D1-82-0348

BOEING SCIENTIFIC RESEARCH LABORATORIES

Also available from the author

Ratios of Normal Variables and Ratios of Sums of Uniform Variables

George Marsaglia

Mathematics Research

April 1964

D1-82-0348

RATIOS OF NORMAL VARIABLES AND RATIOS OF SUMS OF UNIFORM VARIABLES

by

George Marsaglia

Mathematical Note No. 348

Mathematics Research Laboratory

BOEING SCIENTIFIC RESEARCH LABORATORIES

April 1964

SUMMARY

The principal part of this paper is devoted to the study of the distribution and density functions of the ratio of two normal random variables. It gives several representations of the distribution function in terms of the bivariate normal distribution and Nicholson's V function, both of which have been extensively studied, and for which tables and computational procedures are readily available. One of these representations leads to an easy derivation of the density function in terms of the Cauchy density and the normal density and integral. A number of graphs of the possible shapes of the density are given, together with an indication of when the density is unimodal or bimodal.

The last part of the paper discusses the distribution of the ratio $(u_1 + \cdots + u_n)/(v_1 + \cdots + v_m)$ where the u's and v's are independent, uniform variables. The exact distribution for all n and m is given, and some approximations discussed.

1. Introduction. The first part of this paper will discuss the distribution of the ratio of normal random variables; the second part, the distribution of the ratio of sums of uniform random variables. There does not seem to be much in the literature concerning the ratio of normal variables - there are some comments by Curtiss in his paper, [2], on the ratios of arbitrary variates, and papers by Fieller [4], and Geary [5], all of which are quite old. It might be thought that the subject is so simple that it was considered long ago, then dropped, but this is not quite the case. Unless the means are zero, where one easily gets the Cauchy distribution, the distribution of the ratio of normal variables does not respond readily to the devices that work so well for other important quotients in statistics, e.g., those of t, z, or F. Curtiss remarks that it is apparently impossible to evaluate the density in closed form, a rather vague statement. We will derive the exact density of the ratio of two arbitrary normal variates by what might be called modern methods - not in the sense of using powerful new techniques, but merely by using properties of distributions that have been extensively studied in the intervening years. The density may be expressed as the product of a Cauchy density and a factor involving the normal density and integral, which might be considered a closed expression (equation (5) of Section 2). At any rate, there are now available a number of methods for handling the functions associated with the distribution and density of the ratio, and with the aid of a computer, we may study them in detail.

Aside from its frequent occurrence in problems involving the ratio of measured quantities with a random, presumably normal, error, the

problem of the ratio of normal variates is of importance in regression theory. In fitting a line to points $(x_1,y_1),\ldots,(x_n,y_n)$, the x's assumed constant and the y's independent normal with $E(y_i) = \alpha + \beta x_i$, one gets $\hat{\alpha}$ and $\hat{\beta}$ as estimates of α and β by least squares. It is natural to estimate the x-intercept of the regression line in the form $-\hat{\alpha}/\hat{\beta}$, and thus the problem of the ratio of normal variates arises.

The following example of this problem occurs in medicine: in order to estimate the life span of the circulating red blood cells of a subject, a number of his red cells are labelled and then, by some means or other, the number of labelled cells still in the circulation is sampled, say, every 5 days for 50 days. This gives a sequence of points which are plotted and fitted with a straight line; the point where the line intercepts the time axis is used as the estimate of the red cell life span. It is important to know the distribution of this estimate about its true value - the normal red cell life span is about 120 days and shortened life spans are associated with various hematological disorders, most of them severe.

We will discuss the distribution and density of the ratio of two normal random variables in Section 2. In Section 4 we will discuss the distribution of ratios of the form $(u_1 +\cdots+ u_n)/(v_1 +\cdots+ v_m)$ where the u's and v's are independent uniform variables; a recent paper [8], on this distribution for $n = m = 3$ led to its being considered here. We will find the exact distribution for all n and m, and examine the closeness of the normal approximation. On the way to finding the distribution of $(u_1 +\cdots+ u_n)/(v_1 +\cdots+ v_m)$ we will need the distribution of a linear combination of uniform variates; some comments on this distribution and its history are in Section 3.

2. Ratios of normal variables We are concerned with the distribution of the ratio of two normal random variables. The problem has been discussed in the past, [2,4,5]. We will bring the problem up to date in this Section - give an explicit representation of the distribution in terms of what are now familiar functions, and discuss in more detail some of the properties of the distribution.

