Proceedings of the 8th WSEAS International Conference on APPLIED MATHEMATICS, Tenerife, Spain, December 16-18, 2005 (pp90-94)

Comparing the kurtosis measures for symmetric-scale distribution functions considering a new kurtosis

Hedieh Jafarpour Rahman Farnoosh
Department of Mathematics
Islamic Azad University of Shiraz - University of Science and Technology
IRAN
http:/www.iaushiraz.ac.ir
http:/www.iust.ac.ir

Abstract: In this paper we consider the different kinds of kurtosis measures. We compare them for symmetric-scale distribution functions. We show the disadvantages of kurtosis measures then by introducing a kurtosis measure we modify the kurtosis measures. Finally we discuss the properties of introduced measure.

Keywords: Kurtosis measure, Scale functions, Symmetric distribution.

1 Introduction

It is the most popular that the kurtosis measure has been computed by:

$$\beta_2 = \frac{E(X-\mu)^4}{E^2(X-\mu)^2} \qquad (1)$$

where μ is the mean of the random variable X.

This is the fourth moment which is divided by square of the second moment. For a continuous random variable X with distribution function F which is symmetric we ca write formula (1) by:

$$\beta_2 = \frac{E(X)^4}{V(X)}$$

But β_2 is not a robust measure, because it is so sensitive to outliers. For example by generating 1000 data from a normal distribution β_2 is computed, 3.01, if we replace one of the generated value by an outlier value β_2 is computed 3.8, which is far from 3. We know that β_2 is equal to 3 for normal distribution.

β_2 does not measure only peaked ness of a distribution. It measures both peaked ness and tail weight of a distribution, so it does not sort the distribution based on the height of distributions each other. If second moment of distribution is infinite, β_2 does not exist for the distribution.

Balanda and McGillivray (1988) showed that for some distributions which have different shape from normal distribution, β_2 is almost 3, so $\beta_2 = 3$ is not a sufficient condition for normality and β_2 does not measure departure from normality.

β_2 is not a good measure for a special mixture distribution function which is introduced by Ali (1974).

$$F_k(x) = \left(1 - \frac{1}{k^2-1}\right)\Phi(x) + \frac{1}{k^2-1}\Phi\left(\frac{x}{k}\right)$$

where $\Phi(x)$ is normal distribution function and k=2,3….

This sequence converges in distribution to the standard normal distribution as $k \to \infty$, however $\beta_2 = \frac{3(k^2+1)}{4} \to \infty$.

Statistician tried to introduce another kurtosis measure which does not have the disadvantages.

Hogss (1974) introduced a kurtosis measure which measures of tail weight. For $0 < p \le 0.5$ define

$$\bar{U}_p(F) = \int_{F^{-1}(p)}^{\infty} x dF(x)$$

and

$$\bar{L}_p(F) = \int_{-\infty}^{F^{-1}(p)} x dF(x)$$

where F is the distribution function and $F^{-1}(p)$ is the pth quantile of F.

$$Q_p(F) = \frac{\bar{U}_p(F) - \bar{L}_p(F)}{\bar{U}_{0.5}(F) - \bar{L}_{0.5}(F)}$$

He used $Q_{0.2}(F)$. For symmetric F we have

$$Q_p(F) = \frac{\bar{U}_p(F)}{\bar{U}_{0.5}(F)}$$

$Q_p(F)$ is not a robust measure either, but it is not so sensitive as β_2

Ruppert (1987) construct a robust kurtosis measure. For $0 < p < \eta < 0.5$ he defined

$$R_{\eta,p}(F) = \frac{F^{-1}(1-p) - F^{-1}(p)}{F^{-1}(1-\eta) - F^{-1}(\eta)}$$

For symmetric F

$$R_{\eta,p}(F) = \frac{F^{-1}(1-p)}{F^{-1}(1-\eta)}$$

Groeneveld (1988) defined a kurtosis measure for symmetric distributions. For 0<p<0.5 we have:

$$\gamma_2(p,F) = \frac{F^{-1}\left(1-\frac{p}{2}\right) + F^{-1}\left(\frac{1+p}{2}\right) - 2F^{-1}(0.75)}{F^{-1}\left(1-\frac{p}{2}\right) - F^{-1}\left(\frac{1+p}{2}\right)}$$

2 Properties of the kurtosis measures

Oja (1981) called the invariant scale location functional T as a kurtosis measure if for two symmetric distribution function F and G which G has at least as kurtosis s as F, $F \le_s G$, we can conclude that $T(F) \le T(G)$.

Kurtosis measures have two following properties.

1) $T(aX + b) = T(X) \qquad a > 0$

2) $F \le_s G \Rightarrow T(F) \le T(G)$

Van Zwet (1964) introduced for the class of symmetric distributions an ordering defined by $F \le_s G$ iff $R_{F,G}(x) = G^{-1}(F(x))$ is convex for $x > m_F$ where m_F is the point of symmetry of F. Van Zwet ordered the following distributions as:

Uniform < laplace < Logistic < Laplace

β_2 is computed 0.0125, 3, 45.46 and 6 respectively. He showed that β_2 has the properties of kurtosis measure. We have to note that $\beta_2 > 3$ does not mean that the density function is higher than normal distribution. It may the density function has much mass in the tails rather than normal density function.

