight. Letting $t \to 0$, it is straightforward to see that m has to be the *median* of X, uniquely defined as the midpoint of the interval formed by all ν such that $\mathbb{P}\{X \leq \nu\} \geq 1/2$ and $\mathbb{P}\{X \geq \nu\} \geq 1/2$. It is also immediate to see that symmetry is *location-scale invariant*: if X is symmetric and Y is a positive affine transformation of X, that is, $Y = \alpha + \beta X$ with $\beta > 0$, then Y is also symmetric. Further considerations will be eased by the introduction of a suitable distributional framework.

Let F be the distribution function of X. We define the *support interval* of F as the open interval $]a, b[$ with left endpoint $a = \inf\{x \in \mathbb{R} \mid F(x) > 0\}$ and right endpoint $b = \sup\{x \in \mathbb{R} \mid F(x) < 1\}$. Note that we can possibly have $a = -\infty$ or $b = +\infty$. We assume that F is continuous on the real line and strictly increasing on its support interval. More specifically, we assume that F is obtained from a probability density function f that is continuous on $]a, b[$ and such that $f(x) > 0$ for all $x \in]a, b[$, while $f(x) = 0$ if $x \leq a$ or $x \geq b$. Let $\mathcal{F}_0$ be the class of all such distributions. Piecewise continuous density functions could be allowed to enlarge $\mathcal{F}_0$, but we here favor simplicity over generality. Nonetheless, we point out that such an enlargement would provide scope for probability density histograms (representing an important class of data based distributions). Interesting subclasses of $\mathcal{F}_0$ are obtained by assuming that X has finite moments up to some order; let $\mathcal{F}_k = \{F \in \mathcal{F}_0 \mid \mathbb{E}|X|^k < \infty\}$, $k = 1, 2, \ldots$, be such classes. If $F \in \mathcal{F}_1$, we denote by $\mu = \mathbb{E}(X)$ the *mean* of X. If $F \in \mathcal{F}_2$, we denote by $\sigma^2 = \mathbb{E}(X - \mu)^2$ the *variance* of X (with σ denoting the *standard deviation* of X); note that $\sigma^2 > 0$ because F is continuous.

Assuming $F \in \mathcal{F}_0$, the condition for symmetry can be written as

$$F(m - t) + F(m + t) = 1 \quad \text{for all } t > 0, \tag{1}$$

where $m = F^{-1}(1/2)$ with F^{-1} uniquely defined on the open interval $]0, 1[$ as the inverse function of F. Note that F^{-1}, called the *quantile function* of X, is also continuous and strictly increasing. Equivalently, condition (1) can be rewritten as

$$f(m-t) = f(m+t) \quad \text{for all } t > 0; \tag{2}$$

differentiating (1) with respect to t gives (2), while integrating (2) with respect to t gives (1). Note that symmetry implies either $a = -\infty$ and $b = +\infty$ or $a > -\infty$ and $b =< +\infty$ with $m = (a+b)/2$. If $F \in \mathcal{F}_1$, it is apparent from (2) that $\mu = m$ and μ can replace m in (1) as well as in (2).

A special case of interest is given by unimodal distributions. Following Dharmadhikari and Joag-Dev (1988, p. 2), we say that X is *unimodal* at $x_\star \in]a, b[$ if F is convex on $]a, x_\star[$ and concave on $]x_\star, b[$, which corresponds to f increasing on $]a, x_\star[$ and decreasing on $]x_\star, b[$. We say that X is unimodal (*tout court*) if it is unimodal at some $x_\star$. In this case, the *mode* of X (denoted by M) can be uniquely defined as the midpoint of the interval formed by all $x_\star$ such that X is unimodal at $x_\star$. For all $k = 0, 1, 2, \ldots$, we define $\mathcal{F}_k^\star = \{F \in \mathcal{F}_k \mid X \text{ is unimodal}\}$. It is apparent that (2) implies $M = m$, assuming $F \in \mathcal{F}_0^\star$, and M can replace m in (1) as well as in (2). Hence, for all symmetric X with $F \in \mathcal{F}_1^\star$, we have $M = m = \mu$ and the three classical measures of central tendency coincide.

If X is not symmetric, we say that X is *asymmetric*. While all symmetric distributions are alike in symmetry, each asymmetric distribution is asymmetric in its own way. Skewness relaxes symmetry to allow for a specific type of asymmetry: a random variable is left-skewed when its left tails are heavier than its right tails, it is right-skewed when its right tails are heavier than its left tails. A strictly skewed variable is a skewed variable that is not symmetric and, as such, it represents a way of being asymmetric. We formalize these notions in the following: Section 2.1 deals with assessing when a random variable is manifestly left-skewed or right-skewed; Section 2.2 deals with assessing how much of a skew a given random variable exhibits (to the left or to the right) even though such a skew may not be manifest.

2.1 Qualitative assessment of skewness

We start with a simplifying remark: since the left-tails of a random variable X are the right tails of the opposite random variable $-X$, and vice versa, it will be enough to assess when X is right-skewed; X will be left-skewed when $-X$ is right-skewed. We proceed with a quick presentation of the approach by Doksum (1975), which we will not adopt, but which provides an interesting starting point for our discussion.

According to Doksum (1975), a random variable X is *skew to the right* when $\mathbb{P}\{X \leq m-t\} \leq \mathbb{P}\{X \geq m+t\}$ for all $t > 0$, that is, when the right tails of X with respect to its median m are uniformly heavier than the corresponding left tails. This condition is completely general, requiring no assumption on X, and easy to interpret, because it is based on a specific comparison of tails. However, the latter feature is also its main limitation, in that a specific measure of central tendency (the median) is used. If the mean or mode of X are also available, different tails can be compared and, in general, they will lead to different assessments of skewness; see Sato (1997) for an illustration of the different implications of different choices. We will eventually get to a definition of skewness that does not rely on a specific measure of central tendency, but we first reformulate the definition of Doksum (1975) in our distributional framework. If $F \in \mathcal{F}_0$, the condition for skewness to the right can be written as

$$F(m-t) + F(m+t) \leq 1 \quad \text{for all } t > 0, \tag{3}$$

which is known as *van Zwet's condition* (Abadir, 2005) after van Zwet (1979) introduced it to prove the celebrated mode-median-mean inequality. This inequality is usually considered a sign of right skewness and (3) will justify this consideration, because we will eventually adopt a definition of right skewness that implies (3). Note that the distribution function of $-X$ is $\bar{F}(x) = 1 - F(-x)$, $x \in \mathbb{R}$, while the median of X is $-m$, so that the condition for skewness to the left is obtained from (3) simply by changing the direction of the inequality; in this case the mean-median-mode inequality holds for X.

We now consider the harder problem of assessing when a random variable Y is more right-skewed than another random variable X. Then, as recommended by MacGillivray (1986), we will say that X is right-skewed when it is more right-skewed than $-X$. Since the harder problem can be solved without relying on any specific measure of central tendency, this strategy overcomes the main limitation of the approach by Doksum (1975). Furthermore, the harder problem is interesting in its own and its solution provides a useful formal notion of relative skewness.