Let

$$(1) \qquad w = \frac{a + x}{b + y}$$

where a,b are non-negative constants and x,y are independent standard normal random variables. It is easy to see that if $w' = x_1/y_1$ is the ratio of two arbitrary normal random variables, correlated or not, then there are constants c_1 and c_2 such that $c_1 + c_2 w'$ has the same distribution as w. It thus suffices to study the distribution of (1); translations and changes of scale will provide the distributions of the general ratio x_1/y_1.

The set of points (x,y) for which $\frac{a + x}{b + y} < t$ is a region bounded by straight lines, and the normal probability measures of such regions have been extensively studied in the past few years. We should thus be able to express the distribution of w in terms of functions associated with those measures, particularly the bivariate normal distribution function

$$L(h,k,\rho) = P[\xi > h, \eta > k]$$

where ξ and η are standard normal with covariance ρ, and the V

function of Nicholson [11]:

$$V(h,q) = \int_0^h \int_0^{qx/h} \varphi(x)\varphi(y)dydx,$$

where φ is the standard normal density. We have

$$P[w < t] = P[a + x < t(b + y), b + y > 0] + P[a + x > t(b + y), b + y < 0]$$

$$= P[-x + ty > a - bt, y > -b] + P[x - ty > -a + bt, y > b]$$

$$= L\left(\frac{a - bt}{\sqrt{1 + t^2}}, -b, \frac{t}{\sqrt{1 + t^2}}\right) + L\left(\frac{-a + bt}{\sqrt{1 + t^2}}, b, \frac{t}{\sqrt{1 + t^2}}\right).$$

Then using the elementary properties of the L and V functions (see, for example, the NBS table [10], p. vii),

$$L(-h,-k,\rho) = L(h,k,\rho) + \int_0^h \varphi(x)dx + \int_0^k \varphi(x)dx,$$

$$L(-h,-k,\rho) + L(h,k,\rho) = 2V\left(h, \frac{k - \rho h}{\sqrt{1 - \rho^2}}\right) + 2V\left(k, \frac{h - \rho k}{\sqrt{1 - \rho^2}}\right) + \frac{1}{2} + \frac{\sin^{-1}\rho}{\pi},$$

we have several representations of $F(t) = P[\frac{a + x}{b + y} < t]$:

$$(2)\quad F(t) = L\left(\frac{a - bt}{\sqrt{1 + t^2}}, -b, \frac{t}{\sqrt{1 + t^2}}\right) + L\left(\frac{-a + bt}{\sqrt{1 + t^2}}, b, \frac{t}{\sqrt{1 + t^2}}\right),$$

$$(3)\quad F(t) = \int_0^{(bt-a)/\sqrt{1+t^2}} \varphi(x)dx + \int_0^b \varphi(x)dx + 2L\left(\frac{bt - a}{\sqrt{1 + t^2}}, b, \frac{t}{\sqrt{1 + t^2}}\right),$$

$$(4)\quad F(t) = \frac{1}{2} + \frac{1}{\pi}\tan^{-1}t + 2V\left(\frac{bt - a}{\sqrt{1 + t^2}}, \frac{b + at}{\sqrt{1 + t^2}}\right) - 2V(b,a).$$

