

On Probability Density (f) based New Measure of Kurtosis K_f

Shivagaje Ashok¹ Narkhede Vijay² Kasture Madhukar³

¹Dept. of Statistics, MPKV, Rahuri

²Joint Director, Higher Education, Govt. of Maharashtra, Pune

³Post Graduate Dept. of Statistics, NAC&SC College Ahmednagar

Abstract: It is known that $N(0,1)$ has wide range of real life applications, whereas, Cauchy distribution is used mainly for theoretical studies with rare applications. Normal distribution has many interesting properties and its moment based kurtosis is treated as standard to compare kurtosis of other distribution. On the contrary moments for Cauchy distribution does not exist and hence its moment based kurtosis is undefined. This paper proposes a probability density (f) based new measure of kurtosis (K_f) and is further applied to compare kurtosis of $N(0,1)$ with $C(0,1)$, and $N(0, \lambda)$ with $C(0, \lambda)$

Keywords: Probability density, Distribution Function, Kurtosis, Newton Raphson method

I. Introduction

It is well known that measure of central tendency, measure of dispersion, measure of skewness and measure of kurtosis are the basic statistical tools often used to describe the nature of distribution. Measure of central tendency gives the information about location, measure of dispersion throws light on variation or the spread of the data. On the other hand symmetry of the distribution can be captured by measures of skewness, whereas peakedness of the distribution relative to the normal distribution is understood by measure of kurtosis. Usually first four moments describes the nature of the distribution and to compare two distributions. There are variety of measures of central tendencies and measures of dispersion. Further, Karl Perasons coefficient of skewness, Bowleys coefficient of Skewness and measures of skewness based on moments are options to understand the skewness of distribution. Unfortunately, measure of kurtosis based on moments is the only measure found in literature. Therefore, if any distribution does not have moments then it is very difficult to understand the kurtosis of such distribution. Similarly, measure of kurtosis β_2 is more justified if underlying distribution is near symmetry. Lawrence T. DeCarlo (1997) described various aspects of β_2 . The author further pointed out three common errors concerning kurtosis; (a) kurtosis is defined solely in terms of peakedness, with no mention of importance of the tails; (b) the relation between peak and tails of a distribution with excess kurtosis is described or illustrated incorrectly and (c) descriptions and illustrations fail to distinguish between kurtosis and the variance. David C. Blest (2003) suggested a new measure of kurtosis from which the effects of any skewness are deleted, which allows comparison of distribution on the basis of kurtosis. J.F.Rosco etal.(2011) reconsidered the derivation of Blest's measure of kurtosis adjusted for skewness. The author further proposed an adoption of it which generally eliminates the effects of asymmetry more successfully. The paper also gives lower bounds for the two skewness adjusted kurtosis measures using Sinh-arcsinh transformation.

The common feature of $N(0,1)$ and $C(0,1)$ is that both are symmetric about zero. The discriminating feature of the two distributions is existence of moments and defining kurtosis. As, theoretically it is well established that moments for Cauchy distribution does not exist hence measures of kurtosis based on moments is undefined. It leads to difficulty in comparing kurtosis of Cauchy distribution with other distribution in general and Normal distribution in particular.

In this paper, new measure of kurtosis K_f based on probability density (f) is defined. The most important feature of the newly proposed measure of kurtosis K_f is that it does not require moments. It is based on Distribution Function and is suitable for any distribution. Therefore, one can compare kurtosis of any two distributions by using proposed measure of kurtosis K_f .

Measure of kurtosis based on Probability Density Function

Let $f(x)$ is the probability density function (pdf) of a random variable X. Figure-1 shows the graph of pdf $f(x)$.