Following van Zwet (1964), we compare X with Y by means of the function $R(x) = G^{-1}(F(x))$, $a < x < b$, where G is the distribution function of Y. We call $R = G^{-1} \circ F$ the *quantile-quantile function* of Y against X, because it is the function whose graph is represented in the Q-Q plot with X on the horizontal axis and Y on the vertical axis. We say that X is *less right-skewed* than Y, or Y is *more right-skewed* than X, and write $F \precsim G$, or $G \succsim F$, if R is convex, which indicates that the left tails of X are progressively heavier than the left tails of Y and the right tails of Y are progressively heavier that the right tails of X. Note that R is strictly increasing (as well as continuous) on $]a, b[$ and $R(m)$ is the median of Y. The quantile-quantile function of X against Y is given by $R^{-1}(y) = F^{-1}(G(y))$, $c < y < d$, where $]c, d[$ is the support interval of G. Since $R^{-1} = F^{-1} \circ G$ is convex if and only if R is concave, we have $F \succsim G$ if and only if R is concave. An interesting characterization of this kind of comparison is that $F \precsim G$ if and only if Y is equal in distribution to a strictly increasing convex transformation of X: on the one hand, the variable $R(X)$ has the same distribution as Y; on the other hand, if $Y = \varphi(X)$ with φ strictly increasing and convex, then $R = \varphi$. If both $F \precsim G$ and $F \succsim G$, we write $F \sim G$. Clearly, this happens when R is a positive affine function, that is, when Y is equal in distribution to a positive affine transformation of X. This makes relative skewness a property of location-scale models rather than individual distributions. Finally, since $F \precsim G$ if and only if $\bar{F} \succsim \bar{G}$, we find that X is less right-skewed than Y if and only if $-X$ is more right-skewed than $-Y$; this means that we can safely interpret $F \preceq G$ as X being *more left-skewed* than Y, or Y being *less left-skewed* than X.

It is straightforward to check that $\precsim$ is *reflexive* and *transitive*: $F \precsim F$ and $F \precsim G, G \precsim H$ implies $F \precsim H$, where H denotes the distribution function of a third variable Z. This means that the relationship $\precsim$ is a *preorder* on $\mathcal{F}_0$, which justifies its common name of *convex ordering* of distributions and qualifies $\sim$ as the equivalence relationship defined by $\precsim$. Several other orderings of distributions have been proposed in the skewness literature (Oja, 1981; MacGillivray, 1986; Arnold and Groeneveld, 1993) and it turns out that the convex ordering is the strongest one. This is because it only looks at the convexity of the quantile-quantile function, without reference to any measure of central tendency, and thus it only signals the most manifest cases of relative skewness.

As anticipated, we say that X is *right-skewed* when $-X \precsim X$ (and *left-skewed* when $X \precsim -X$); note that here, for the sake of expressiveness, we apply $\precsim$ to $-X$ and X rather than to $\bar{F}$ and F. Since the convex ordering of distributions actually compares location-scale models, we can compare $m - X$ with $X - m$ instead of $-X$ with X, which we find convenient to investigate the connection to the approach by Doksum (1975). Indeed, the quantile-quantile function of $X - m$ against $m - X$ is given by

$$\bar{R}(x) = F^{-1}(\bar{F}(x - m)) - m, \quad -(b - m) < x < m - a, \tag{4}$$

because $\bar{F}(x - m)$, $x \in \mathbb{R}$, is the distribution function of $m - X$ and $F^{-1} - m$ is the quantile function of $X - m$; clearly $\bar{R}(0) = 0$. At the same time, the slope of $\bar{R}$ is given by

$$\bar{r}(x) = \frac{f(m - x)}{f(F^{-1}(\bar{F}(x - m)))}, \quad -(b - m) < x < m - a, \tag{5}$$

and clearly $\bar{r}(0) = 1$. It then follows from the convexity of $\bar{R}$ that $x \leq \bar{R}(x)$ for all x between $-(b-m)$ and $m-a$, which in turn implies (3). Hence, if X is right-skewed according to our chosen definition, it is also skew to the right in the sense of Doksum (1975). The converse is not true: convexity is clearly a stronger requirement than asking for the graph of $\bar{R}$ to lie entirely above the graph of the identity function; it is even stronger than requiring $\bar{r}(x) \geq 1$ for all x between 0 and $m-a$, equivalently $\bar{r}(x) \leq 1$ for all x between $-(b-m)$ and 0, which represents the notion of *strong skewness to the right* given by Doksum (1975).

We conclude our discussion of qualitative skewness by remarking that a random variable X is symmetric if and only if it is both right-skewed and left-skewed ($-X \sim X$). A notion of *strict skewness* can then be defined by excluding symmetric distributions ($-X \prec X$ or $X \prec -X$). In this way, we are able to partition $\mathcal{F}_0$ in four groups of distributions: symmetric distributions, strictly left-skewed distributions, strictly right-skewed distributions, and other asymmetric distributions. Note that, since symmetry and skewness are location-scale invariant, all F in a given location-scale model will belong to the same group. We will see that the other asymmetric distributions in the fourth group can be adjudicated as cases of negative (left) or positive (right) skewness through the choice of a suitable measure of skewness, which is the topic of the next subsection.

2.2 Quantitative assessment of skewness

We here consider the more ambitious problem of measuring how much a given distribution is skewed. This means associating to every distribution of interest a real number, whose sign captures the direction of skewness and whose absolute value is larger when skewness is more pronounced. More formally, within our distributional framework, for distributions with finite moments up to order $k \in \{0, 1, 2, \dots\}$, we aim at specifying a functional $\mathrm{Sk} : \mathcal{F}_k \to \mathbb{R}$, which will be called a *measure of skewness* of order k and will be required to satisfy the following properties:

(P1) $\mathrm{Sk}(\bar{F}) = -\mathrm{Sk}(F)$ for all $F \in \mathcal{F}_k$;

(P2) $\mathrm{Sk}(F) \leq \mathrm{Sk}(G)$ whenever $F \precsim G$.

If we are only interested in unimodal distributions, we can replace $\mathcal{F}_k$ with $\mathcal{F}_k^\star$ and specify a *unimodal measure of skewness* of the same order. We will say that Sk is *valid* to stress that (P1) and (P2) hold.

The meaning of (P1) and (P2) is that we want our quantitative assessment to respect our qualitative assessment. Indeed, it follows from (P1) and (P2) that $\mathrm{Sk}(F) \geq 0$ if F is right-skewed, while $\mathrm{Sk}(F) \leq 0$ if F is left-skewed, so that $\mathrm{Sk}(F) = 0$ if F is symmetric. Furthermore, it follows from (P2) that $\mathrm{Sk}(F) = \mathrm{Sk}(G)$ if $F \sim G$, that is, we require any valid measure of skewness to be location-scale invariant. Hence, to all effects, we are making the same assumptions as Groeneveld and Meeden (1984), which are rooted in the foundational work of Oja (1981) and have been consistently used in later work (Groeneveld, 1991a,b; Arnold and Groeneveld, 1995; Tajuddin, 1996, 1999; Groeneveld and Meeden, 2009). If we define *positive skewness* as $\mathrm{Sk}(F) \geq 0$ and *negative skewness* as $\mathrm{Sk}(F) \leq 0$, we are able to partition $\mathcal{F}_k$ (or $\mathcal{F}_k^\star$) in three groups of distributions: distributions with strictly positive skew, distributions with strictly negative skew, and distributions with null skew. Properties (P1) and (P2) guarantee that such a partition does not contradict the partition obtained in the previous subsection in terms of left and right skewness. We will call two measures *equivalent* when they give rise to the same partition.