3 Introducing a modified kurtosis measure

β_2 is so sensitive with respect to the outliers, so we offer the following measure:

$$\beta_p^q(F) = \frac{E_F\left[(X - E_F(x)) I_{(F^{-1}(p), F^{-2}(1-q))}(x)\right]^4}{E^2{}_F\left[(X - E_F(x))^2 I_{(F^{-1}(p), F^{-2}(1-q))}(x)\right]}$$

where $F^{-1}(p), F^{-2}(1-q)$ are pth and qth quantiles of F which $X \sim F(.)$.

$\beta_p^q(F)$ is always finite and is symmetric distribution function q = 1-p.

First of all we show that the introduced measure have two properties. First property is intuitive. For showing the

second property, we show that $\frac{G^{-1}(\alpha)}{F^{-1}(\alpha)}$, $\alpha \in (0,1)$ is decreasing and if R(x) is convex then $\frac{R(x)}{x}$ is a non decreasing function of $x \neq 0$.
Without loss of generality we can take $m_F = 0$. In this paper we consider symmetric distribution, so the median of distributions are zero.

$$R(0) = G^{-1}(F(0)) = G^{-1}\left(\frac{1}{2}\right) = 0$$

For $x > 0$ we take the first derivative of $\frac{R(x)}{x}$. We obtain that $\frac{R(x)}{x} \leq R'(x)$. For showing the last equality we note that $R(0) = 0$.
By using the mean value theorem we have

$$R'(x_1) = \frac{R(x)}{x} \qquad 0 < x_1 < x$$

By the convexity of R(x), $R''(x) > 0$ and hence $R'(x)$ is non decreasing and

$$\frac{R(x)}{x} \leq R'(x_1) \leq R'(x)$$

So $\frac{R(x)}{x}$ is a non decreasing for x>0.
A similar proof holds for x<0.
So

$$\frac{G^{-1}(F(F^{-1}(\alpha)))}{F^{-1}(\alpha)} = \frac{G^{-1}(\alpha)}{F^{-1}(\alpha)} \qquad \alpha \in (0,1)$$

is non decreasing.
For proving second property we have to show that

$$\beta_{F^{-1}(p)}^{F^{-1}(q)}(F) \leq \beta_{G^{-1}(p)}^{G^{-1}(q)}(G)$$

It is sufficient to show that

$$\frac{E_F\left(X^4 I_{(F^{-1}(p),F^{-1}(q))}(X)\right)}{\left(E_F\left(X^2 I_{(F^{-1}(p),F^{-1}(q))}(X)\right)\right)^2} \leq \frac{E_G\left(X^4{}_{(G^{-1}(p),G^{-1}(q))} I(X)\right)}{\left(E_G\left(X^2 I_{(G^{-1}(p),G^{-1}(q))}(X)\right)\right)^2}$$

We can take $x = F^{-1}(p)$.
So

$$E_F\left(X^4 I_{(F^{-1}(p),F^{-1}(q))}(X)\right)$$
$$= \int_{F^{-1}(p)}^{F^{-1}(q)} x^4 dF(x)$$
$$= \int_p^q \left(F^{-1}(u)\right)^4 d(x)$$

The last inequality is equivalent to:

$$\frac{\int_p^q \left(F^{-1}(u)\right)^4 d(u)}{\int_p^q \left(G^{-1}(u)\right)^4 d(u)} \leq \frac{\left(\int_p^q \left(F^{-1}(u)\right)^2\right)^2}{\left(\int_p^q \left(G^{-1}(u)\right)^2\right)^2}$$

By using the mean value theorem the last inequality is replaced by

$$\frac{\left(F^{-1}(u_2)\right)^4}{\left(G^{-1}(u_2)\right)^4} \leq \frac{\left(\left(F^{-1}(u_1)\right)\right)^2}{\left(\left(G^{-1}(u_1)\right)\right)^2} \qquad 0 < u_1 < u_2 < 1$$

Because we know that $\frac{F^{-1}(\alpha)}{G^{-1}(\alpha)}$ is non decreasing when $F \leq_s G$ the above inequality is hold.

Proposition :
The kurtosis measure $\beta_p^q(F_k)$ for F_k, which has been introduced by Ali (1974), converges to $\beta_p^q(\Phi_k)$.