Representation (4) appears best for numerical purposes, unless b

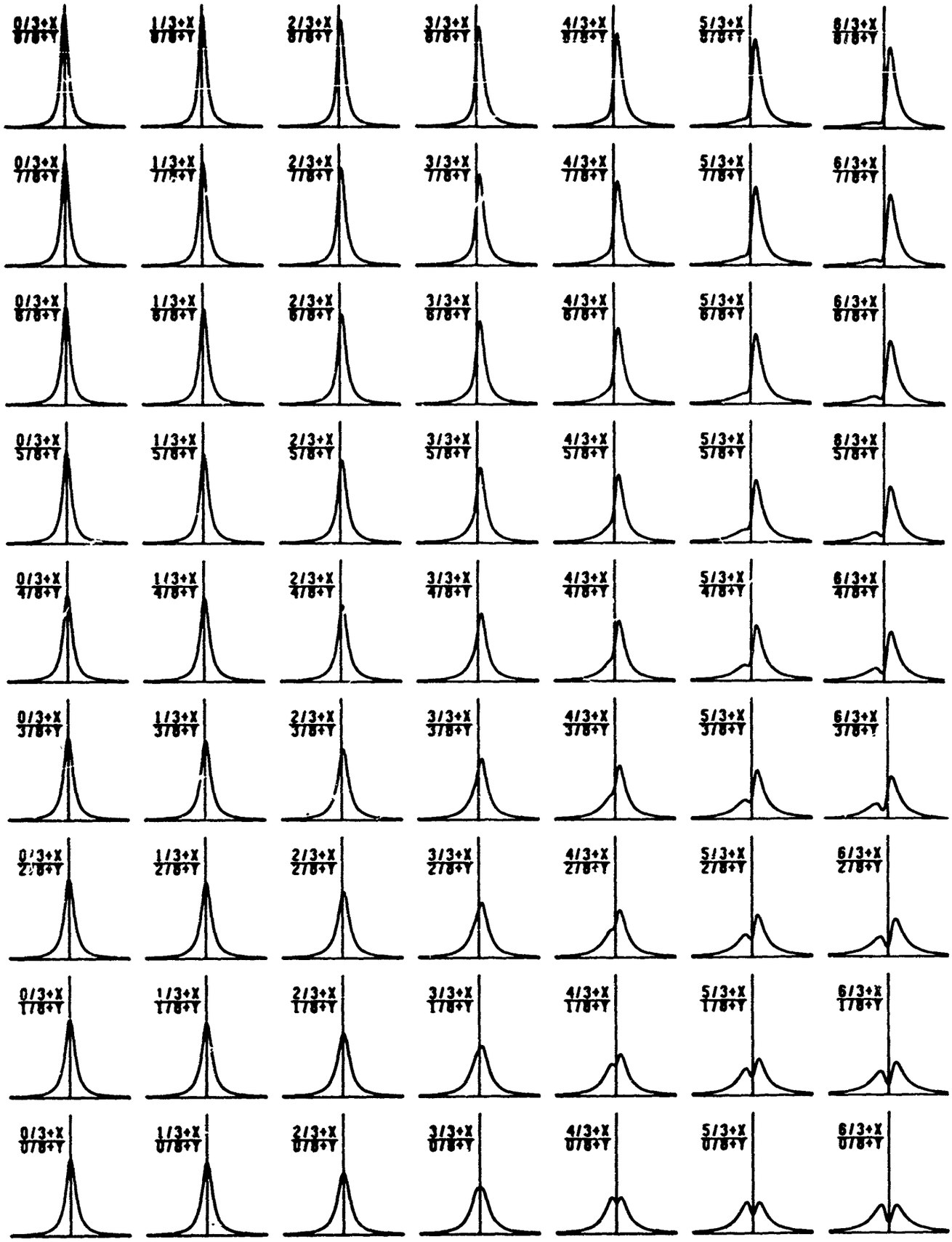

Figure 1. Graphs of the density of $(a + x)/(b + y)$, where $a > 0$, $b > 0$ and x,y are independent, standard normal random variables. The formula for the density is in equation (5). Values $a = 0/3,1/3,\ldots,6/3$ and $b = 0/8,1/8,\ldots,8/8$ were chosen so as to represent the possible shapes of the density function

is large, say $b > 3$, since we have good methods for providing values of V and $\frac{1}{\pi}\tan^{-1}t$, [9],[10], and [13]. However, when b is large, the second and third terms of (3) may be replaced by .5 and 0, so that

$$P[\frac{a+x}{b+y} < t] \cong \frac{1}{2} + \int_0^{(bt-a)\sqrt{1+t^2}} \varphi(x)dx = \int_{-\infty}^{(bt-a)\sqrt{1+t^2}} \varphi(x)dx$$

provides very good numerical approximations to $F(t)$, plus the additional information that $(bw - a)\sqrt{1 + w^2}$ is approximately normally distributed.

Now we turn to the density of $(a + x)/(b + y)$. Let

$$h = \frac{bt - a}{\sqrt{1 + t^2}}, \quad q = \frac{b + at}{\sqrt{1 + t^2}}, \quad \lambda = \frac{q}{h} = \frac{b + at}{b - at}.$$

Using primes to indicate differentiation with respect to t, so that $h' = q/(1 + t^2)$, $\lambda' = -(a^2 + b^2)/(bt - a)^2$, we differentiate (4) to get

$$f(t) = F' = \frac{1}{\pi(1 + t^2)} + 2h'\varphi(h)\int_0^q \varphi(y)dy + 2\lambda'\int_0^h x\varphi(x)\varphi(\lambda x)dx.$$

Integrating the last term and simplifying, we get this form for $f(t)$, the density function of the ratio $\frac{a + x}{b + y}$:

$$(5) \quad f(t) = \frac{e^{-.5(a^2+b^2)}}{\pi(1 + t^2)}[1 + \frac{q}{\varphi(q)}\int_0^q \varphi(y)dy], \qquad q = \frac{b + at}{\sqrt{1 + t^2}}.$$

