# THE ESTIMATION OF A QUOTIENT WHEN THE DENOMINATOR IS NORMALLY DISTRIBUTED

By Robert D. Gordon

*Scripps Institution of Oceanography, La Jolla, Calif.*

**1. Introduction.** In an oceanographic investigation we have to deal with a time series consisting of single pairs of observed values $x$, $y$, of two independent stochastic variables, whose true (mean) values we shall denote respectively by $a$, $b$. Of interest is the corresponding time series of quotients $(b/a)$, which it is required to estimate from the observations $x$, $y$. Both $x$ and $y$ are approximately normally distributed about their mean values $a$, $b$ with rather large variances $\sigma_x^2$, $\sigma_y^2$ which can be estimated. It is easily possible for $x$ to vanish or even to be of opposite sign to $a$, although $a$ cannot itself vanish. The required estimates of $(b/a)$ should have the property that they can be numerically integrated, i.e. that an arbitrary sum of such estimates shall equal the corresponding estimate of the true sum.

Let us define a function $\gamma(x)$ to have the property that its mathematical expectation $E\{\gamma(x)\}$ is exactly $1/a$, where $a = E(x)$. If such a function exists we shall have

$$E\{y\cdot\gamma(x)\} = E(y)\cdot E\{\gamma(x)\} = b\cdot(1/a) = b/a \tag{1}$$

so that $y\cdot\gamma(x)$ will be an estimate of $b/a$ which has the required property: namely such estimates can be added, and we have

$$E\{y_1\gamma(x_1) + y_2\gamma(x_2)\} = E\{y_1\gamma(x_1)\} + E\{y_2\gamma(x_2)\} = b_1/a_1 + b_2/a_2$$

as required. It turns out that if $x$ is normally distributed with non-zero mean such a function $\gamma(x)$ does exist, and is given by the formula

$$\gamma(x) = \frac{1}{\sigma_x} \exp\,(x^2/2\sigma_x^2) \int_{x/\sigma_x}^{\infty} e^{-t^2/2}\,dt = \frac{1}{\sigma_x} R_{x/\sigma_x} \tag{2}$$

where $R_u$ is the "ratio of the area to the bounding ordinate" which is tabulated by J. P. Mills,[1] also in Pearson's tables.[2] Equation (2) holds if $a$ is positive; if $a$ is negative the integration should extend over $(x/\sigma_x, -\infty)$. It is easy to verify that

$$E(\gamma(x)) = \frac{1}{\sqrt{2\pi}\,\sigma_x} \int_{-\infty}^{\infty} \gamma(x) \exp\left(-\frac{(x-a)^2}{2\sigma_x^2}\right) dx = \frac{1}{a} \tag{3}$$

by direct substitution from (2).

[1] J. P. Mills, "Table of ratio: area to bounding ordinate, for any portion of the normal curve," *Biometrika*, Vol. 18 (1926), pp. 395–400.

[2] Karl Pearson, *Tables for Statisticians and Biometricians*, part II, table III, Cambridge Univ. Press.

**2. The law of large numbers for $\gamma(x)$.** The function $\gamma(x)$ defined by (2) has mean value $1/a$ as required, but its second moment (hence variance) does not exist, as may readily be verified. By a theorem of Khinchine[3] however, its values satisfy a law of large numbers. It will be of interest to inquire about the "strength" of this law of large numbers for $\gamma(x)$. Namely, given a positive number $\epsilon$, how many "observations" (independent estimates) $\gamma(x)$ will suffice to guarantee probabilities of .50, .90, .95, etc. for the following inequality to hold

$$\left| \frac{\gamma(x_1) + \gamma(x_2) + \cdots + \gamma(x_n)}{n} - \frac{1}{a} \right| < \epsilon \tag{4}$$

where $n$ is the number of "observations."