In the following, we review the literature on measuring skewness with (P1) and (P2) in mind. We also care about the values taken by Sk. Let $\underline{s} = \inf_F \mathrm{Sk}(F)$ and $\bar{s} = \sup_F \mathrm{Sk}(F)$ be its extrema. Note that $\underline{s} = -\bar{s}$ by (P1). If Sk is a valid measure of skewness and φ is an odd and (strictly) increasing real function defined on the image of Sk, then $\varphi \circ \mathrm{Sk}$ is another (equivalent) valid measure of skewness. This means that any measure of skewness can be transformed so as to have -1 and $+1$ as extrema, if $\bar{s}$ is known. However, this is not necessarily the case and, even if it is, there is an interest for understanding whether the extrema can be attained or not and which distributions get close to them. Finally, we pay attention to the estimation of $\mathrm{Sk}(F)$ when a random sample $X_1, \dots, X_n$ from F is available. A general solution, in our distributional framework, is kernel density estimation on a suitable interval, but it is typically sufficient (and easier) to estimate finite dimensional summaries of F (mean, variance, …).

2.2.1 Higher-order measures

The study of skewness was pioneered by Pearson (1895, 1901, 1916). Indeed, the most classical measure of skewness goes under the name of *Pearson's moment coefficient of skewness*:

$$\gamma_j = \frac{\mathbb{E}(X-\mu)^{2j+1}}{\sigma^{2j+1}}, \tag{6}$$

where typically $j = 1$, but possibly $j = 2, 3, \dots$; since γ_j is well-defined when the distribution of X has finite moments up to order $2j + 1$, the choice $j = 1$ is the least demanding in terms of distributional assumptions. The rationale behind (6) is to use higher order moments to gauge how much the right tails of X are heavier than its left tails. This strategy obtains a valid $(2j+1)$th order measure of skewness: it is apparent that (6) satisfies (P1) and it was shown by van Zwet (1964) that (6) satisfies (P2); see also Oja (1981) and MacGillivray (1986).

The third order measure of skewness γ_1 (standardized third central moment of X) is also called *Fisher-Pearson coefficient of skewness* (Doane and Seward, 2011); see Arnold and Groeneveld (1995) for its historical attribution and Holgersson (2010) for the link to Fisher (1929). As an aside, note that Karl Pearson was not interested in the sign of skewness and used $\beta_1 = \gamma_1^2$ in place of γ_1. In order to compute γ_1, we can write

$$\mathbb{E}(X-\mu)^3 = \mathbb{E}(X^3) - 3\mu\,\mathbb{E}(X^2) + 2\mu^3. \tag{7}$$

As pointed out by Groeneveld and Meeden (1984), if X follows a Pareto distribution with unit scale and large enough shape, that is, $f(x) = \theta/x^{\theta+1}$ for $x > 1$ and $f(x) = 0$ for $x \leq 1$, with $\theta > 3$, then $\mathbb{E}(X^k) = \theta/(\theta-k)$, for $k = 1, 2, 3$, and $\mathbb{E}(X^3)$ is arbitrarily large for θ close to 3, while $\mathbb{E}(X^2)$ tends to 3, μ tends to $3/2$ and σ^2 tends to $3/4$; it follows that γ_1 is arbitrarily large for θ close to 3 and we conclude that $\bar{s} = +\infty$.

In the right hand side of (7), for $k = 1, 2, 3$, we will estimate $\mathbb{E}(X^k)$ by $\langle X^k \rangle = n^{-1}\sum_{i=1}^{n} X_i^k$ (corresponding sample moment) if $X_1, \ldots, X_n$ are observed; note that for $k = 1$ we will estimate μ by the sample mean $\hat{\mu} = \langle X \rangle$. Then, in (6), we will estimate σ by the square root $\hat{\sigma}$ of the sample variance $\hat{\sigma}^2 = \langle X^2 \rangle - \langle X \rangle^2$. The value $\hat{\gamma}_1$ obtained in this way (sample moment coefficient of skweness) can be adjusted for sample size, but we are not interested in such an adjustment here; see Doane and Seward (2011) for information and references on this topic. Egon Sharpe Pearson, together with H. O. Hartley, provided tables to use $\hat{\gamma}_1$ as a test for departure from normality (Doane and Seward, 2011); see also Holgersson (2010) on testing asymmetry. Finally, the sharp algebraic bound $|\hat{\gamma}_1| \leq (n-2)/(n-1)^{1/2}$ holds for all samples of size n (Wilkins, 1944; Kirby, 1974; Cox, 2010) even though we have seen that γ_1 can take arbitrarily large values.

As illustrated by Li and Morris (1991), in some cases γ_1 may not express asymmetry well. Furthermore, being based on the third order moment, γ_1 is strongly influenced by outliers; see for instance Groeneveld (1991a). This lack of robustness, together with an appetite for broadening the domain of definition, motivates the investigation of alternative measures of skewness.

2.2.2 Unimodal measures

A second measure of skewness that dates back to the pioneering work of Pearson (1895) is called *Pearson's mode coefficient of skewness* or *Pearson's first coefficient of skewness*:

$$S'_K = \frac{\mu - M}{\sigma}, \tag{8}$$

well-defined for $F \in \mathcal{F}_2^\star$. For instance, if $f(x) = x^{\alpha-1}e^{-x}/\Gamma(\alpha)$ for $x > 0$ and $f(x) = 0$ for $x \leq 0$, with $\alpha > 1$, we find $\mu = \alpha$, $M = \alpha - 1$ and $\sigma = \sqrt{\alpha}$, so that $S'_K = 1/\sqrt{\alpha}$. Remarkably, in this case (gamma distribution with unit scale), the equality $S'_K = \gamma_1/2$ holds (Arnold and Groeneveld, 1995). The quantity $\alpha_1 = \gamma_1/2$ is called *coefficient of momental skewness* (Zwillinger and Kokoska, 1999, p. 18) and is clearly equivalent to γ_1. In general, of course, α_1 and γ_1 are not equivalent to S'_K. The rationale behind (8) is that, as discussed in Section 2.1, the mode-median-mean inequality is a sign of right skewness. If $-X \preceq X$, then $S'_K > 0$ and (8) gauges the width of inequality. It is immediate to see that S'_K satisfies (P1). However, as illustrated by Arnold and Groeneveld (1995), the coefficient S'_K does not satisfy (P2): compatibility with right skewness does not extend to full compatibility with the convex ordering of distributions. Hence, we cannot regard S'_K as a valid measure of skewness.

Arnold and Groeneveld (1995) proposed to replace (8) by

$$\gamma_{AG} = \mathbb{P}\{X \geq M\} - \mathbb{P}\{X \leq M\} = 1 - 2F(M), \tag{9}$$

which we name *Arnold-Groeneveld coefficient of skewness*. The coefficient γ_{AG} is well-defined for all $F \in \mathcal{F}_0^\star$, because it does not involve any moment of X, which is an improvement in itself. The rationale behind (9) is an implicit comparison between M and m (in place of μ): if $M \leq m$ then $F(M) \leq 1/2$ and $\gamma_{AG} \geq 0$. In this way, like before, right skewness implies $\gamma_{AG} \geq 0$ through the mode-median-mean inequality. In addition, differently from before, property (P2) is satisfied. This was shown by Arnold and Groeneveld (1995) assuming differentiable probability density functions, but it holds for all $F, G \in \mathcal{F}_0^\star$ with modes M_X and M_Y, respectively, that $F \preceq G$ implies $F(M_X) \geq G(M_Y)$; this follows from $F = G \circ R$, where $R = G^{-1} \circ F$, and the definition of unimodality. As for property (P1), it follows from the equality $\bar{F}(-M) = 1 - F(M)$. We conclude that γ_{AG} is a valid unimodal measure of skewness of order zero (best possible order).