Figure 1 shows $f(t)$, the density of $(a + x)/(b + y)$, for various values of a and b. The curves in Figure 1 were drawn by a computer; it also drew the identification for each density in the form $\frac{a + x}{b + y}$, where a is a multiple of $\frac{1}{3}$ and b a multiple of $\frac{1}{8}$. The values of a and

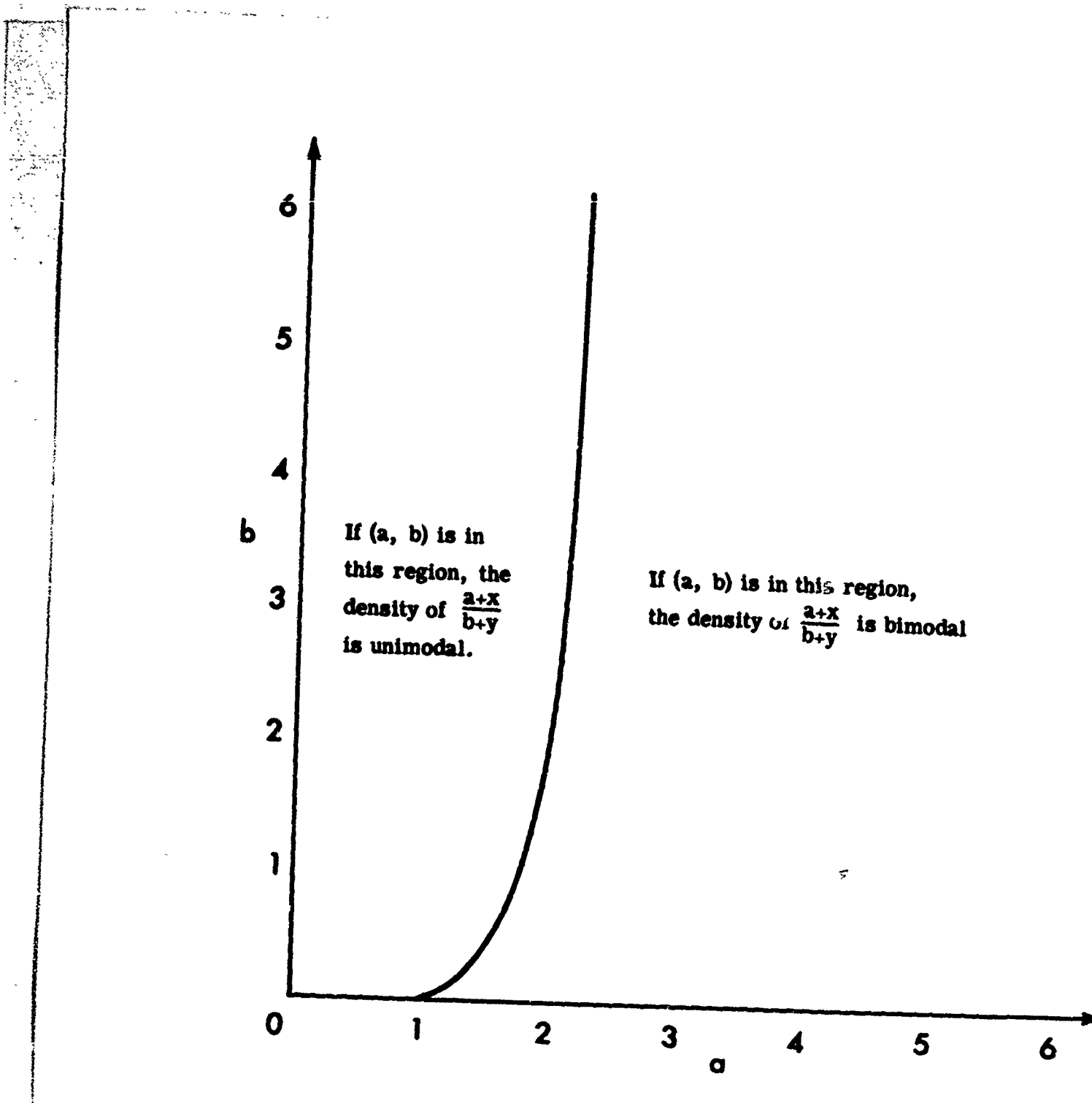

Figure 2. The density of $(a + x)/(b + y)$ is unimodal or bimodal according to the region of the positive quadrant in which the point (a,b) falls.

b were chosen so as to give a rough indication of the possible shapes of the densities given by formula (5). As you can see, some unusual shapes are encountered. The positive a,b quadrant may be divided into two regions according to whether the density of $(a + x)/(b + y)$ is unimodal or bimodal, as in Figure 2. The curve that determines the two regions is asymptotic to $a \cong 2.257$. Thus when $a > 2.257$, the density of $(a + x)/(b + y)$ is bimodal, even though it may not appear so. For example, the density of $(10^6 + x)/(10^6 + y)$, x and y independent standard normal, would appear to be a single spike at $t = 1$, but in fact it has another mode somewhere in the vicinity of $t = -10^{12}$.

