

The Standard Fourth Moment Coefficient of Kurtosis and Its Influence Function: An Early Intuition by L. Faleschini

Il Momento Quarto Standardizzato Come Indice di Curtosi e la Sua Curva di Influenza: Un'Intuizione di L. Faleschini

Anna Maria Fiori, Michele Zenga

Dipartimento di Metodi Quantitativi per le Scienze Economiche ed Aziendali,
Università di Milano-Bicocca, e-mail: anna.fiori@unimib.it

Riassunto: Che cosa misura la curtosi? Quali trasformazioni permettono di conseguire aumenti o riduzioni della curtosi? In un pionieristico lavoro del 1948, L. Faleschini rispondeva a questi interrogativi calcolando la derivata parziale del coefficiente di curtosi β_2 rispetto ad una frequenza puntuale assoluta. Oltre a rivalutare la portata della “derivata di Faleschini”, il presente lavoro ne esplicita un’inattesa identità formale con il concetto di curva di influenza, coniato da Hampel nel 1968 ed applicato per la prima volta a β_2 da Ruppert soltanto nel 1987.

Keywords: kurtosis, peakedness, tail behaviour, influence function.

1. Introduction

The oldest and most common measure of kurtosis is the standard fourth moment coefficient:

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = E\left(\frac{X - \mu}{\sigma}\right)^4 \quad (1)$$

which was introduced by K. Pearson in 1905 (Pearson, 1905). There has been much discussion over whether β_2 measures peakedness, tail weight or the “shoulders” of a distribution. A convenient way of grasping the meaning of β_2 is to ask how β_2 changes as observations are added to an existing distribution. To our knowledge, the first author who answered this question was L. Faleschini in 1948.

This work rediscovers Faleschini’s methodology and shows how it can be used to assess the relative importance of tail heaviness and peakedness in determining β_2 . An unexpected connection is established between a partial derivative computed by Faleschini and the influence function of β_2 , which appeared much later in statistical literature (Hampel, 1968 and 1974; Ruppert, 1987).

2. From Faleschini’s derivative...

In his 1948 paper, Faleschini considered a statistical variable X consisting of n real values, $a_1, a_2, \dots, a_n$, and the corresponding frequencies $f_1, f_2, \dots, f_n$. The total size of the distribution was denoted by $P = \sum_{r=1}^n f_r$.

To observe how β_2 is changed by observations added at some particular point, say a_r , Faleschini computed the partial derivative of β_2 with respect to the altered frequency f_r . His fundamental result, which is restricted to symmetric distributions, is sketched in Table 1.

Table 1: *Outline of Faleschini's methodology.*

BACKGROUND	The kurtosis coefficient β_2 is written as the moment ratio: $\beta_2 = \frac{\mu_4}{\mu_2^2}$. The s-th central moment of X is represented by: $\mu_s = \frac{1}{P} \sum_{r=1}^n (a_r - \mu)^s = \sum_{i=0}^s (-1)^i \binom{s}{i} m_{s-i} \cdot \mu^i, \quad (2)$ where: $m_{s-i} = \frac{1}{P} \sum_{r=1}^n a_r^{s-i} \cdot f_r$ and μ stands for the arithmetic mean.
QUESTION	«We want to investigate the behaviour of β_2 when a frequency f_r is altered».
PROPOSED METHODOLOGY	Using the combinatorial formula (2) for central moments, Faleschini computed the partial derivative of μ_s with respect to f_r: $\frac{\partial \mu_s}{\partial f_r} = \sum_{i=0}^s (-1)^i \binom{s}{i} \frac{\partial (m_{s-i} \cdot \mu^i)}{\partial f_r} = \frac{1}{P} [(a_r - \mu)^s - \mu_s - s \mu_{s-1} (a_r - \mu)]$ where: $\frac{\partial m_{s-i}}{\partial f_r} = \frac{1}{P} (a_r^{s-i} - m_{s-i}) \quad \text{and:} \quad \frac{\partial \mu^i}{\partial f_r} = \frac{1}{P} i \mu^{i-1} (a_r - \mu)$ He then found the partial derivative of β_2 with respect to f_r: $\frac{\partial \beta_2}{\partial f_r} = \frac{\partial (\mu^4 / \mu_2^2)}{\partial f_r} = \frac{1}{P} \cdot \left\{ \left[\left(\frac{a_r - \mu}{\sigma} \right)^2 - \beta_2 \right]^2 - \beta_2 (\beta_2 - 1) \right\}$