Proof:

$$\beta_p^q(F_k(x)) = \frac{\left(1 - \frac{1}{k^2 - 1}\right)\int_p^q x^4 \phi(x) dx + A(k)}{\left(\left(1 - \frac{1}{k^2 - 1}\right)\int_p^q x^2 \phi(x) dx + B(k)\right)^2}$$

where

$$A(k) = \frac{1}{k(k^2 - 1)}\int_p^q x^4 \phi\left(\frac{x}{k}\right) dx$$
$$= \frac{k^5}{k(k^2 - 1)}\int_{\frac{q}{k}}^{\frac{p}{k}} x^4 \phi(x) dx$$

and

$$B(k) = \frac{1}{k(k^2 - 1)}\int_p^q x^2 \phi\left(\frac{x}{k}\right) dx$$

$$= \frac{k^3}{k(k^2-1)} \int_{\frac{q}{k}}^{\frac{p}{k}} x^2 \phi(x) dx$$

Note that

$$0 \le \lim_{k\to\infty} A(k) \le \lim_{k\to\infty} \frac{k^5}{k(k^2-1)} \int_{\frac{q}{k}}^{\frac{p}{k}} x^4 dx$$

$$= \lim_{k\to\infty} \frac{k^5}{k^6(k^2-1)} \left(q^5 - p^5\right) = 0$$

Similarly we have $\lim_{k\to\infty} B(k) = 0$. So

$$\lim_{k\to\infty} \beta_p^q (F_k(x))$$

$$= \lim_{k\to\infty} \frac{\left(1 - \frac{1}{k^2-1}\right)\int_p^q x^4 \phi(x) dx + A(k)}{\left(\left(1 - \frac{1}{k^2-1}\right)\int_p^q x^2 \phi(x) dx + B(k)\right)^2}$$

$$= \frac{\int_p^q x^4 \phi(x) dx}{\left(\int x^2 \phi(x) dx\right)^2}$$

$$= \beta_p^q(\phi(x)).$$

4 Difference between β_2 and standardized fourth moment

we consider the standardized fourth central moment given by

$$\gamma_2 = E\left(\frac{X-\alpha}{\beta}\right)^4$$

As a new measure of kurtosis, where α and $\beta > 0$ are location and scale parameters of the distribution.

Two kurtosis measure are called shape parameter that measures peaked- ness and tailed ness of distributions.

The denominator of β_2 is fourth power of standard deviation, but the denominator of γ_2 is the fourth power of scale parameter. If the scale parameter, β, and the standard deviation, σ, are equal in addition if the location parameter, α, and the mean, μ, are equal then $\beta_2 = \gamma_2$.

For the class of family of normal distributions β_2 and γ_2 are equal, because α and β in γ_2 are mean and standard deviation.

We note that for the family of normal distribution $\beta_2 = \gamma_2 = 3$.

Not only for standardized normal distribution but also for any desirable μ and σ in normal distribution we have $\beta_2 = \gamma_2 = 3$.

Note that location and square of scale parameters are different from mean and variance. Standard deviation is one of the estimators of scale parameter.

For example for standard Laplace distribution function $\beta_2 = 6$ and $\gamma_2 = 24$. Because

$$f_X(x) = \frac{1}{2} e^{-|x|} \qquad -\infty < x < \infty$$

We know that $\alpha = 0$, $\beta = 1$, $\mu = 0$, $\sigma = \sqrt{2}$.

So,

$$\gamma_2 = E(X^4) = 24$$

$$\beta_2 = \frac{E(X^4)}{E^2(X^2)} = \frac{E(X^4)}{E(X^2) + \mathrm{var}(X)} = \frac{24}{4} = 6$$

So in general:

$$\frac{E(X-\mu)^4}{E^2(X-\mu)^2} \ne E\left(\frac{X-\alpha}{\beta}\right)^4$$

5 Concluson

In this paper the kurtosis measure has been reviewed. The properties of kurtosis measure have been considered. The usual kurtosis measure has been modified and the properties of a kurtosis measure have been proved for this kurtosis measure. The introduced kurtosis measure does not have the disadvantages of the other kurtosis measures especially it works well for the mixture normal distribution function which has been introduced by Ali. The difference of two types of kurtosis measures β_2 and γ_2 has been shown. If the mean and the standard deviation be the location and scale parameter then $\beta_2 = \gamma_2$, otherwise $\beta_2 \neq \gamma_2$.

References:

[1] Ali, M.M, Stochastic Ordering and Kurtosis Measure, JASA, Vol.69, No.346, 1974, pp.543-545.

[2] Brys, Hubert, M and Stryyf, A, Robust measure of tail weight, Computational Statistics and Data Analysis, 2004, pp.1-27.

[3] Groeneveld, R. A., A class of Quantile measure for kurtosis, The American Statistician, Vol.24, 1970, pp.19-22.

[4] Groeneveld, R. A. and Meeden, G., Measuring skewness and kurtosis, The Statistician, Vol. 33, 1984, pp.391-399.

[5] Kevin P. Balanda and H. L. MacGillivray, Kurtosis: a critical review, JASA, Vol.42, No.2, 1988, pp.111-119.

[6] Oja, H. On location, scale, skewness and kurtosis of univariate distributions, Scand J Statist, Vol.8, 1981, pp.154-168.

[7] Ruppert, D, What is Kurtosis?, The American Statistician, Vol.41,1987, pp.1-5.

[8] Van Zwet, W.R., Convex transformation of random variables, Mathematical Centrum, Amsterdam, 1964.