We conclude this Section with a summary.

Summary of the properties of the ratio $w = \frac{a + x}{b + y}$, where $a > 0$, $b > 0$, x and y independent standard normal.

1. If $w' = x_1/y_1$ is the ratio of any two jointly normal variables, then there are constants c_1 and c_2 so that $c_1 + c_2 w$ has the same distribution as w.

2. The distribution of w, say $F(t) = P[\frac{a + x}{b + y} < t]$, may be expressed in terms of the bivariate normal distribution, or Nicholson's V function in several ways - formulas (2), (3), and (4) above.

3. When b is large, say $b > 3$, then $(bw - a)/\sqrt{1 + w^2}$ is approximately normally distributed, and

$$P[w \leq t] = P[\frac{a + x}{b + y} \leq t] \cong \int_{-\infty}^{(bt-a)/\sqrt{1+t^2}} \varphi(u)du.$$

4. The density of $\frac{a + x}{b + y}$ is given by formula (5). This density is plotted for various a and b in Figure 1.

5. The density of $\frac{a + x}{b + y}$ is unimodal or bimodal according to the region of Figure 2 in which (a,b) lies. When $a > 2.257$, the density is bimodal, although one of the modes may be insignificant.

3. The distribution of $c_1u_1 + \cdots + c_nu_n$.

Let $u_1,\ldots,u_n$ be independent random variables, each uniformly distributed over the interval $(0,1)$. In the next Section we will need the distribution of a linear contribution of the u's,

$$c_1u_1 + c_2u_2 + \cdots + c_nu_n \tag{6}$$

with the c's positive. The general linear form in the u's can readily be reduced to (6), for example

$$3u_1 - 2u_2 + 5u_3$$

has the same distribution as

$$3u_1 - 2(1 - u_2) + 5u_3 = 3u_1 + 2u_2 + 5u_3 - 2,$$

since $1 - u_2$ has the same distribution as u_2.

There have been a number of discussions of the distribution of (6) in the literature - the problem (for equal c's) dates back to Laplace [7], who solved it as a limiting form of the discrete case*,and, again with equal c's, the result is in standard textbooks, e.g., Uspensky [17], who inverted the characteristic function,and Cramer [1], proof by successive convolution. For unequal c's the result was given by Olds [12], and the distribution appeared as a problem on volumes, [3], with subsequent remarks on its proof - particularly a development of Schoenberg [15], using recursive relations for spline curves.

*The discrete case of the problem, which may be viewed as the problem of finding the sum on n "dice", each one having a certain number of faces, has an even more curious history. In 1710 Montmart solved the problem for equal dice, as did DeMoivre in 1711, Simpson in 1740, LaGrange around 1770, and Laplace in 1774. Montmart attempted, but did not solve, the problem of unequal dice. See Todhunter's History [16],Articles 148,149,364,888,915,987.

More recently, Roach [14], offered a geometric argument.

Thus the problem is now well known, and it is not particularly difficult, although notational difficulties, plus the fact that the problem may be viewed as one of probability, geometry, or spline functions, have led to a variety of proofs.

Roughly, the distribution of $c_1u_1 + \cdots + c_nu_n$ may be described as follows: Let S be the set of all 2^n numbers which can be formed as a sum of different c's:

$$S = \{0, c_1, \ldots, c_n, c_1 + c_2, \ldots, c_1 + \cdots + c_n\}.$$

Then

$$P[c_1u_1 + \cdots + c_nu_n < a] = \frac{1}{n!c_1c_2\cdots c_n} \sum_{s \in S, s<a} \pm (a - s)^n,$$

the $+$ or $-$ being according to whether there are an even or odd number of c's used to form s. For example,

$$P(2u_1 + 3u_2 + 8u_3 < 7] = \frac{1}{3!(48)}[7^3 - (7 - 2)^3 - (7 - 3)^3 + (7 - 5)^3]$$

and

$$(7) \quad P[2u_1 + 3u_2 + 8u_3 < 12] = \frac{1}{3!(48)}[12^3 - (12 - 2)^3 - (12 - 3)^3 - (12 - 8)^3 + (12 - 5)^3 + (12 - 10)^3 + (12 - 11)^3].$$