Solving the quartic equation: $\partial \beta_2 / \partial f_r = 0$, Faleschini obtained the four roots:

$$a_r = \mu \pm \sigma \sqrt{\beta_2 \pm \sqrt{\beta_2(\beta_2 - 1)}} \quad (3)$$

which are always real owing to the inequality: $\beta_2 \geq 1$ (cf. Stuart and Ord, 1994). He could therefore decompose the range of the statistical variable X into five regions according to the sign of $\partial \beta_2 / \partial f_r$:

$$\text{Region I:} \quad a_1 \leftrightarrow \mu - \sigma \sqrt{\beta_2 + \sqrt{\beta_2(\beta_2 - 1)}} \quad \partial \beta_2 / \partial f_r \text{ is positive}$$

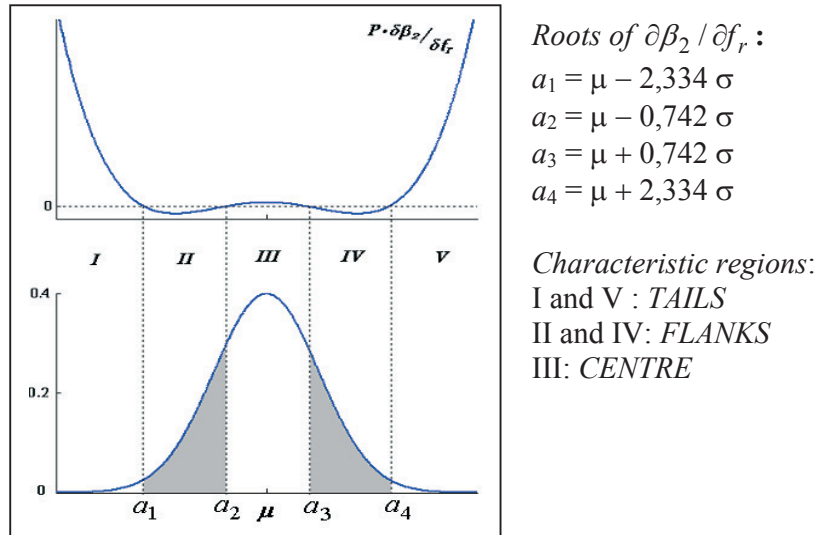
$$\text{Region II:} \quad \leftrightarrow \mu - \sigma \sqrt{\beta_2 - \sqrt{\beta_2(\beta_2 - 1)}} \quad \partial \beta_2 / \partial f_r \text{ is negative}$$

Region III:	$\leftrightarrow \mu + \sigma \sqrt{\beta_2 - \sqrt{\beta_2(\beta_2 - 1)}}$	$\partial\beta_2 / \partial f_r$ is positive
Region IV:	$\leftrightarrow \mu + \sigma \sqrt{\beta_2 + \sqrt{\beta_2(\beta_2 - 1)}}$	$\partial\beta_2 / \partial f_r$ is negative
Region V:	$\leftrightarrow a_n$	$\partial\beta_2 / \partial f_r$ is positive

Symmetrically-added observations within regions *II* and *IV* lower β_2 , those outside raise it. Regions *II* and *IV* can therefore be identified with the *flanks*, region *III* with the *centre*, regions *I* and *V* with the *tails*. Since $\partial\beta_2 / \partial f_r$ is of order a_r^4 , it grows very rapidly as a_r decreases below $\mu - \sigma \sqrt{\beta_2 + \sqrt{\beta_2(\beta_2 - 1)}}$ or increases beyond $\mu + \sigma \sqrt{\beta_2 + \sqrt{\beta_2(\beta_2 - 1)}}$. Hence it is clear that β_2 is primarily a measure of tail behaviour and only to a lesser extent of peakedness.