The coefficient γ_{AG} takes values in $[-1, 1]$ and the equality $\gamma_{AG} = 1$ is attained when $M = a$, which requires $a > -\infty$ in the support interval of F, while $\gamma_{AG} = -1$ when $M = b$, which requires $b < +\infty$. It follows that all decreasing densities exhibit maximal positive skewness, while all increasing densities exhibit maximal negative skewness. This is clearly a limitation, because γ_{AG} cannot discriminate between monotone densities of the same type. A sample version $\hat{\gamma}_{AG}$ of (9) can be obtained from an estimator $\hat{M}$ of the mode and an estimator $\hat{F}$ of the distribution function; the latter can be the empirical distribution function, for simplicity, while the former can be one of the estimators of the mode implemented in package `modeest` (Poncet, 2019) for `R` (R Core Team, 2019); references on these estimators for unimodal distributions can be found in the same package.

2.2.3 First order measures

A third classical measure of skewness is called *Pearson's median coefficient of skewness* or *Pearson's second coefficient of skewness*:

$$S_K'' = 3\,\frac{\mu - m}{\sigma}, \tag{10}$$

where the leading (arbitrary) multiplicative constant stems from an approximation of (8); see Yule (1911, p. 150). Equation (10) is well-defined for $F \in \mathcal{F}_2$ and is based, like (8), on the mode-median-mean inequality. It is clear that S_K'' satisfies (P1), but like S_K', as shown by van Zwet (1964), S_K'' does not satisfy (P2). We therefore cannot consider S_K'' a valid measure of skewness. However, a valid replacement for (10) is provided by Groeneveld and Meeden (1984):

$$\gamma_{GM} = \frac{\mu - m}{\mathbb{E}\,|X - m|}, \tag{11}$$

which is well-defined for all $F \in \mathcal{F}_1$ and we name *Groeneveld-Meeden coefficient of skewness*. The broader domain of definition is an advantage in itself, property (P1) is clearly preserved and, moreover, γ_{GM} satisfies (P2), as shown by Groeneveld and Meeden (1984). The coefficient γ_{GM} is thus a valid measure of skewness of order one (best possible order using the mean). The mean absolute error turns out to be the right denominator for the difference between the mean and the median, if this is to be used as a measure of skewness.

We know from Jensen's inequality that $|\,\mathbb{E}(X - m)| \leq \mathbb{E}\,|X - m|$ with equality if and only if $\mathbb{P}\{X \geq m\} = 1$ or $\mathbb{P}\{X \leq m\} = 1$. It follows that $-1 < \gamma_{GM} < 1$ and the extrema of γ_{GM} are unattainable by continuous distributions; see Groeneveld (1991b) for the case of discrete distributions. A sample version $\hat{\gamma}_{GM}$ of (11) will be obtained by replacing m with the sample median $\hat{m}$ and $\mathbb{E}\,|X - m|$ with its sample counterpart $\langle |X - \hat{m}| \rangle$, as well as μ with $\hat{\mu}$. Finally, we point out an interesting interpretation of (11):

$$\gamma_{GM} = \frac{\mathbb{E}(X - m\,|X \geq m) - \mathbb{E}(m - X\,|X \leq m)}{\mathbb{E}(X - m\,|X \geq m) + \mathbb{E}(m - X\,|X \leq m)}; \tag{12}$$

see Groeneveld and Meeden (1984). In words, assuming for simplicity $m = 0$, we can say that γ_{GM} is the normalized difference between the mean gain conditional on a gain and the mean loss conditional on a loss.

A simple alternative first order measure of skewness was suggested by Tajuddin (1999) in parallel to γ_{AG}:

$$\gamma_T = \mathbb{P}\{X \leq \mu\} - \mathbb{P}\{X \geq \mu\} = 2F(\mu) - 1. \tag{13}$$

We call γ_T in (13) the *Tajuddin coefficient of skewness*, noting that Tajuddin (1996) had previously suggested the equivalent measure $\log\left(F(\mu)/\{1 - F(\mu)\}\right) = \log(1 + \gamma_T)/(1 - \gamma_T)$. Equation (13) is clearly well-defined for all $F \in \mathcal{F}_1$, it satisfies (P1), because $\bar{F}(-\mu) = 1 - F(\mu)$, and it satisfies (P2), because Jensen's inequality gives

$\mathbb{E}(Y) = \mathbb{E}(G^{-1}(F(X))) \geq G^{-1}(F(\mu))$ if $X \preceq Y$ ($F \preceq G$); see also Tajuddin (1996). It follows that γ_T is a valid alternative to γ_{GM}.

The rationale behind (13) is again the mode-median-mean inequality for right-skewed distributions: if a return is right-skewed, then it is probably below average. It may sound counterintuitive that investors like such returns, but a different wording is possible: if a return is right-skewed, then on average it is in the right tail of its distribution. This may sound more palatable, but neither formulation has any impact on the validity of γ_T. As for the values that γ_T can take, it is immediate to see that $-1 < \gamma_T < 1$. The estrema cannot be attained, because F is continuous, but we will present in Section 3 an example where $\gamma_T = 1 - 2\lambda \to 1$ as $\lambda \downarrow 0$. Finally, a sample version $\hat{\gamma}_T$ of γ_T can be obtained from (13) by estimating F with the empirical distribution function and μ with the $\hat{\mu} = \langle X \rangle$.

2.2.4 Zeroth order measures

None of the measures of skewness presented until now is well-defined for all $F \in \mathcal{F}_0$. A possibility in this sense is offered by the *quantile coefficient of skewness*

$$B_\alpha = \frac{\{F^{-1}(1-\alpha) - m\} - \{m - F^{-1}(\alpha)\}}{F^{-1}(1-\alpha) - F^{-1}(\alpha)} = \frac{\{F^{-1}(1-\alpha) + F^{-1}(\alpha)\}/2 - m}{\{F^{-1}(1-\alpha) - F^{-1}(\alpha)\}/2}, \tag{14}$$

where $\alpha \in \,]0, 1/2[$ and a typical choice is $\alpha = 1/4$. The quartile coefficient of skewness $B_{1/4}$ dates back to Bowley (1920, p. 116) and is called *Bowley-Yule coefficient of skewness*, because the coefficient $2B_{1/4}$ can be traced back to Yule (1911, p. 150). Groeneveld and Meeden (1984) introduced B_α, inspired by Hinkley (1975), and also let $\alpha \downarrow 0$ to obtain the coefficient $B_0 = (a + b - 2m)/(b - a)$ for distributions with bounded support interval, that is, with $a > -\infty$ and $b < +\infty$. Remarkably, if both the numerator and denominator in (14) are integrated with respect to α from 0 to $1/2$, before taking their ratio, the coefficient γ_{GM} in (11) emerges (assuming $F \in \mathcal{F}_1$). It was shown by Groeneveld and Meeden (1984) that B_α satisfies (P1) and (P2) for all $\alpha \in [0, 1/2[$. Hence, we have a family of valid measures of skewness that can be used without any assumption on the moments of the distribution (zeroth order measures of skewness).