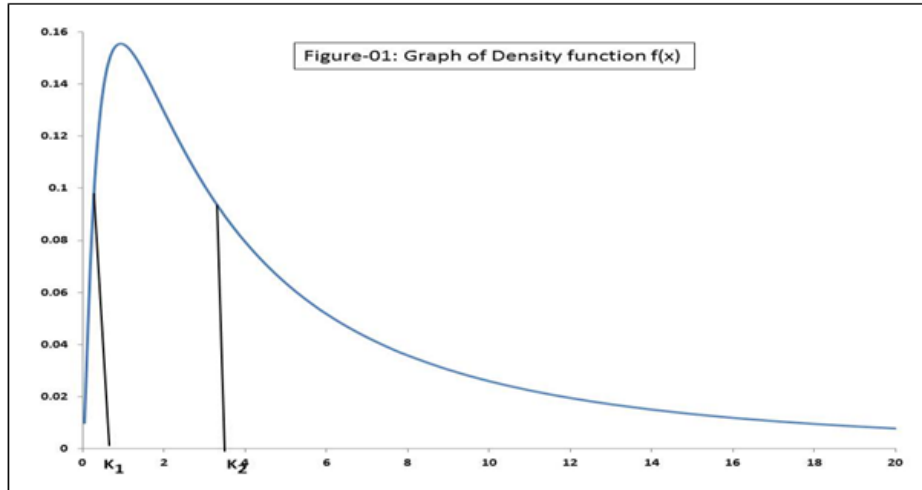


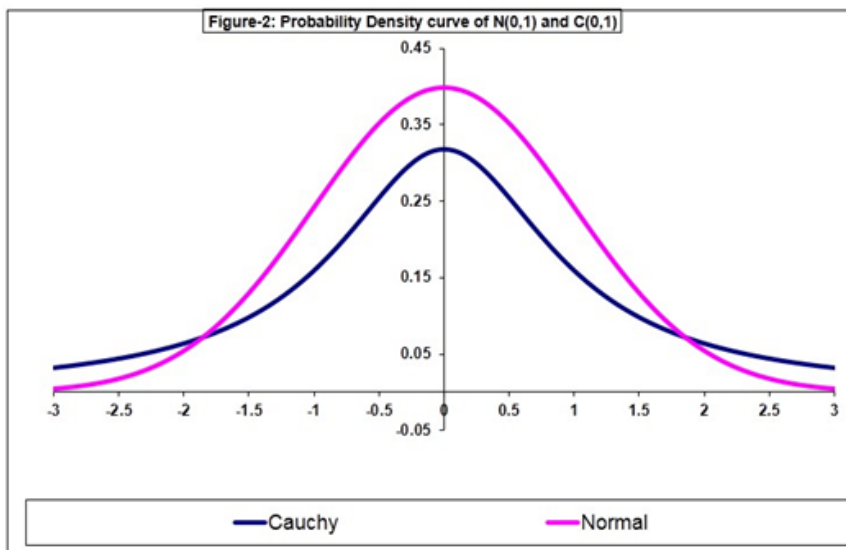
Figure-1 shows peak of the distribution between k_1 and k_2 . The intuition suggest that if distribution has high peak then probability density in the peaked portion will be high, and if distribution has low peak then probability density in the peaked portion will be less. By using this simple intuition, the measure of kurtosis based on Probability Density Function is proposed as:

$$K_f = \frac{F(k_2) - F(k_1)}{k_2 - k_1}$$

Where $F(\cdot)$ denotes Distribution Function corresponding to pdf $f(x)$. K_f can also be interpreted as ‘per unit probability density’. It is remarkable to note that K_f is directly proportional to the probability density covered in the interval $[k_1, k_2]$ and is inversely proportional to the length of the interval $(k_2 - k_1)$. It is further consistent with the thought expressed by Balanda and MacGillivray(1988).

Comparison of Kurtosis using K_f of $N(0,1)$ and $C(0,1)$

In literature $N(0,1)$ and $C(0,1)$ are Standard Normal and Standard Cauchy distributions symmetric about zero. The Figure-2 depicts the $N(0,1)$ and $C(0,1)$ curve intersect at two points and curve of $N(0,1)$ is above the curve of $C(0,1)$.



The graphical order of the two underlying distributions indicate that $C(0,1)$ is flat as compared to $N(0,1)$. The β_2 of normal distribution is 3 whereas that of $C(0,1)$ is undefined. Therefore, Kurtosis of $N(0,1)$ and $C(0,1)$ can be compared by using K_f .