To investigate how β_2 relates to non-normality, Figure 1 graphs Faleschini's derivative for a standard normal distribution ($\beta_2 = 3$). Elaborating on Faleschini's discussion, "we can therefore conclude that a symmetric curve with $\beta_2 > 3$ will be *iperbinomial* (i.e. *leptokurtic*) because some values at the centre and in the tails will have higher frequency than in the normal curve with the same mean and variance, and some values in the flanks will be underweighted compared to the corresponding normal curve. The situation is reversed for distributions with $\beta_2 < 3$ " (Faleschini, 1948).

Figure 1: Faleschini's derivative for a standard normal distribution, with indication of the ranges corresponding to the centre, flanks and tails.



3. ... to the influence function of β_2

Following the lines of Hampel (1968, 1974), Ruppert (1987) suggested how the influence function (*IF*) might be used to assess the properties of a kurtosis measure. His approach is outlined in Table 2, focusing on a continuous symmetric random variable X with distribution function F .

Table 2: Outline of the influence function approach to β_2 for a one-point mass “contaminant” at the value x .

BACK- GROUND	The kurtosis coefficient β_2 is viewed as the functional on the set of distribution functions with finite fourth moment defined by: $\beta_2 = \beta_2(F) = \frac{\mu_4(F)}{[\mu_2(F)]^2} \quad \text{where: } \mu_r(F) = \int_{-\infty}^{\infty} (x - \mu)^r dF$
QUES- TION	«How does β_2 change if we throw in an additional observation at some point x ?»
PROPOSED METHODOLOGY	For δ_x a point mass at the value x, the “contaminated” distribution: $F_\varepsilon = (1 - \varepsilon) F + \varepsilon \delta_x \quad (0 \leq \varepsilon \leq 1)$ is defined. The total change in β_2, computed as: $\beta_2(F_\varepsilon) - \beta_2(F)$, is “standardized” by ε, the proportion of contamination. Then: $\lim_{\varepsilon \downarrow 0} \frac{\beta_2(F_\varepsilon) - \beta_2(F)}{\varepsilon} = \left[\left(\frac{x - \mu}{\sigma} \right)^2 - \beta_2 \right]^2 - \beta_2(\beta_2 - 1) = IF(\beta_2)$ is the derivative of β_2 at F in the direction toward δ_x and is called the influence function of β_2.

We notice here that the IF of β_2 differs from Faleschini’s derivative by minor deviations only (x replaces a_r and the $1/P$ factor disappears). Although Faleschini moved from different assumptions (a statistical rather than a continuous random variable), he obtained in practice the same result as Ruppert. Faleschini’s work can thus be seen as the first intuition of the IF of β_2 , as well as an accessible approach to its interpretation.

References

- Faleschini L. (1948) Su Alcune Proprietà dei Momenti Impiegati nello Studio della Variabilità, Asimmetria e Curtosi, *Statistica*, 8, 503-513.
- Hampel F. (1968) *Contributions to the Theory of Robust Estimation*, Unpublished Ph.D. dissertation, University of California, Berkeley, Dept. of Statistics.
- Hampel F. (1974) The Influence Curve and Its Role in Robust Estimation, *Journal of the American Statistical Association*, 69, 383-393.
- Pearson K. (1905) Skew Variation, A Rejoinder, *Biometrika*, 4, 169-212.
- Ruppert D. (1987) What is Kurtosis? An Influence Function Approach, *The American Statistician*, 41, 1-5.
- Stuart A., Ord J. K. (1994) *Kendall's Advanced Theory of Statistics, Volume I: Distribution Theory*, London: Edward Arnold.