Groeneveld and Meeden (2009) suggest a variant of (14) that is appropriate when the direction of skewness is known *a priori*, but we do not deal with this case here. Brys *et al.* (2003) argue that the octile coefficient $B_{1/8}$ is more appropriate to detect asymmetry than the quartile coefficient $B_{1/4}$, because it uses more information from the tails of the distribution, but they also note that $B_{1/4}$ is less sensitive to outliers (more robust) than $B_{1/8}$. In the end, this tension between sensitivity and robustness is at the heart of the choice of α in (14) and, more generally, of a measure of skewness or any other distributional summary.

It is easy to see that $-1 < B_\alpha < 1$ for all $\alpha \in [0, 1/2[$; the extreme values are unattainable by continuous distributions, because $B_\alpha = -1$ would require $F^{-1}(1-\alpha) = m$ and $B_\alpha = 1$ would require $F^{-1}(\alpha) = m$, but see Groeneveld (1991b) for discrete distributions. A sample version of (14) will be obtained by replacing all quantiles of F by their sample counterparts (quantiles of the empirical distribution function); in particular, of course, $\hat{m}$ will replace m. Finally, the coefficient $B_{1/4}$ admits an interpretation analogous to that of γ_{GM}, but with conditional medians in place of conditional means, while the coefficient B_0 can be interpreted in terms of conditional ranges (Groeneveld and Meeden, 1984).

The coefficient B_α defined by (14) features in an interesting decomposition of the quantile function:

$$\begin{aligned} F^{-1}(\alpha) &= m - S_\alpha(1 - B_\alpha)/2, \\ F^{-1}(1/2) &= m, \\ F^{-1}(1-\alpha) &= m + S_\alpha(1 + B_\alpha)/2, \end{aligned} \tag{15}$$

where α varies in $]0, 1/2[$ and $S_\alpha = F^{-1}(1-\alpha) - F^{-1}(\alpha)$ is the αth inter-quantile range; see Benjamini and Krieger (1996). The decomposition (15) links the measure of skewness B_α to the measure of scale S_α and the measure of location m. These three measures, together, determine the quantile function (letting α vary in $]0, 1/2[$). Note that $B_\alpha > 0$ implies $(1 + B_\alpha)/2 > 1/2$ and $(1 - B_\alpha)/2 < 1/2$, while $B_\alpha < 0$ implies $(1 + B_\alpha)/2 < 1/2$

and $(1-B_\alpha)/2 > 1/2$. If $B_\alpha = 0$, then $F^{-1}(1-\alpha) - m = m - F^{-1}(\alpha)$ and (1) holds for their common value t (a sign of symmetry). Indeed, setting $B_\alpha = 0$ in (1) produces a symmetric distribution; see also Doksum (1975).

3 The risk asymmetry index

Let X be a random variable with mean μ and standard deviation σ. The centered variable $X - \mu$ can be written as the sum of its positive part $(X-\mu)_+ = \max(0, X-\mu)$ and its negative part $(X-\mu)_- = \max(0, \mu - X)$. Accordingly, the variance of X can be written as $\sigma^2 = \sigma_U^2 + \sigma_D^2$, where $\sigma_U^2 = \mathbb{E}(X-\mu)_+^2$ is called the *upside variance* of X and $\sigma_D^2 = \mathbb{E}(X-\mu)_-^2$ is called the *downside variance* of X; the quantities σ_U and σ_D are called the *upside standard deviation* and *downside standard deviation* of X, respectively. From a financial viewpoint, σ_U represents "good" volatility and σ_D represents "bad" volatility, while σ represents "total" volatility. The *risk asymmetry index* (Elyasiani *et al.*, 2018) is defined as $\mathrm{RAX} = (\sigma_U - \sigma_D)/\sigma$ and represents the relative excess of "good" volatility (with respect to "bad" volatility) in the distribution of returns modeled by X. The rationale behind this definition is to compare above average returns with below average returns in terms of their root mean squared residuals. We show in the following that such a comparison results in a valid measure of skewness.

The risk asymmetry index can be rewritten as

$$\mathrm{RAX} = \sqrt{\frac{\sigma_U^2}{\sigma^2}} - \sqrt{1 - \frac{\sigma_U^2}{\sigma^2}} = \sqrt{1 - \frac{\sigma_D^2}{\sigma^2}} - \sqrt{\frac{\sigma_D^2}{\sigma^2}}, \tag{16}$$

that is, as a strictly increasing function of the *relative upside variance* σ_U^2/σ^2 or, alternatively, as the opposite function of the *relative downside variance* $\sigma_D^2/\sigma^2 = 1 - \sigma_U^2/\sigma^2$. This rewriting is useful to show that RAX is a valid measure of skewness. Indeed, property (P1) follows directly from the fact that the upside variance of $-X$ is the downside variance of X. As for property (P2), we first note that (16) is location-scale invariant. This allows us to focus on the standard case $\mu = 0$ and $\sigma = 1$. In this case, we have $\sigma_U^2 = \mathbb{E}\,X_+^2$ and it follows from Theorem 5.3 in Oja (1981) that $X \preceq Y$ implies $\mathbb{E}\,X_+^2 \le \mathbb{E}\,Y_+^2$. Then, by (16), the same inequality holds for RAX and (P2) holds. We conclude that RAX is a valid measure of skewness and we point out that it is a second order one, because (16) is well-defined for $F \in \mathcal{F}_2$. As such, RAX fills a gap in the literature reviewed in Section 2.2.

It is clear from the decomposition of variance in its upside and downside components that $0 < \sigma_U^2/\sigma^2 < 1$. If X is symmetric, then $\sigma_U^2 = \sigma_D^2$ and $\sigma_U^2/\sigma^2 = 1/2$, so that $\mathrm{RAX} = 0$. The following example shows that the relative upside variance can get arbitrarily close to 1 for a suitable choice of F in $\mathcal{F}_2$. Let X be a random variable with probability density function defined by

$$\begin{array}{rcl} f(x) & = & (1-\lambda) f_-(x) + \lambda f_+(x), \\ f_+(x) & = & \lambda e^{-\lambda x}, \quad x \ge 0, \\ f_-(x) & = & \frac{2\lambda^2}{(1-\lambda)^2} f_u(x) + \frac{1-2\lambda-\lambda^2}{(1-\lambda)^2} f_t(x), \\ f_t(x) & = & (1-\lambda)\{1 - |1 + (1-\lambda)x|\}, \quad -\frac{2}{(1-\lambda)} < x < 0, \\ f_u(x) & = & \frac{1-\lambda}{2}, \quad -\frac{2}{(1-\lambda)} < x < 0, \end{array} \tag{17}$$

assuming $f_+(x) = 0$ for $x < 0$ and $f_t(x) = f_u(x) = 0$ for $x \le -2/(1-\lambda)$ or $x \ge 0$, while $\lambda \in]0, \sqrt{2}-1[$; note that f is continuous at 0 and therefore on its support interval $]a, b[\, = \,] - 2/(1-\lambda), +\infty[$. It can be seen from (17) that $\mathbb{E}(X) = 0$, $\mathbb{E}(X_+^2) = 2/\lambda$ and $\mathbb{E}(X_-^2) \le \{2/(1-\lambda)\}^2(1-\lambda) = 4/(1-\lambda)$, so that $\sigma_D^2/\sigma_U^2 \le 2\lambda/(1-\lambda)$. It follows that $\sigma_U^2/\sigma^2 = 1/(1 + \sigma_D^2/\sigma_U^2) \to 1$, as $\lambda \downarrow 0$, while $\sigma_D^2/\sigma^2 \to 0$ and thus the relative upside variance of $-X$ can get arbitrarily close to 0. As a consequence, we have $\mathrm{RAX} \to 1$ for X and $\mathrm{RAX} \to -1$ for $-X$. Note that $\mathbb{P}\{X \le 0\} = 1 - \lambda$, while $\mathbb{P}\{X \ge 0\} = \lambda$, so that $\gamma_T = 1 - 2\lambda$ as anticipated in Section 2.2.

