SOME EFFICIENT MEASURES OF RELATIVE DISPERSION[1]

BY NILAN NORRIS

For some time it has been known that the coefficient of variation (in the sense of the ratio of the standard deviation to the arithmetic mean) is not an efficient statistic for distributions departing materially from normality.[2] At various times there have been proposed certain supplementary estimates of relative variation, such as those involving ratios between sums and differences of upper and lower quartiles, and ratios of mean deviations to medians or to arithmetic means. Some of these have appeared in certain textbooks.[3] But there appears to have been no attempt to found their use on considerations of minimum sampling variance.

The point of departure of this paper is that of using the Method of Maximum Likelihood to derive two efficient measures of relative dispersion, together with expressions for their standard errors. These optimum estimates of true or parametric variation are the ratio of the arithmetic mean to the geometric mean (the arithmetic-geometric ratio) for Pearson Type III distributions, and the ratio of the geometric mean to the harmonic mean (the geometric-harmonic ratio) for Pearson Type V distributions. The usefulness of these measures is suggested by the generalized-mean-value-function approach to the analysis of averages, especially the theorem of inequalities among averages.[4]

---

[1] Presented before a joint meeting of the American Statistical Association and the Institute of Mathematical Statistics at Chicago, Illinois on December 28, 1936.

[2] The term "efficient statistic" is used here in the sense of R. A. Fisher, that is, of a parameter-estimate which tends towards normality of distribution with the least possible standard deviation. For a discussion of the inefficiency of certain commonly used statistics as applied to distributions departing from normality, see R. A. Fisher, "On the Mathematical Foundations of Theoretical Statistics," *Philosophical Transactions of the Royal Society of London*, Series A, Vol. 222, 1922, pp. 332–336.

[3] See, for example, William Vernon Lovitt and Henry F. Holtzclaw, *Statistics* (Prentice-Hall, Inc., New York, 1929), p. 134; Herbert Arkin and Raymond R. Colton, *Statistical Methods* (Barnes and Noble, Inc., New York, 1935), revised ed., p. 41; and Herbert Sorenson, *Statistics for Students in Psychology and Education* (McGraw-Hill Book Company, Inc., New York, 1936), pp. 153 *f.*

[4] Nilan Norris, "Inequalities among Averages," *Annals of Mathematical Statistics*, Vol. VI, No. 1, March, 1935, pp. 27–29; and "Convexity Properties of Generalized Mean Value Functions," *Annals of Mathematical Statistics*, Vol. VIII, No. 2, June, 1937, pp. 118–120. Professor John B. Canning appears to have been the first to point out the possibility of making use of certain ratio measures of relative variation. See "The Income Concept and Certain of Its Applications," *Papers and Proceedings of the Eleventh Annual Conference of the Pacific Coast Economic Association* (Edwards Brothers, Ann Arbor, 1933), p. 64.

This theorem states that if $t_1 < t_2$, then $\phi(t_1) < \phi(t_2)$, where the unit weight or simple sample type of generalized mean value function is defined as

$$\phi(t) = \left(\frac{x_1^t + x_2^t + \cdots + x_n^t}{n}\right)^{\frac{1}{t}}. \tag{1}$$

The x_i are restricted to positive real numbers not all equal, but t may take any real value. A necessary and sufficient condition that $\phi(-\infty) = \phi(t) = \phi(\infty)$ is the excluded trivial case that $x_1 = x_2 = \cdots = x_n$. When the x_i are not all equal, the ratios between various pairs of averages as generated by $\frac{\phi(t_2)}{\phi(t_1)}$ yield ratio measures of relative dispersion, the usefulness of which depends, in part, on their efficiency as estimates of population-characterizing constants (parameters). The arithmetic-geometric ratio may be written $\frac{\phi(1)}{\phi(0)} = \frac{A}{G}$; and the geometric-harmonic ratio may be written $\frac{\phi(0)}{\phi(-1)} = \frac{G}{H}$. In certain cases it may be of convenience to reverse the order of each of the ratios. The standard errors for the two forms which each of the ratios may assume are presented below.