In computing K_f the first step is to determine the x co-ordinate of two points of intersection k_1 and k_2 of two underlying distributions. In this case $k_1=-k_2=k$ due to symmetry. Consider $f(x)$ as density of $N(0,1)$ and $g(x)$ as density of $C(0,1)$. Then point of intersection of $N(0,1)$ and $C(0,1)$ is the solution of the equation $f(x)=g(x)$.

Therefore,
$$\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} = \frac{1}{\pi} \frac{1}{1+x^2} \dots\dots\dots(1)$$

Simplification of Equation (1) is,

$$x^2 - 2L_n(1+x^2) + L_n\left(\frac{7}{11}\right) = 0 \dots\dots\dots(2)$$

Equation (2) is a transcendental equation which can be solved by using Newton Raphson method

Let us define Equation (2) as

$$h(x) = 0 \dots\dots\dots(3)$$

where
$$h(x) = x^2 - 2L_n(1+x^2) + L_n\left(\frac{7}{11}\right) \dots\dots\dots(4)$$

on differentiating Equation (4)

$$h'(x) = 2x - \frac{4x}{1+x^2} \dots\dots\dots(5)$$

If $x = x_0$ is treated as initial root of Equation (3) then by applying Newton Raphson method the improved solution is given by

$$x_n = x_{n-1} - \frac{h(x_{n-1})}{h'(x_{n-1})} \quad n=1,2,\dots \dots\dots(6)$$

Figure-2, justifies the initial root $x_0 = -1$ in seven iterations the solution of Equation (3) is $x = -1.85142735926976$ produces an error of $7.82005743493741E-09$.

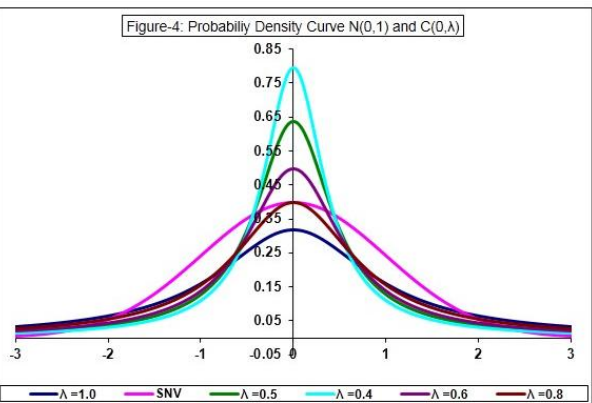
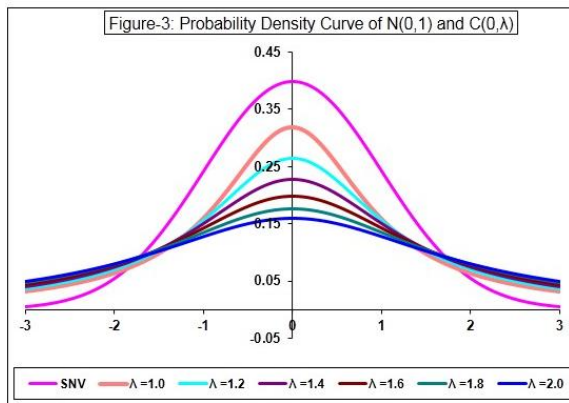
Thus $k = -1.85142735926976$ and the value of K_f is 0.2527 and K_g is 0.1848 showing more per unit probability density of $N(0,1)$ as compared to $C(0,1)$, hence $C(0,1)$ is flat as compared to $N(0,1)$.

Comparison of Kurtosis using K_f of $N(0,1)$ and $C(0,\lambda)$

This section describes the comparative account of kurtosis of $N(0,1)$ and $C(0,\lambda)$ for various values of λ . Figure-3 and Figure-4 elaborates probability density curve of $N(0,1)$ and $C(0,\lambda)$ for $\lambda > 1$ and $\lambda < 1$ respectively.

It is seen from Figure-3 and Figure-4 that peakedness of $C(0,\lambda)$ increases as λ decreases i.e. as λ increases $C(0,\lambda)$ becomes more and more flatter.