If a random sample $X_1, \ldots, X_n$ from the distribution of X is available, a sample version of RAX can be obtained from (16) by replacing σ^2 with the sample variance $\hat\sigma^2 = \langle X^2 \rangle - \langle X \rangle^2$ and σ_U^2 with the *sample upside variance* $\hat\sigma_U^2 = \langle (X - \hat\mu)_+^2 \rangle$ or σ_D^2 with the *sample downside variance* $\hat\sigma_D^2 = \langle (X - \hat\mu)_-^2 \rangle$, where of course $\hat\mu = \langle X \rangle$ is the sample mean. An advantage of RAX is that risk-neutral versions of these quantities are easy to obtain from option data in a model-free setup (Bakshi *et al.*, 2003; Muzzioli, 2013a,b); this was indeed the original setting of Elyasiani *et al.* (2018). In this setting σ_U and σ_D represent the upside corridor implied volatility and the downside corridor implied volatility, respectively (Carr and Madan, 1998; Muzzioli, 2013a,b). From an economic point of view the upside corridor is associated to "good" volatility, because it refers to the possibility of large gains. On the other hand, the downside corridor is associated to "bad" volatility due to the chance of large losses for investors. Note that, in principle, the whole risk-neutral distribution function of returns can be recovered from option data, because it is the discounted first derivative of the European put price, but in practice it can be tricky to go beyond the first moments; see Birru and Figlewski (2012) for an example of work in this direction.

4 Empirical findings

In this section we follow the approach of Brys *et al.* (2003) to examine the performance of RAX (introduced in Section 3) together with other valid measures of skewness. In particular, we consider simulated data from a gamma distribution and we test our measures for different degrees of the shape parameter α. Recall that a gamma distribution has probability density function given by

$$f_{\alpha,\beta}(x) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} x^{\alpha-1} e^{\frac{-x}{\beta}}, \qquad 0 < x < \infty, \tag{18}$$

where $\alpha > 0$ determines the shape of the distribution, while $\beta > 0$ is a scale parameter and can therefore be ignored since we are interested in skewness ($\beta = 1$ for concreteness). Arnold and Groeneveld (1995) highlight that the parameterization of gamma distributions in terms of α respects the convex ordering of distributions. Indeed, the Fisher-Pearson coefficient of skewness is straightforward to calculate: $\gamma_1 = 2/\sqrt{\alpha}$.

In Figure 1 several coefficients are presented for the gamma case. In detail, we generated 1000 samples, each of $n = 100$ observation, from the gamma distribution with α ranging from 0 to 3, and we plot the average estimated skewness versus the value of α. For our purposes, we take into consideration four measures of skewness: the Groeneveld-Meeden (γ_{GM}), the Bowley-Yule ($B_{1/4}$), the Risk Asymmetry indeX (RAX) and the Tajuddin (γ_T) coefficient of skewness. We expect that all four measures decrease monotonically in α and Figure 1 confirms this expectation: each measure start from 1 when α is near to zero (indicating high skewness) and drop towards zero as α grows (indicating low skewness). We see that $B_{1/4}$ and γ_T attain the smallest values, while RAX maintains between γ_{GM} (upper limit) and $B_{1/4}$ or γ_T (lower limit) for every α. This means that RAX has an intermediate sensitivity to changes in the shape of the gamma distribution. Table 1 reports the average estimates of the four skewness measures (with their standard errors in parenthesis) for several values of α and for $n = 30, 100, 1000$. We can see that the four measures of skewness behave more or less as in Figure 1 for all sample sizes.

As a further step, we analyze the robustness of the four measures upon varying the shape parameter of the gamma distribution. Specifically, we compare their behavior under the influence of a few outliers. To this purpose, in Figure 2(a) we propose the boxplots of the skewness estimates on 1000 random data sets for $\alpha = 1.5$ and considering $n = 1000$ observations, while Figure 2 (b) depicts the same boxplots where we replaced 15% of the data with outliers distant 8 standard deviations to the right of the mean. It can be seen that the median values increase for all skewness measures, bringing the boxes with them, but for RAX, which instead shows a decrease. We deeper explore robustness in Figure 3, where further simulations are depicted. In particular, Figure 3(a) shows for each measure and for several values of α the difference between the average estimated value at the contaminated and at the original data sets. As in Figure 2(b), we replaced a percentage of the original data with outliers under different contamination levels. In detail, we contaminated our data at 5% in Figure 3 (a) and at 15% in Figure 3 (b). We see that a decrease in skewness upon contamination is also possible for γ_T, but not for $B_{1/4}$ nor for γ_{GM}. This aspect deserves, in our opinion, further investigations. Finally, focussing on the absolute skewness change upon contamination, we note that $B_{1/4}$ stands out for being rather insensitive to the presence of outliers, while the performance of RAX is competitive with that of γ_T and γ_{GM}. We can conclude that RAX strikes a good balance between robustness to outliers and sensitivity to changes in the shape of the distribution.

Table 1: Average estimated skewness and standard error for 1000 samples of $n = 30, 100, 1000$ observations and for α ranging from 0.01 to 3.

	n	γ_{GM}	$B_{1/4}$	RAX	γ_T
$\alpha = 0.01$	30	1.0000 (0.0000)	1.0000 (0.0000)	0.7663 (0.0014)	0.8756 (0.0019)
	100	1.0000 (0.0000)	1.0000 (0.0000)	0.8413 (0.0011)	0.9109 (0.0007)
	1000	1.0000 (0.0000)	1.0000 (0.0000)	0.8881 (0.0005)	0.9197 (0.0002)
$\alpha = 0.2$	30	0.8780 (0.0030)	0.7185 (0.0058)	0.5210 (0.0030)	0.5105 (0.0035)
	100	0.9027 (0.0013)	0.7558 (0.0031)	0.5704 (0.0019)	0.5218 (0.0019)
	1000	0.9116 (0.0004)	0.7730 (0.0009)	0.5928 (0.0007)	0.5282 (0.0006)
$\alpha = 0.5$	30	0.6121 (0.0047)	0.4004 (0.0074)	0.3925 (0.0030)	0.3478 (0.0040)
	100	0.6262 (0.0026)	0.4119 (0.0041)	0.4281 (0.0019)	0.3623 (0.0021)
	1000	0.6343 (0.0009)	0.4198 (0.0013)	0.4473 (0.0007)	0.3651 (0.0007)
$\alpha = 1$	30	0.4202 (0.0051)	0.2430 (0.0075)	0.3029 (0.0030)	0.2530 (0.0041)
	100	0.4365 (0.0028)	0.2562 (0.0041)	0.3308 (0.0019)	0.2634 (0.0023)
	1000	0.4413 (0.0009)	0.2592 (0.0013)	0.3423 (0.0006)	0.2633 (0.0007)
$\alpha = 3$	30	0.2382 (0.0051)	0.1315 (0.0074)	0.1860 (0.0028)	0.1436 (0.0044)
	100	0.2444 (0.0029)	0.1341 (0.0040)	0.2037 (0.0016)	0.1512 (0.0024)
	1000	0.2473 (0.0009)	0.1372 (0.0014)	0.2094 (0.0005)	0.1545 (0.0008)