The demonstration that these ratio measures of relative dispersion are 100% efficient statistics for their appropriate distributions, and the derivation of useful expressions for their respective standard errors both may be accomplished by the ordinary method of differentiating the logarithm of the likelihood.

$$\text{Let digamma of } x = F_D(x) = \frac{d}{dx}\log x!,$$

$$\text{and trigamma of } x = F_T(x) = \frac{d^2}{dx^2}\log x!$$

For Pearson Type III distributions, the frequency with which the variate x falls into the range dx is given by

$$df = \frac{1}{p!}\left(\frac{x}{a}\right)^p e^{-\frac{x}{a}}\frac{dx}{a}. \tag{2}$$

The parameter a measures the absolute dispersion of the distribution, and the parameter p determines the general shape of the frequency curve. The relative variation may be regarded as a population parameter, θ, defined as the ratio of the population arithmetic mean to the population geometric mean. Let the logarithm of the likelihood for this distribution be represented by L, we have

$$L = -n\log p! - n(p+1)\log a + \Sigma \log x_i - \frac{1}{a}\Sigma x_i, \tag{3}$$

where the summation is taken over the n individuals of the sample. It follows that

$$\frac{\partial L}{\partial a} = -\frac{n}{a}(p+1) + \frac{1}{a^2}\Sigma x_i; \quad \text{and} \quad \frac{\partial^2 L}{\partial a^2} = \frac{n}{a^2}(p+1) - \frac{2}{a^3}\Sigma x_i. \tag{4}$$

When L is maximized with respect to a by equating to zero the first derivative of L with respect to a, we find

$$\frac{\Sigma x_i}{n} = a(p+1). \tag{5}$$

It also follows that

$$\begin{aligned} \frac{\partial L}{\partial p} &= -nF_D(p) - n \log a + \Sigma \log x_i ; \\ \frac{\partial^2 L}{\partial p^2} &= -nF_T(p); \quad \text{and} \quad \frac{\partial^2 L}{\partial a\, \partial p} = \frac{\partial^2 L}{\partial p\, \partial a} = -\frac{n}{a}. \end{aligned} \tag{6}$$

When L is maximized with respect to p by equating to zero the first derivative of L with respect to p, we find

$$(\Pi x_i)^{\frac{1}{n}} = ae^{F_D(p)} \tag{7}$$

The optimum estimate $\hat{p}$ of p is therefore found from (5) and (7) to be given by the equation

$$(p+1)e^{-F_D(p)} = \frac{\Sigma x_i}{n} \Big/ (\Pi x_i)^{\frac{1}{n}} = \frac{A}{G} \tag{8}$$

But $(p+1)e^{-F_D(p)}$ is the parameter θ. Hence we find the optimum estimate of θ to be $\frac{A}{G}$, which can be expressed in terms of the generalized mean value function as $\frac{\phi(1)}{\phi(0)}$. Therefore, for distributions well graduated by a Type III curve the optimum estimate of θ, the ratio of the arithmetic mean to the geometric mean, is given by $\frac{A}{G}$.

If only p is being estimated, (a given) the variance, or square of the standard deviation of p is obtained from $\frac{\partial^2 L}{\partial p^2}$, and is $V(p) = \frac{1}{nF_T(p)}$. To a first approximation, the variance of $\frac{A}{G}$, the estimate of θ, is found from the usual relation between the variance of a function and the variance of the argument, namely

$$V[f(x)] = \left[\frac{df(x)}{dx}\right]^2 V(x). \tag{9}$$

Since

$$\frac{d}{dp}(\theta) = \theta\left[\frac{1}{p+1} - F_T(p)\right], \tag{10}$$

therefore

$$V\left(\frac{A}{G}\right) = \theta^2 \frac{\left[F_T(p) - \frac{1}{p+1}\right]^2}{nF_T(p)}, \tag{11}$$

or the standard error of $\frac{A}{G}$ is the square root of the last expression, if only p is being estimated. If it is more convenient to do so, one may reverse the terms in the ratio to obtain

$$(12) \qquad V\left(\frac{G}{A}\right) = \theta^{-2} \frac{\left[F_T(p) - \frac{1}{p+1}\right]^2}{nF_T(p)},$$

and extract the square root of the last expression to obtain the standard error of $\frac{G}{A}$.

If a and p are being estimated simultaneously, there exists the matrix of negative mean values

$$(13) \qquad \begin{Vmatrix} -E\left(\frac{\partial^2 L}{\partial a_2}\right) & -E\left(\frac{\partial^2 L}{\partial a \partial p}\right) \\ -E\left(\frac{\partial^2 L}{\partial a \partial p}\right) & -E\left(\frac{\partial^2 L}{\partial p^2}\right) \end{Vmatrix} \equiv \begin{Vmatrix} \frac{n}{a^2}(p+1) & \frac{n}{a} \\ \frac{n}{a} & nF_T(p) \end{Vmatrix}$$

from which the variance of $\frac{A}{G}$ can be computed. In fact we have

$$(14) \qquad V(\dot{p}) = \frac{p+1}{n[(p+1)F_T(p) - 1]} = \frac{1}{n\left[F_T(p) - \frac{1}{p+1}\right]},$$

and consequently

$$(15) \qquad V\left(\frac{A}{G}\right) = \theta^2 \frac{\left[F_T(p) - \frac{1}{p+1}\right]}{n}.$$