This section will provide the values of K_f for $C(0,\lambda)$ for different values of λ and are compared with that of



$N(0,1)$. To determine the point of intersection of $N(0,1)$ and $C(0,\lambda)$ one has to solve ,

$$\frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2} \frac{x^2}{2}\right\} = \frac{\lambda}{\pi} \frac{1}{\lambda^2 + x^2} \dots\dots\dots(7)$$

To determine the point of intersection of $N(0,1)$ and $C(0,\lambda)$ by Newton Raphson method, Equation (7) can be simplified as,

$$x^2 = 2L_n(\lambda^2 + x^2) - 2L_n(\lambda) - L_n\left(\frac{7}{11}\right) \dots\dots\dots(8)$$

Which further gives $u(x)$ and $u'(x)$ as,

$$u(x) = x^2 - 2\log(\lambda^2 + x^2) + 2\log \lambda + \log\left(\frac{7}{11}\right) = 0 \dots\dots\dots(9)$$

$$u'(x) = 2x - \frac{4x}{\lambda^2 + x^2} \dots\dots\dots(10)$$

Then by taking x_0 as initial root the further improved roots are given by,

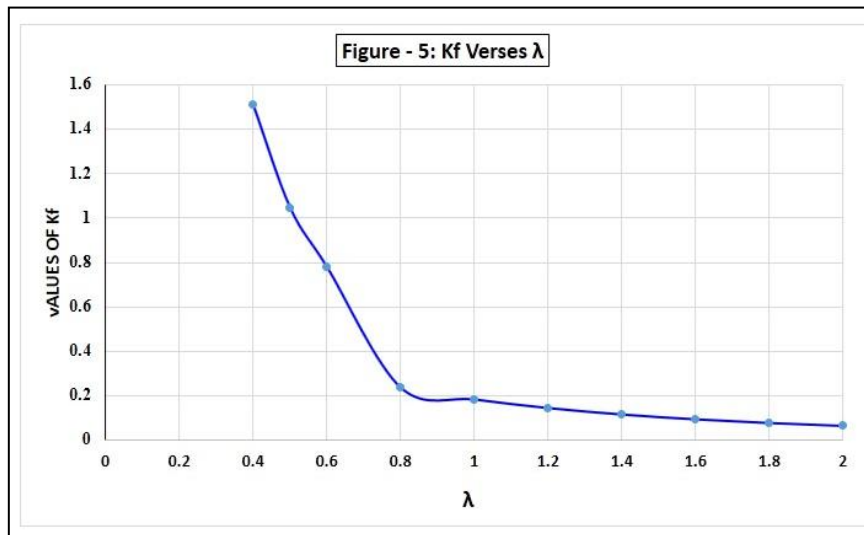
$$x_1 = x_0 - \left(\frac{x_0^2 - 2\log(\lambda^2 + x_0^2) + 2\log \lambda + \log\left(\frac{7}{11}\right)}{2x_0 - \left(\frac{4x_0}{\lambda^2 + x_0^2}\right)} \right) \dots\dots\dots(11)$$

On applying the formula for kurtosis K_f , kurtosis can be computed and it helps in comparison. Table-1 gives the summary of point of intersection and values of K_f for $N(0,1)$ and $C(0,\lambda)$ for different values of λ .