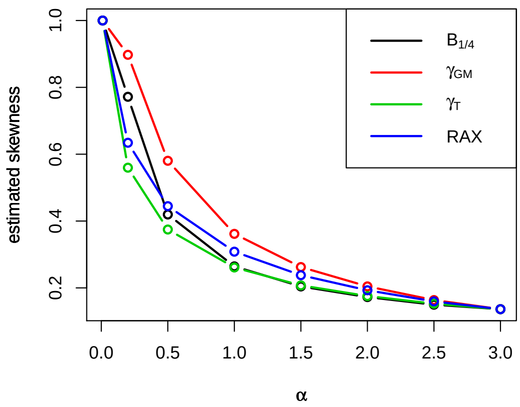

Figure 1: Average of four skewness measures over 1000 samples of size $n = 100$ for several values of the shape parameter α of the gamma distribution.

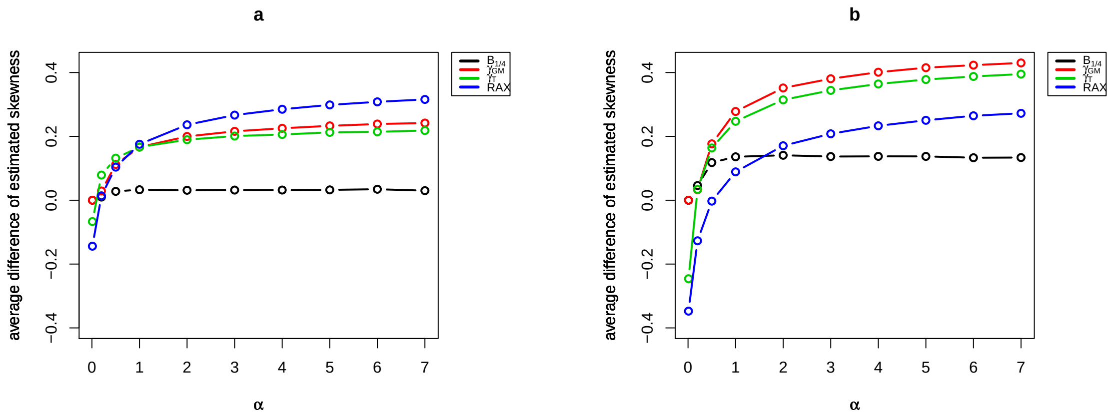

Figure 2: Boxplots of skewness estimates on 1000 random samples of $n = 1000$ observations from a gamma distribution with shape parameter $\alpha = 1.5$ without contamination (a) and with 15% contamination (b).

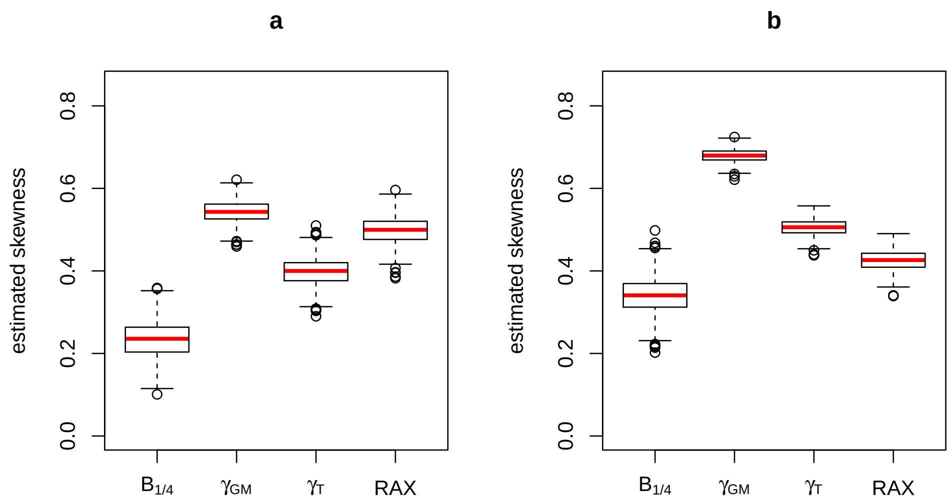

Figure 3: Difference between average skewness estimates at contaminated and at uncontaminated data for different values of α and 5% (a) and 15% (b) contamination respectively.

5 Discussion

We presented in this paper a comprehensive framework for the assessment of univariate skewness. We did not deal with multivariate skewness; see for instance Khattree and Bahuguna (2019) or Franceschini and Loperfido (2019) on this broader topic. We reviewed the statistical literature in search of principled and broadly applicable measures of skewness, but there are of course further alternatives. For instance, if a parametric model is adopted, a skewness parameter may be included in its parameterization. This is the case of the skew normal model that led Eling *et al.* (2010) to suggest *Azzalini's skewness parameter* as a measure of skewness; see Azzalini (2005) for a first-hand review with discussion of the skew normal distribution. This is also the case of the stable model, which has the advantage of allowing for very heavy tails; see e.g. Nolan (2003) for an introduction to stable distributions with focus on financial data. Another alternative approach was taken by Kashlak (2018), who suggested to measure distributional asymmetry with Wasserstein distance and Rademacher symmetrization.

We proposed a new measure of skewness, called RAX and based on work by Elyasiani *et al.* (2018), showing that our proposal is a valid second order measure of skeweness. RAX is the relative difference between upside and downside volatility. We used volatility, following Elyasiani *et al.* (2018), because of its high standing in finance. In principle, we could also compare above average returns with below average returns in terms of their mean absolute residuals, rather than root mean squared residuals, but we would still need volatility in the denominator to satisfy (P2).

In Section 2.2, the presentation of first order measures of skewness led us to remark that the mean absolute error was the right denominator to turn the difference between the mean and the median into a valid measure of skewness; this turned (10) into (11). We wonder whether there is right denominator for the difference between the mean and the mode to turn (8) into a valid measure of skewness.

Acknowledgements

The authors gratefully acknowledge financial support from the FAR2017 project "The role of Asymmetry and Kolmogorov equations in financial Risk Modelling (ARM)".

References

Abadir, K. M. (2005). The mean-median-mode inequality: counterexamples. *Econometric Theory*, **21**(2), 477–482.

Arnold, B. C. and Groeneveld, R. A. (1993). Skewness and kurtosis orderings: an introduction. In M. Shaked and Y. L. Tong, editors, *Stochastic Inequalities*, volume 22 of *Lecture Notes-Monograph Series*, pages 17–24. Institute of Mathematical Statistics, Hayward.

Arnold, B. C. and Groeneveld, R. A. (1995). Measuring skewness with respect to the mode. *The American Statistician*, **49**(1), 34–38.

Azzalini, A. (2005). The skew-normal distribution and related multivariate families. *Scandinavian Journal of Statistics*, **32**(2), 159–200. With discussion.

Bakshi, G., Kapadia, N., and Madan, D. (2003). Stock return characteristics, skew laws, and the differential pricing of individual equity options. *Review of Financial Studies*, **16**(1), 101–143.