The standard error of $\frac{A}{G}$ is equal to the square root of the expression in (15), if both a and p are being estimated. If the terms in the ratio are reversed, one obtains

$$(16) \qquad V\left(\frac{G}{A}\right) = \theta^{-2} \frac{\left[F_T(p) - \frac{1}{p+1}\right]}{n}.$$

The square root of the last expression may be taken to derive the standard error of $\frac{G}{A}$. Since the digamma and the trigamma functions have been tabulated for considerable ranges,[5] these standard error formulae, and those developed below for the Type V case should be quite useful.

[5] *British Association for the Advancement of Science: Mathematical Tables* (Office of the British Association, London, 1931), Vol. I, pp. 42–51.

For Pearson Type V distributions, the frequency with which the variate x falls into the range dx is given by

$$df = \frac{1}{p!}\left(\frac{a}{x}\right)^{p+2} e^{-\frac{a}{x}} \frac{dx}{a}. \tag{17}$$

The parameter a measures the absolute dispersion of the distribution, and the parameter p determines the general shape of the frequency curve. The relative dispersion may be regarded as a population parameter, θ', defined as the ratio of the population geometric mean to the population harmonic mean. Let the logarithm of the likelihood for this distribution be represented by L. Then

$$L = -n \log p! + n(p+1) \log a - (p+2) \Sigma \log x_i - a\Sigma \frac{1}{x_i}, \tag{18}$$

the summation being taken over the sample of n individuals. It follows that

$$\begin{aligned} \frac{\partial L}{\partial p} &= -nF_D(p) + n \log a - \Sigma \log x_i; \\ \frac{\partial^2 L}{\partial p^2} &= -n F_T(p); \\ \frac{\partial L}{\partial a} &= \frac{n}{a}(p+1) - \Sigma \frac{1}{x_i}; \\ \frac{\partial^2 L}{\partial a^2} &= -\frac{n}{a^2}(p+1); \\ \frac{\partial^2 L}{\partial p \partial a} &= \frac{\partial^2 L}{\partial a \partial p} = \frac{n}{a}. \end{aligned} \tag{19}$$

Let L be maximized with respect to p to derive the geometric mean, and let L be maximized with respect to a to derive $\phi(-1)$, or H, the harmonic mean. It is clear that for the Type V distribution, the relative dispersion, as we have defined it, is the population parameter $\theta' = \frac{e^{F_D(p)}}{p+1}$. Therefore, if $\phi(0) = G = (\Pi x_i)^{\frac{1}{n}}$, and $\phi(-1) = H = \frac{1}{\frac{1}{n}\Sigma \frac{1}{x_i}}$, it follows, by an argument similar to that used in the case of a Type III curve, that the geometric-harmonic ratio, $\frac{G}{H}$, is an optimum estimate of the parameter θ', for distributions well graduated by the Pearson Type V curve.

If only p is being estimated, the variance of $\hat{p}$ is given by $V(\hat{p}) = \frac{1}{nF_T(p)}$, and

$$V\left(\frac{G}{H}\right) = \theta'^2 \frac{\left[F_T(p) - \frac{1}{p+1}\right]^2}{nF_T(p)}, \tag{20}$$

or the standard error of $\frac{G}{H}$ is the square root of the last expression, if p alone is being estimated, a being given. If the terms in the ratio are reversed,

$$V\left(\frac{H}{G}\right) = \theta'^2 \frac{\left[F_T(p) - \frac{1}{p+1}\right]^2}{nF_T(p)}, \tag{21}$$

and the square root of the last expression yields the standard error of $\frac{H}{G}$.

If a and p are being estimated simultaneously, there exists the matrix

$$\begin{Vmatrix} -E\left(\frac{\partial^2 L}{\partial a^2}\right) & -E\left(\frac{\partial^2 L}{\partial a \partial p}\right) \\ -E\left(\frac{\partial^2 L}{\partial a \partial p}\right) & -E\left(\frac{\partial^2 L}{\partial p^2}\right) \end{Vmatrix} = \begin{Vmatrix} \frac{n}{a^2}(p+1) & -\frac{n}{a} \\ -\frac{n}{a} & nF_T(p) \end{Vmatrix} \tag{22}$$

from which the variance of $\frac{G}{A}$ can be found. In fact

$$V(\hat{p}) = \frac{1}{n\left[F_T(p) - \frac{1}{p+1}\right]} \tag{23}$$

and hence

$$V(\theta') = \frac{\theta'^2}{n}\left[F_T(p) - \frac{1}{p+1}\right]. \tag{24}$$

The standard error of $\frac{G}{H}$ is then given by the square root of the expression for $V(\theta')$. If the terms in the ratio are reversed,

$$V\left(\frac{H}{G}\right) = \theta'^{-2}\left[F_T(p) - \frac{1}{p+1}\right], \tag{25}$$

the square root of which yields the standard error of $\frac{H}{G}$.

