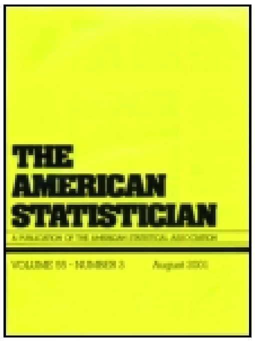

The American Statistician

ISSN: 0003-1305 (Print) 1537-2731 (Online) Journal homepage: http://www.tandfonline.com/loi/utas20

The Complications of the Fourth Central Moment

Yadolah Dodge & Valentin Rousson

To cite this article: Yadolah Dodge & Valentin Rousson (1999) The Complications of the Fourth Central Moment, The American Statistician, 53:3, 267-269

To link to this article: http://dx.doi.org/10.1080/00031305.1999.10474471

Teacher's Corner

The Complications of the Fourth Central Moment

Yadolah DODGE and Valentin ROUSSON

This article illustrates the additional complications in the mathematical formulas involving the fourth central moment in comparison with those involving the first moment, the second central moment, and the third central moment.

KEY WORDS: Central moments; Cumulants; Descriptive statistics; Moments; Unbiased estimation.

1. INTRODUCTION

In an introductory statistics course, *moments* and *central moments* of a random variable are interpreted descriptively. The first moment (the mean) and the second central moment (the variance) are introduced to describe, respectively, the location and the dispersion of a random variable. Then the third and the fourth central moments are said to be in turn the skewness and the kurtosis of the random variable. A natural and legitimate question that students ask at this point is "What about the higher order central moments? What do they describe? Why do we not go beyond the fourth one?" The fact is that the fifth, the sixth, and other higher order central moments have not yet received any descriptive interpretation in the statistical literature and one can wonder why.

The absence of interpretation for high order central moments is perhaps simply due to the mathematical complications involved. In this article we present some mathematical formulas dealing with moments and central moments. Looking at these formulas we come to the conclusion that mathematical complications arise not with the fifth, but already with the fourth central moment. In comparison, formulas involving the first moment, the second central moment, and the third central moment are beautifully simple and certainly give joy to teachers and graduates of statistics.

Yadolah Dodge is Professor of Statistics and Operations Research, and Valentin Rousson is Research Assistant, Statistics Group, University of Neuchâtel, Case Postal 1825, CH-2002 Neuchâtel, Switzerland (Email: yadolah.dodge@seco.unine.ch). The authors thank Sir David Cox, Joe Whittaker, the Editor, and the two referees for their careful reading of the manuscript and for their valuable comments that helped to improve the article.

2. MOMENTS AND CUMULANTS

Let X be a real valued random variable with distribution F. The first moment μ of X is defined by

$$\mu = \mu(X) = \int_{-\infty}^{+\infty} x dF(x).$$

Similarly the moments of higher order μ_r' $(r \geq 2)$ are defined by

$$\mu_r' = \mu_r'(X) = \int_{-\infty}^{+\infty} x^r dF(x)$$

and the central moments μ_r $(r \geq 1)$ are defined by

$$\mu_r = \mu_r(X) = \int_{-\infty}^{+\infty} (x-\mu)^r dF(x).$$

The first central moment μ_1 is always equal to zero and as a result is not an interesting concept. This is the reason why the central moments μ_r $(r \geq 2)$ are often "associated" with the first moment μ. Thus, in what follows, we shall simply call *moments* the set of constants constituted by the first moment μ and by the central moments μ_r $(r \geq 2)$.

Another set of constants of interest are the *cumulants*. Formally the cumulants $\kappa_r = \kappa_r(X)$ $(r \geq 1)$ of X are defined by the identity in t

$$\exp\left\{\kappa_1 t + \kappa_2 \frac{t^2}{2!} + \cdots + \kappa_r \frac{t^r}{r!} + \ldots\right\} = 1 + \mu t + \mu_2' \frac{t^2}{2!} + \cdots + \mu_r' \frac{t^r}{r!} + \cdots$$

If $\phi(t) = \mu(\exp(itX))$ is the characteristic function of X, we can also define the cumulants κ_r to be the coefficient of $(it)^r/r!$ in the expansion in power series of $\log(\phi(t))$ as follows

$$\log(\phi(t)) = \kappa_1(it) + \kappa_2 \frac{(it)^2}{2!} + \cdots + \kappa_r \frac{(it)^r}{r!} + \cdots \quad (1)$$