Note also that the distribution of $2u_1 + 3u_2 + 8u_3$ is symmetric (any linear combination of independent symmetric random variables is symmetric), and that, rather than compute expression (7), one might

consider

$$P[2u_1 + 3u_2 + 8u_3 < 12] = P[2(1 - u_1) + 3(1 - u_2) + 8(1 - u_3) < 12]$$

$$= P[2u_1 + 3u_2 + 8u_3 > 1] = 1 - \frac{1}{3!(48)} .$$

We may formally describe the distribution of $c_1u_1 + \cdots + c_nu_n$ as follows:

Theorem 1. Let $u_1,u_2,u_3,\ldots,u_n$ be independent random variables, each uniformly distributed over the interval $(0,1)$, and let $c_1,c_2,\ldots,c_n$ be positive constants. Let

$$F_n(a) = \text{Prob}[c_1u_1 + \cdots + c_nu_n \leq a]$$

and let

$$g_n(x) = \begin{cases} 0 & \text{if } x \leq 0, \\ \dfrac{x^n}{n!c_1c_2\cdots c_n} & \text{if } 0 < x. \end{cases}$$

Then, for $0 < a < c_1 + \cdots + c_n$,

$$F_n(a) = g_n(a) - \sum_i g_n(a-c_i) + \sum_{i<j} g_n(a-c_i-c_j) - \sum_{i<j<k} g_n(a-c_i-c_j-c_k) + \cdots .$$

The theorem may be easily proved by induction, using the elementary results:

$$F_{n+1}(a) = \frac{1}{c_{n+1}} \int_0^{c_{n+1}} F_n(a - x)dx$$

and

$$\frac{1}{c_{n+1}} \int_0^{c_{n+1}} g_n(b - x)dx = g_{n+1}(b) - g_{n+1}(b - c_{n+1}).$$

When the c's are all equal to 1, the result takes the following form:

$$P[u_1 + \cdots + u_n < a] = \frac{1}{n!}[a^n - \binom{n}{1}(a-1)^n + \binom{n}{2}(a-2)^n \cdots],$$

where the terms are taken as long as $a, a-1, a-2, \cdots,$ are positive. More formally, for $0 \leq a \leq n$, and with the greatest integer notation,

$$P[u_1 + \cdots + u_n < a] = \frac{1}{n!} \sum_{i=0}^{[a]} (-1)^i (a-i)^n.$$

4. The distribution of $\frac{u_1 + \cdots + u_n}{v_1 + \cdots + v_m}$.

Let $u_1, u_2, \cdots, u_n, v_1, \cdots, v_m$ be independent random variables, each uniform over $(0,1)$. We want the distribution of

$$\frac{u_1 + \cdots + u_n}{v_1 + \cdots + v_m}. \tag{8}$$

The distribution of (8) is of interest in studying round-off error propagation in numerical analysis, see [6], [18]. The particular case $m = n = 2$ was worked out in detail in [8]. We will find the distribution of (8) for all n and m, by applying the results of the previous Section, and will, in addition, discuss approximations to the distribution.

Since $1 - v_i$ is distributed as v_i, we have

$$P[\frac{u_1 + \cdots + u_n}{v_1 + \cdots + v_m} < a] = P[\frac{u_1 + \cdots + u_n}{(1 - v_1) + \cdots + (1 - v_m)} < a]$$

$$= P[u_1 + \cdots + u_n + av_1 + av_2 + \cdots + av_m < ma]$$

and hence a direct application of Theorem 1 gives, (after a little thought about how the terms combine):

$$P[\frac{u_1 + \cdots + u_n}{v_1 + \cdots + v_m} < a] = \frac{1}{(n+m)!a^m} \sum_{i=0}^{[ma]} \sum_{j=0}^{[(ma-i)/a]} (-1)^{i+j} \binom{n}{i}\binom{m}{j}[(m-j)a - i]^{n+m}.$$