Just as the coefficient of variation is an efficient statistic only for distributions well graduated by the normal, or Pearson Type VII curve, so also the two maximum likelihood estimates of relative dispersion herein developed are efficient only when applied to their appropriate distributions. One may expect to obtain an optimum degree of efficiency only when the arithmetic-geometric ratio is used for series well specified by the Type III function, and the geometric-harmonic ratio is used for series well specified by the Type V function.

It may be recalled that Karl Pearson proposed the use of the coefficient of variation late in the nineteenth century.[6] Since that time there appears to have been some tendency to rely on it as a measure of relative variation, regard-

[6] "Regression, Heredity, and Panmixia," *Philosophical Transactions of the Royal Society of London*, Series A, Vol. 187, 1896, p. 277. For materials pertaining to the Pearson-Thorndike controversy resulting from the latter's suggestion that the ratio of the standard deviation to the square root of the arithmetic mean is often a more suitable device than is

less of whether or not it extracts from the sample a relatively large amount of the pertinent information concerning the parent population.[7] There are several cases in which the coefficient of variation is not an optimum estimate of relative dispersion. For example, in a comparison of the true or parametric variation of the weights of humans of given age levels, the arithmetic-geometric ratio is often the appropriate statistic to use, since weights tend to be distributed according to the Pearson Type III law. Frequently the distribution of weights is very well graduated by the Type V function, if the origin is fixed at 0 in advance. Although this procedure yields a special two-parameter Type V function, the principle of using the geometric-harmonic ratio as an optimum estimate of relative dispersion is still valid. Again, in a comparison of the relative variation of the personal distribution of wealth and income in certain modern countries, the arithmetic-geometric ratio will be found to have a smaller sampling variance than that of the coefficient of variation, since the personal distribution of wealth and income in these countries tends to be in accordance with the Type III law, rather than the normal law. Similarly, the distribution of the number of trials required to obtain r successes of an event having a given probability usually follows the Type III function, and requires the use of the arithmetic-geometric ratio, if the maximum amount of the relevant information is to be extracted from the sample.

It seems clear that in practice the usefulness of the arithmetic-geometric ratio and the geometric-harmonic ratio will depend on the type of the distribution with which one is dealing, and on the extent to which added efficiency is desired. In certain cases there is doubtless room for some difference of opinion as to whether or not the degree of added efficiency achieved by the use of these maximum likelihood estimates of relative dispersion will merit departing from the use of such a time-honored statistic as the coefficient of variation. If one is interested in avoiding the assumption of normality implicit in methods customarily used in the more general problem of analysis of variance, an alternative is the use of ranks.[8] Although the efficiency of these rank-correlation methods is not always 100%, their economy of effort is sometimes a great advantage.

HUNTER COLLEGE OF THE CITY OF NEW YORK.

the coefficient of variation see Edward L. Thorndike, "Empirical Studies in the Theory of Measurement," *Archives of Psychology* (The Science Press, New York, 1907), Vol. I, No. 3, April, 1907, pp. 9-13; and *An Introduction to the Theory of Mental and Social Measurements* (Teachers College, Columbia University, New York, 1913), 2d. ed. pp. 133 *f.*, or 1st. ed., 1904, pp. 102 *f.* See also Helen M. Walker, *Studies in the History of Statistical Method* (The Williams and Wilkins Company, Baltimore, 1929), p. 178.

[7] *Cf.* Walter A. Hendricks and Kate W. Robey, "The Sampling Distribution of the Coefficient of Variation," *Annals of Mathematical Statistics*, Vol. VII, No. 4, December, 1936, pp. 129-132.

[8] Harold Hotelling and Margaret Richards Pabst, "Rank Correlation and Tests of Significance Involving No Assumption of Normality," *Annals of Mathematical Statistics*, Vol. VII, No. 1, March, 1936, pp. 29-43. See also Milton Friedman, "The Use of Ranks to Avoid the Assumption of Normality Implicit in the Analysis of Variance," *Journal of the American Statistical Association*, Vol. 32, No. 200, December, 1937, pp. 675-701.