In order to arrive at a rough answer to this question we have made use of certain inequalities due to Tshebysheff (Tshebysheff's "method of moments", cf. Uspensky[4]). Let $u$ be an arbitrary stochastic variable whose distribution has moments of the first and second order which are known. Denote by $m$ its first moment, by $\sigma^2$ its variance, then it results from Tshebysheff's theory that the probability $P(u_1, u_2)$ for a value of $u$ to lie between $u_1$ and $u_2$ (i.e. $u_1 \leqq u \leqq u_2$) satisfies the inequality

$$P(u_1, u_2) > 1 - \frac{\sigma^2}{(u_1 - m)^2 + \sigma^2} - \frac{\sigma^2}{(u_2 - m)^2 + \sigma^2}. \tag{5}$$

This inequality is independent of the values, or even the existence, of further moments of the $u$-distribution beyond the second, and depends only on the condition that the cumulant of the distribution function shall have at least three "points of increase."

Although $\gamma(x)$ does not have a second moment, a second moment does exist for those values of $\gamma(x)$ which correspond to $x \geqq -\theta > -\infty$, where $\theta$ is an arbitrary number, positive or negative. If we can estimate the first two moments of $\gamma(x) \sim 1/x$ corresponding to a given value of $\theta$, then for a given number $n$ of observations we need only to divide the corresponding variance by $n$ to obtain $\sigma^2$ in (5), then multiply (5) by the $n$th power of the (normal) probability for the inequality $x \geqq -\theta$, in order to obtain a lower bound for the probability of the inequality (4). $\theta$ is to be determined so as to yield a maximum result.

The first moment $m_1$ of $\gamma(x)$ for values of $x \geqq -\theta$ is easily computed, and is given by the formula

$$\sigma_x m_1(\theta) = \frac{\sigma_x}{a}\left\{1 - \frac{R_{-\theta/\sigma_x}}{R_{-(\theta+a)/\sigma_x}}\right\}. \tag{6}$$

[3] J. V. Uspensky, *Introduction to Mathematical Probability*, pp. 195, McGraw-Hill (1937).

[4] J. V. Uspensky, l.c. pp. 365 ff.

The second moment is harder to compute, but if we place

$$\phi(\theta) = K\cdot(m_2 - m_1^2) = \frac{1}{\sqrt{2\pi}\,\sigma_x}\int_{-\theta}^{\infty}[\gamma(x)]^2 \exp\left(-\frac{(x-a)^2}{2\sigma_x^2}\right)dx$$

$$(7) \qquad - \frac{\left[\int_{-\theta}^{\infty}\gamma(x)\exp\left(-\frac{(x-a)^2}{2\sigma_x^2}\right)dx\right]^2}{\sqrt{2\pi}\,\sigma_x \int_{-\theta}^{\infty}\exp\left(-\frac{(x-a)^2}{2\sigma_x^2}\right)dx}$$

where

$$K = \frac{1}{\sqrt{2\pi}\,\sigma_x}\int_{-\theta}^{\infty}\exp\left(-\frac{(x-a)^2}{2\sigma_x^2}\right)dx = \frac{1}{\sqrt{2\pi}}\int_{-(\theta+a)/\sigma_x}^{\infty} e^{-t^2/2}\,dt$$

we easily obtain the relationship

$$(8) \qquad \phi'(\theta) = \frac{1}{\sqrt{2\pi}\,\sigma_x^3}\exp\left(-\frac{(\theta+a)^2}{2\sigma_x^2}\right)\left\{R_{(-\theta/\sigma_x)} - \frac{\sigma_x}{a}\left(1 - \frac{R_{-\theta/\sigma_x}}{R_{-(\theta+a)/\sigma_x}}\right)\right\}^2.$$