Table-1: Summary

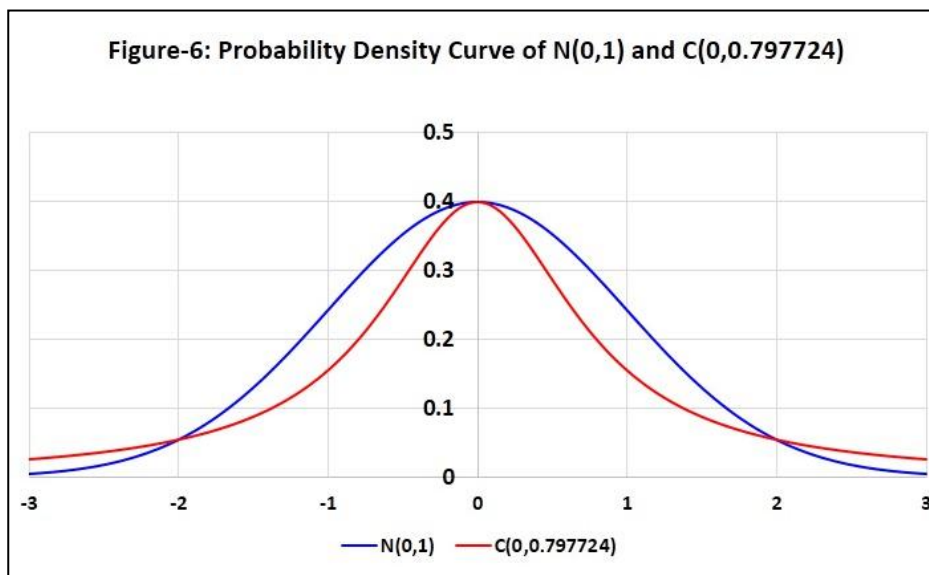
λ	0.4	0.5	0.6	0.8	1	1.2	1.4	1.6	1.8	2
K	0.4371	0.4338	0.3975	0.0054	1.8514	1.7674	1.7232	1.7061	1.7056	1.7148
K_f	1.5099	1.0483	0.7806	0.2379	0.1848	0.1462	0.1172	0.0953	0.0786	0.0658

Table-1 gives the quantification of measure of kurtosis based on the interpretation per unit probability density i.e K_f . It is clear to note that for $\lambda < 1$ Cauchy distribution has high peak than $N(0,1)$ showing excess per unit probability density. For $\lambda \geq 1$ $C(0,\lambda)$ seems to have low peak than $N(0,1)$. Further increase in λ shows increase in excess of per unit probability density as compared to $N(0,1)$. The Figure-5 gives the decreasing trend of K_f for increasing in λ .



Is Kurtosis means peakedness.....?

Figure-6 depicts probability density curve of $N(0,1)$ and $C(0,0.797724)$. The important feature of this figure is that both densities are symmetric about zero, with mode zero. Similarly, these two densities intersect at three points, out of which the point $(0, 0.3988)$ is the center point indicating maximum value of both density functions at $x=0$ is 0.3988. As height of both densities or peakedness of both densities is same, does it indicate same kurtosis of both densities?. The other two x co-ordinates of points of intersection are ± 1.987525 . From Figure-6 it is further clear that the probability covered by both densities between the points -1.987525 to 1.987525 is different, indicating that the two densities differ in kurtosis. The K_f values for $N(0,1)$ is 0.2527 and $C(0,0.797724)$ is 0.2386. Thus among the two underlying distribution $N(0,1)$ has high kurtosis than $C(0,0.797724)$. Thus the proposed measure of kurtosis K_f has ability to capture the distinction between the kurtosis of two distributions reflecting same peakedness.



II. Conclusion

Thus in conclusion it can be stated that the proposed measure of kurtosis K_f based on per unit probability density is a powerful measure applicable to all distributions and hence will facilitate the comparison of kurtosis of distributions if moments does not exists. Further, K_f is best suit measure of kurtosis discriminating two distributions with same peakedness. In addition K_f based on probability density covered in middle part of the distribution indirectly gives a focus on tail probabilities and hence it provides a good remedy to overcome the common error (a) mentioned by Lawrence T. DeCarlo (1997).

References

- [1]. Kelvin P. Balanda and H.L. MacGillivray. "Kurtosis: A Critical Review". The American Statistician 42:2, May 1988, Pp 111-119.
- [2]. Lawrence T. DeCarlo. "On the meaning and use of kurtosis", Psychological methods, 2 : 3, 1997, Pp 292-307.
- [3]. David C. Blest. "A new measure of kurtosis adjusted for skewness", Aust. N. Z. J. Stat. 45:2, 2003, Pp 175-179.
- [4]. Rosco J.F., Arthur Pewsey and Jones M.C., "On Blest's measure of kurtosis adjusted for skewness", preprint submitted to Computational Statistics and Data Analysis, 2011.