It turns out that the first three cumulants are the first three moments (see, e.g., Kendall and Stuart 1963, p. 71)

$$\left.\begin{aligned} \kappa_1 &= \mu \\ \kappa_2 &= \mu_2 \\ \kappa_3 &= \mu_3 \end{aligned}\right\} \quad (2)$$

but that differences arise from the fourth one

$$\kappa_4 = \mu_4 - 3\mu_2^2.$$

3. SUM OF INDEPENDENT RANDOM VARIABLES

Let X and Y be two independent random variables with characteristic functions $\phi_X(t)$ and $\phi_Y(t)$, respectively. The characteristic function of the sum of X and Y is given by

$$\phi_{X+Y}(t) = \phi_X(t) \cdot \phi_Y(t)$$

and thus

$$\log(\phi_{X+Y}(t)) = \log(\phi_X(t)) + \log(\phi_Y(t)).$$

Using (1), we find that for any $r \geq 1$

$$\kappa_r(X+Y) = \kappa_r(X) + \kappa_r(Y).$$

By (2) we therefore deduce that

$$\left.\begin{array}{rcl} \mu(X+Y) & = & \mu(X) + \mu(Y) \\ \mu_2(X+Y) & = & \mu_2(X) + \mu_2(Y) \\ \mu_3(X+Y) & = & \mu_3(X) + \mu_3(Y) \end{array}\right\}, \qquad (3)$$

but we have

$$\mu_4(X+Y) = \mu_4(X) + \mu_4(Y) + 6\mu_2(X)\mu_2(Y).$$

Thus, the first three moments have an additivity property that the fourth one does not have.

Moreover, if $\overline{X}$ is the arithmetic mean of n independent random variables identically distributed like a random variable X, we have, as a consequence of (3), simple formulas for the first three moments of $\overline{X}$

$$\left.\begin{array}{rcl} \mu(\overline{X}) & = & \mu(X) \\ \mu_2(\overline{X}) & = & \mu_2(X)/n \\ \mu_3(\overline{X}) & = & \mu_3(X)/n^2 \end{array}\right\} \qquad (4)$$

but the fourth one is more complicated

$$\mu_4(\overline{X}) = \mu_4(X)/n^3 + 3(n-1)\mu_2^2(X)/n^3.$$

Note that if the n random variables in question constitute a sample of observations drawn without replacement from a finite population of size N with moments μ and μ_r $(r \geq 2)$, we have simple formulas too for the first three moments (see, e.g., Kendall and Stuart 1963, p. 302)

$$\left.\begin{array}{rcl} \mu(\overline{X}) & = & \mu \\ \mu_2(\overline{X}) & = & \frac{\mu_2}{n} \cdot \frac{N-n}{N-1} \\ \mu_3(\overline{X}) & = & \frac{\mu_3}{n^2} \cdot \frac{(N-n)(N-2n)}{(N-1)(N-2)} \end{array}\right\} \qquad (5)$$

but a more cumbersome formula for the fourth one

$$\mu_4(\overline{X}) = \frac{(N-n)\left\{(N^2 - 6nN + N + 6n^2)\mu_4 + 3N(n-1)(N-n-1)\mu_2^2\right\}}{n^3(N-1)(N-2)(N-3)}.$$

4. UNBIASED ESTIMATION

Let $X_1, \ldots, X_n$ be a sample of n independent observations drawn from a population with distribution F. From this sample, we want to estimate the moments μ and μ_r $(r \geq 2)$ of F. The theory of unbiased estimation (see, e.g., Halmos 1946) leads to the following natural unbiased estimators for, respectively, the first three moments μ, μ_2, and μ_3

$$\left.\begin{array}{rcl} \overline{X} & = & \frac{\sum X_i}{n} \\ S^2 & = & \frac{\sum (X_i - \overline{X})^2}{n} \cdot \frac{n}{n-1} \\ M_3 & = & \frac{\sum (X_i - \overline{X})^3}{n} \cdot \frac{n^2}{(n-1)(n-2)}. \end{array}\right\} \qquad (6)$$

The simplicity of the sequence of these estimators is again broken with the fourth one. One might have thought that a multiple of $\sum(X_i - \overline{X})^4/n$ were an unbiased estimator of μ_4 but the reality is that

$$\mu\left\{\frac{\sum(X_i - \overline{X})^4}{n}\right\} = \frac{(n-1)(n^2-3n+3)}{n^3}\mu_4 + \frac{3(2n-3)(n-1)}{n^3}\mu_2^2.$$

Note that in the case of a sample of observations drawn without replacement from a finite population of size N, the sequence of estimators of (6) must be replaced by (see, e.g., Kendall and Stuart 1963, p. 300)

$$\left.\begin{array}{l} \frac{\sum X_i}{n} \\ \frac{\sum (X_i - \overline{X})^2}{n} \cdot \frac{n}{n-1} \cdot \frac{N-1}{N} \\ \frac{\sum (X_i - \overline{X})^3}{n} \cdot \frac{n^2}{(n-1)(n-2)} \cdot \frac{(N-1)(N-2)}{N^2}. \end{array}\right\} \qquad (7)$$

Let us also mention that the unbiased estimation of the fourth cumulant κ_4 of F is not really simple too since it involves the estimator (see Kendall and Stuart 1963, p. 281)

$$\left\{\frac{n+1}{n} \cdot \frac{\sum(X_i - \overline{X})^4}{n} - \frac{3(n-1)}{n} \cdot \left(\frac{\sum(X_i - \overline{X})^2}{n}\right)^2\right\} \cdot \frac{n^3}{(n-1)(n-2)(n-3)}. \qquad (8)$$

As $\kappa_4 = \mu_4 - 3\mu_2^2$, one might have thought that the sequence of estimators of (6) for the cumulants could be continued with (8) but without the factors $(n+1)/n$ and $(n-1)/n$.