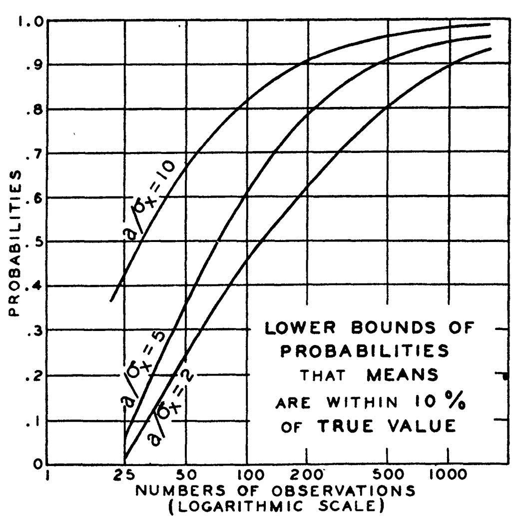


FIG. 1

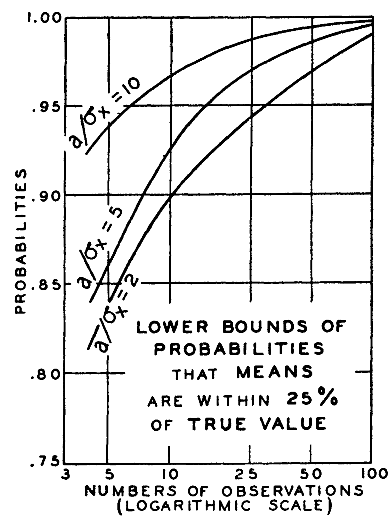


FIG. 2

From (7), using a table of the probability integral, it can be verified that $\phi(-a - 3\sigma_x) \ll 0.001$. Assume, therefore, as a boundary condition $\phi(-a - 3\sigma_x) = 0$ then (8) can be integrated graphically or numerically. It is by this means that the curves shown in Figs. 1 and 2 were determined. Computations were also attempted for $a/\sigma_x = \frac{1}{2}$, $a/\sigma_x = 1$, but it was not possible to obtain significant results: it would be necessary in these cases to take more than two moments into account, which would lead to hopeless complications. In these figures the ordinates represent probabilities for an observation to fall between *.90a* and *1.11a* (Figure 1), and between *.75a* and *1.33a* (Figure 2), respectively.

**3. Two practical formulas for computations.** It seems worthwhile to note here two simple formulas in connection with Mills' ratio (2) which will be useful for computations. The first is the obvious relationship

$$R_{-u} = \sqrt{2\pi}\, e^{u^2/2} - R_u = 1/z - R_u \tag{9}$$

in the notation of Pearson's tables. The second applies to large values of $x$, and may be written

$$\frac{x}{x^2 + \sigma_x^2} < \gamma(x) = \frac{1}{\sigma_x} R_{x/\sigma_x} < \frac{1}{x} \tag{10}$$

(10) is true for $x > 0$, and can be proved by means of the differential equation which $\gamma(x)$ satisfies.

**4. Remarks.** The estimate $\gamma(x)$ has the following inadequacy: If only a single observation $x$ is known, then it is unknown whether $a$ is of like or unlike sign compared to $x$. It turns out then that the mathematical expectation for the value of $\gamma(x)$ vanishes identically. This difficulty of course disappears if more than one observation is available. Methods of avoiding this difficulty for time series, e.g. by noting relative frequencies for observations separated by 1, 2, 3 etc. intervals to agree in sign, will be discussed elsewhere in connection with practical applications.

It may be worthwhile to note that Geary[5] developed certain characteristics of the distribution of a quotient, which however are not adapted to our purposes.

# NOTE ON CONFIDENCE LIMITS FOR CONTINUOUS DISTRIBUTION FUNCTIONS

By A. Wald* and J. Wolfowitz

In a recent paper [1] we discussed the following problem: Let $X$ be a stochastic variable with the cumulative distribution function $f(x)$, about which nothing is known except that it is continuous. Let $x_1, \cdots, x_n$ be $n$ independent, random observations on $X$. The question is to give confidence limits for $f(x)$. We gave a theoretical solution when the confidence set is a particularly simple and important one, a "belt."

A particularly simple and expedient way from the practical point of view is to construct these belts of uniform thickness ([1], p. 115, equation 50). If the appropriate tables, as mentioned in our paper, were available, the construction of confidence limits, no matter how large the size of the sample, would be immediate.

Our formulas (11), (16), (19), (27) and (29) are not very practical for computation, particularly when the samples are large. We have recently learned that

[5] Geary, R. C., "The Frequency Distribution of a Quotient," *Jour. Roy. Stat. Soc.*, Vol. 93 (1930), pp. 442–446.

* Columbia University, New York City.