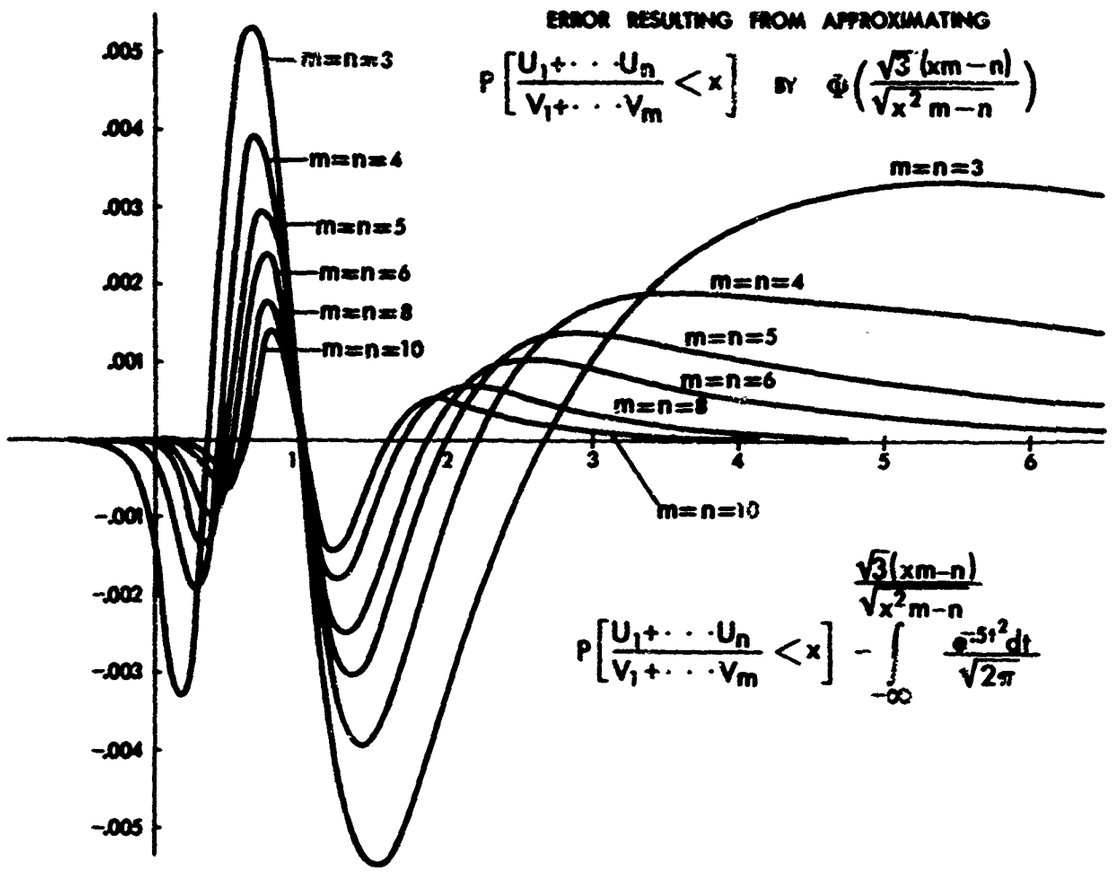
ERROR RESULTING FROM APPROXIMATING
$P\left[\frac{U_1+\cdots U_n}{V_1+\cdots V_m}<x\right]$ BY $\Phi\left(\frac{\sqrt{3}(xm-n)}{\sqrt{x^2m-n}}\right)$
.005
.004
.003
.002
.001
-.001
-.002
-.003
-.004
-.005
m=n=3
m=n=4
m=n=5
m=n=6
m=n=8
m=n=10
1
2
3
4
5
6
$P\left[\frac{U_1+\cdots U_n}{V_1+\cdots V_m}<x\right] - \int_{-\infty}^{\frac{\sqrt{3}(xm-n)}{\sqrt{x^2m-n}}} \frac{e^{-5t^2}dt}{\sqrt{2\pi}}$

Figure 3

For example,

$$P\left[\frac{u_1+\cdots+u_7}{v_1+\cdots+v_5}<.9\right]=\frac{1}{12!(.9)^5}\left\{\begin{array}{l}\binom{7}{0}\left[\binom{5}{0}(4.5)^{12}-\binom{5}{1}(3.6)^{12}+\binom{5}{2}(2.7)^{12}-\binom{5}{3}(1.8)^{12}+\binom{5}{4}(.9)^{12}\right]\\ -\binom{7}{1}\left[\binom{5}{0}(3.5)^{12}-\binom{5}{1}(2.6)^{12}+\binom{5}{2}(1.7)^{12}+\binom{5}{3}(.8)^{12}\right]\\ +\binom{7}{2}\left[\binom{5}{0}(2.5)^{12}-\binom{5}{2}(1.6)^{12}+\binom{5}{3}(.7)^{12}\right]\\ -\binom{7}{3}\left[\binom{5}{0}(1.5)^{12}-\binom{5}{2}(.6)^{12}\right]\\ +\binom{7}{4}\left[\binom{5}{0}(.5)^{12}\right]\end{array}\right\}$$