5. COVARIANCES BETWEEN ESTIMATORS

Let us compute now the covariances involving the three estimators—$\overline{X}$, S^2, and M_3—introduced in the previous section. Note that all results given here are exact results and not approximations to order n^{-1} as in Kendall and Stuart (1963, chap. 10). In particular, we have the sequence of results

$$\mathrm{cov}(\overline{X}, \overline{X}) = \mu_2/n,$$

$$\mathrm{cov}(\overline{X}, S^2) = \mu_3/n,$$

and

$$\text{cov}(\overline{X}, M_3) = \mu_4/n - 3\mu_2^2/n \quad (= \kappa_4/n)$$

which leads one to note that once μ_4 appears, the formulas become more complicated. The other covariances are still much more cumbersome. We have

$$\text{cov}(S^2, S^2) = \frac{\mu_4}{n} - \frac{n-3}{n(n-1)}\mu_2^2,$$

$$\text{cov}(S^2, M_3) = \frac{\mu_5}{n} - \frac{2(2n-5)}{n(n-1)}\mu_3\mu_2,$$

and

$$\text{cov}(M_3, M_3) = \frac{\mu_6}{n} - \frac{3(2n-5)}{n(n-1)}\mu_4\mu_2 - \frac{n-10}{n(n-1)}\mu_3^2 + \frac{3(3n^2-12n+20)}{n(n-1)(n-2)}\mu_2^3.$$

When the distribution of the population from which observations are drawn is normally distributed with variance σ^2, these complicated formulas become very simple. We have $\text{cov}(\overline{X}, S^2) = \text{cov}(\overline{X}, M_3) = \text{cov}(S^2, M_3) = 0$ and the sequence of results

$$\left.\begin{array}{rcl} \text{cov}(\overline{X}, \overline{X}) & = & \frac{\sigma^2}{n} \\ \text{cov}(S^2, S^2) & = & \frac{2!\sigma^4}{n} \cdot \frac{n}{n-1} \\ \text{cov}(M_3, M_3) & = & \frac{3!\sigma^6}{n} \cdot \frac{n^2}{(n-1)(n-2)}. \end{array}\right\} \qquad (9)$$

This sequence of results is however interrupted because of the absence of a simple unbiased estimator for μ_4.

6. CONCLUSION

The fact that such remarkable sequences of results as (2), (3), (4), (5), (6), (7), or (9) cannot be continued after the third element leads to the conclusion that statistical theory provides beautiful formulas when they involve the first three moments (with a special prize for the insufficiently known formula $\text{cov}(\overline{X}, S^2) = \mu_3/n$) but that serious complications arise once the fourth one is introduced.

A question that may be asked now is "Why do descriptive statistics deal with four concepts, while the mathematics are only tractable at the third level?" This curious feature is perhaps connected with the fact that, contrary to the concepts of location, dispersion, and skewness, the concept of kurtosis (usually seen through the fourth central moment) has met some difficulties of interpretation. From *The American Statistician* we can cite, for example, Darlington (1970) who stated that "a far better term for describing kurtosis is bimodality" and Moors (1986) that "the concept of kurtosis seems to be rather difficult to interpret" or Ruppert (1987) that "there is no agreement on what kurtosis measures". In an article that treats the axiomatic definitions of location, dispersion, and skewness, Rousson (1995) also concluded that only three kinds of descriptive measures (and not four) can be rigorously and harmoniously defined. From this point of view, descriptive statistics and mathematical simplicity agree.

[Received February 1997. Revised December 1997.]

REFERENCES

Darlington, R. B. (1970), "Is Kurtosis Really Peakedness?", *The American Statistician*, 24, 19–22.

Halmos, P. R. (1946), "The Theory of Unbiased Estimation," *The Annals of Mathematical Statistics*, 17, 34–43.

Moors, J. J. A. (1986), "The Meaning of Kurtosis: Darlington Reexamined," *The American Statistician*, 40, 283–284.

Kendall, M. G., and Stuart, A. (1963), *The Advanced Theory of Statistics*, (vol. 1, 2nd ed), London: Charles Griffin.

Rousson, V. (1995), "Axiomatique et Statistique Descriptive," *Student*, 1, 17–34.

Ruppert, D. (1987), "What is Kurtosis? An Influence Function Approach," *The American Statistician*, 41, 1–5.