Benjamini, Y. and Krieger, A. M. (1996). Concepts and measures for skewness with data-analytic implications. *Canadian Journal of Statistics*, **24**(1), 131–140.

Birru, J. and Figlewski, S. (2012). Anatomy of a meltdown: the risk neutral density for the S&P 500 in the fall of 2008. *Journal of Financial Markets*, **15**(2), 151–180.

Bowley, A. L. (1920). *Elements of Statistics*. Scribner, New York, 4 edition.

Brys, G., Hubert, M., and Struyf, A. (2003). A comparison of some new measures of skewness. In R. Dutter, P. Filzmoser, U. Gather, and P. J. Rousseeuw, editors, *Developments in Robust Statistics*, pages 98–113. Springer, Berlin.

Carr, P. and Madan, D. (1998). Towards a theory of volatility trading. In R. A. Jarrow, editor, *Volatility: New Estimation Techniques for Pricing Derivatives*, pages 417–427. Risk Books, London.

Cox, N. J. (2010). Speaking Stata: the limits of sample skewness and kurtosis. *The Stata Journal*, **10**(3), 482–495.

Dharmadhikari, S. and Joag-Dev, K. (1988). *Unimodality, Convexity, and Applications*. Academic Press, San Diego.

Doane, D. P. and Seward, L. E. (2011). Measuring skewness: a forgotten statistic? *Journal of Statistics Education*, **19**(2), 1–18.

Doksum, K. A. (1975). Measures of location and asymmetry. *Scandinavian Journal of Statistics*, **2**(1), 11–22.

Eling, M., Farinelli, S., Rossello, D., and Tibiletti, L. (2010). Skewness in hedge funds returns: classical skewness coefficients vs Azzalini's skewness parameter. *International Journal of Managerial Finance*, **6**(4), 290–304.

Elyasiani, E., Gambarelli, L., and Muzzioli, S. (2018). The risk-asymmetry index as a new measure of risk. *Multinational Finance Journal*, **22**(3/4), 173–210.

Fama, E. F. (1965). The behavior of stock-market prices. *The Journal of Business*, **38**(1), 34–105.

Fisher, R. A. (1929). Moments and product moments of sampling distributions. *Proceedings of the London Mathematical Society: Series 2*, **30**(1), 199–238. Corrigenda in 30(1), 552–552.

Franceschini, C. and Loperfido, N. (2019). MaxSkew and MultiSkew: two R packages for detecting, measuring and removing multivariate skewness. *Symmetry*, **11**(8), 970.

Groeneveld, R. A. (1991a). An influence function approach to describing the skewness of a distribution. *The American Statistician*, **45**(2), 97–102.

Groeneveld, R. A. (1991b). Sharp inequalities for skewness measures. *Journal of the Royal Statistical Society: Series D (The Statistician)*, **40**(4), 387–392.

Groeneveld, R. A. and Meeden, G. (1984). Measuring skewness and kurtosis. *Journal of the Royal Statistical Society: Series D (The Statistician)*, **33**(4), 391–399.

Groeneveld, R. A. and Meeden, G. (2009). An improved skewness measure. *METRON*, **67**(3), 325–337.

Hinkley, D. V. (1975). On power transformations to symmetry. *Biometrika*, **62**(1), 101–111.

Holgersson, T. H. E. (2010). A modified skewness measure for testing asymmetry. *Communications in Statistics: Simulation and Computation*, **39**(2), 335–346.

Kashlak, A. B. (2018). Measuring distributional asymmetry with Wasserstein distance and Rademacher symmetrization. *Electronic Journal of Statistics*, **12**(2), 2091–2113.

Khattree, R. and Bahuguna, M. (2019). An alternative data analytic approach to measure the univariate and multivariate skewness. *International Journal of Data Science and Analytics*, **7**(1), 1–16.

Kirby, W. (1974). Algebraic boundedness of sample statistics. *Water Resources Research*, **10**(2), 220–222.

Lempérière, Y., Deremble, C., Nguyen, T.-T., Seager, P., Potters, M., and Bouchaud, J.-P. (2017). Risk premia: asymmetric tail risks and excess returns. *Quantitative Finance*, **17**(1), 1–14.

Li, X. and Morris, J. M. (1991). On measuring asymmetry and the reliability of the skewness measure. *Statistics & Probability Letters*, **12**(3), 267–271.

MacGillivray, H. L. (1986). Skewness and asymmetry: measures and orderings. *The Annals of Statistics*, **14**(3), 994–1011. Correction in The Annals of Statistics 1987, Vol. 15, No. 2, 884.

Muzzioli, S. (2013a). The forecasting performance of corridor implied volatility in the italian market. *Computational Economics*, **41**(3), 359–386.

Muzzioli, S. (2013b). The information content of option-based forecasts of volatility: evidence from the Italian stock market. *The Quarterly Journal of Finance*, **3**(01), 1350005.

Nolan, J. P. (2003). Modeling financial data with stable distributions. In S. T. Rachev, editor, *Handbook of Heavy Tailed Distributions in Finance*, pages 105–130. Elsevier.

Oja, H. (1981). On location, scale, skewness and kurtosis of univariate distributions. *Scandinavian Journal of Statistics*, **8**(3), 154–168.

Pearson, K. (1895). Contributions to the mathematical theory of evolution. II. Skew variation in homogeneous material. *Philosophical Transactions of the Royal Society of London: Series A*, **186**, 343–414.

Pearson, K. (1901). Mathematical contributions to the theory of evolution. X. Supplement to a memoir on skew variation. *Philosophical Transactions of the Royal Society of London: Series A*, **197**, 443–459.

Pearson, K. (1916). Mathematical contributions to the theory of evolution. XIX. Second supplement to a memoir on skew variation. *Philosophical Transactions of the Royal Society of London: Series A*, **216**, 429–457.

Peiró, A. (1999). Skewness in financial returns. *Journal of Banking & Finance*, **23**(6), 847 – 862.

Poncet, P. (2019). *modeest: Mode Estimation*. R package version 2.4.0.

R Core Team (2019). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna.

Sato, M. (1997). Some remarks on the mean, median, mode and skewness. *Australian Journal of Statistics*, **39**(2), 219–224.

Tajuddin, I. (1996). A simple measure of skewness. *Statistica Neerlandica*, **50**(3), 362–366.

Tajuddin, I. H. (1999). A comparison between two simple measures of skewness. *Journal of Applied Statistics*, **26**(6), 767–774.

van Zwet, W. R. (1964). Convex transformations: a new approach to skewness and kurtosis. *Statistica Neerlandica*, **18**(4), 433–441. Republished in S. van de Geer and M. Wegkamp, editors, Selected Works of Willem van Zwet, pages 3-11. Springer, New York, 2012.

van Zwet, W. R. (1979). Mean, median, mode II. *Statistica Neerlandica*, **33**(1), 1–5.

Whaley, R. E. (2000). The investor fear gauge. *The Journal of Portfolio Management*, **26**(3), 12–17.

Whaley, R. E. (2009). Understanding the vix. *The Journal of Portfolio Management*, **35**(3), 98–105.

Wilkins, J. E. (1944). A note on skewness and kurtosis. *The Annals of Mathematical Statistics*, **15**(3), 333–335.

Yule, G. U. (1911). *An Introduction to the Theory of Statistics*. Griffin, London.

Zwillinger, D. and Kokoska, S. (1999). *CRC Standard Probability and Statistics Tables and Formulae*. CRC Press, Boca Raton.