The variate $(u_1 +\cdots+ u_n)/(v_1 +\cdots+ v_m)$ is approximately a ratio of independent normal variables, and the discussion of Section 2 should apply. We may derive a good normal approximation directly, however, writing

$$P\left[\frac{u_1 +\cdots+ u_n}{v_1 +\cdots+ v_m} < a\right] = P[u_1 +\cdots+ u_n + av_1 +\cdots+ av_m < ma].$$

Since the sum on the right is approximately normal with mean $.5[n + ma]$ and variance $(a^2m + n)/12$, we have

$$P\left[\frac{u_1 +\cdots+ u_n}{v_1 +\cdots+ v_m} < a\right] \approx \Phi\left[\frac{\sqrt{3}(am - n)}{\sqrt{a^2m + n}}\right] .$$

Figure 3 gives some indication of the merits of this approximation.

In case it is necessary to get the tail of the distribution with great precision, it is not too difficult to calculate the exact probabilities: for $0 < a < \frac{1}{m}$,

$$P\left[\frac{u_1 +\cdots+ u_n}{v_1 +\cdots+ v_m} < a\right] = \frac{a^n}{(n + m)!}\left[m^{n+m} - \binom{m}{1}(m - 1)^{n+m} + \binom{m}{2}(m - 2)^{n+m} - \cdots\right],$$

and for $b > n$,

$$P\left[\frac{u_1 +\cdots+ u_n}{v_1 +\cdots+ v_m} > b\right] = \frac{b^{-m}}{(n + m)!}\left[n^{n+m} - \binom{n}{1}(n - 1)^{n+m} + \binom{n}{2}(n - 2)^{n+m} - \cdots\right].$$

REFERENCES

[1] Cramér, Harald, 1946, Mathematical Methods of Statistics, Princeton, pp. 244-246.

[2] Curtiss, J. H., 1943, On the Distribution of the Quotient of Two Chance Variables, Annals Math. Stat., V. 12, pp. 409-421.

[3] Eisenstein, Maurice and Klamkin, M. S., 1959, Problem 59-2, N-dimensional Volume, SIAM Review, Vol. 1, No. 1, p. 69.

[4] Fieller, E. C., 1932, The Distribution of the Index in a Normal Bivariate Population, Biometrika, Vol. 24, pp. 428-440.

[5] Geary, R. C., 1930, The Frequency Distribution of the Quotient of Two Normal Vairates, Journal Roy. Stat. Soc., Vol. 93, pp. 442-446.

[6] Inman, S., 1950, The Probability of a Given Error Being Excluded in Approximate Computations, Math. Gazette, Vol. 34, pp. 99-113.

[7] Laplace, P., 1812, Theorie Analytique des Probabilities, Paris, pp. 253-261.

[8] Locker, John and Perry, N. C., 1962, Probability Functions for Computations Involving More Than One Operation, Mathematics Magazine, Vol. 35, No. 2, pp. 87-89.

[9] Marsaglia, G., 1960, Tables of $\frac{1}{2\pi}\,\mathrm{Tan}^{-1}(\lambda)$ and $\mathrm{Tan}^{-1}(\lambda)$ for λ = .0001,.0002,...,.9999, with Some Remarks on Their Use in Finding the Normal Probability Measure of Polygonal Regions, Boeing Scientific Research Laboratories Document D1-82-0078.

[10] National Bureau of Standards,1959, Tables of the Bivariate Normal Distribution and Related Functions, Applied Math. Series 50, Washington, D. C.

[11] Nicholson, C., 1943, The Probability Integral for Two Variables, Biometrika, V. 33, pp. 59-72.

[12] Olds, E. G., 1952, A Note on the Convolutions of Normal Distributions, Annals Math. Stat., V. 23, pp. 282-285.

[13] Owen, D. B., 1956, Tables for Computing Bivariate Normal Probabilities, Annals Math. Stat., V. 27, pp. 1075-1090.

[14] Roach, S. A., 1963, The Frequency Distribution of the Sample Mean When Each Member of the Sample Is Drawn from a Different Rectangular Distribution, Biometrika, V. 50, pp. 508-513.

[15] Schoenberg, I. J., 1960, Solution to Problem 59-2, N-dimensional Volume, SIAM Review, V. 2, No. 1, pp. 41-45.

[16] Todhunter, I., 1865, A History of the Mathematical Theory of Probability, Chelsea Reprinted Edition, New York, 1949.

[17] Uspensky, J. V., 1937, Introduction to Mathematical Probability, New York, pp. 277-278.

[18] Woodward, R. S., 1906, Probability and Theory of Errors, New York.