5260 UVP

UNCLASSIFIED

NRL Report 6008

The Moments of a Sampled Binomially Distributed Variable

H. M. Suski

Security Systems and Avigation Branch
Electronics Division

December 2, 1963

U.S. NAVAL RESEARCH LABORATORY
Washington, D.C.

The Moments of a Sampled Binomially Distributed Variable

H. M. Suski

Security Systems and Avigation Branch
Electronics Division

A statistical variable consisting of two components is considered. One component has as its density function the universe of the variable; the other component, the distribution of the mean. The use of a two-component variable leads to the introduction of two orders of finite sampling.

For the case of a binomially distributed variable, the theory of moment-generating functions is used to derive equations for the first four moments about an arbitrary origin. The second (μ_2), third (μ_3), and fourth (μ_4) moments about the mean are obtained, as well as the moment numbers β_1, β_2, γ_1, and γ_2. The detailed steps of the derivations, including underlying assumptions, are given.

The results consist of derived equations and of graphs of μ_2, μ_3, and μ_4 as functions of p, the probability of success in a single trial. These graphs show the effect of both orders of sampling.

INTRODUCTION

Statistical methods provide the means for calculating moments, which are the parameters describing the distribution function of a statistical variable (Ref. 1, pp. 50-52). Functions which generate the moments (Ref. 1, pp. 54-55, 81-84, and 90-114; Ref. 2, pp. 26-28, and 35-43; Ref. 3, pp. 248-252) permit the comparison of distribution functions and their associated variables. Use is made, in this report, of moment-generating-function theory to derive expressions for the first four moments of a binomially distributed variable. Such expressions have been derived (Ref. 1, p. 52) for the case of finite sampling. The variable treated in this paper has two components; the two distribution functions associated with the components are the universe itself and the error in the mean. By including the second component of the variable, a second order of sampling is introduced. Thus, it becomes necessary to evaluate, using the law of large numbers, the effect of both orders of sampling. This evaluation is made in terms of an empirical problem in which the (constant) probability p in the case of Bernoulli trials is measured.

Equations for the first four moments about an arbitrary point, for the moment about the mean, and for the moment numbers are derived from fundamentals. Graphs of these equations are presented as the results.

In discussing the problem involved in measuring p, it is found that the choice of sizes for the two orders of sampling essentially establishes *a priori* an upper bound for the error in determining p. This result is demonstrated with a set of graphs.

NRL Problem R03-01; Projects RF 001-10-41-4150, SF 001-10-01-6108. This is an interim report on this problem; work is continuing. Manuscript submitted August 23. 1963.

GENERAL DISCUSSION

A Composite Statistical Variable $w = x + y$

Consider the statistical variable

$$w = x + y, \tag{1}$$

where x and y are two components into which the variable w can be separated. We will be concerned with finite sampling from a universe; let n be the sample size (the first order of sampling). The x component is considered to be the theoretically correct value of the variable, as if it were obtained from the universe without sampling error. The y component is introduced to account for the sampling error, resulting from uncertainty of the mean value because of finite sampling. The sampling error can be represented as a deviation from the mean or the expected value of the x component.

Probability Distribution

In general, for a continuous variable x having a density function $f(x)$, the relation

$$F(x)\Big|_a^b \triangleq \int_a^b f(x)dx \doteq Pr(a \leq x \leq b),\ (a \leq b), \tag{2}$$

is the probability that x will fall within the limits a and b when the sampling is random. The definition given by Eq. (2), called the distribution function, implies

$$F(x)|_{-\infty}^{x} \triangleq F(x) \triangleq \int_{-\infty}^{x} f(u)\,du = Pr(u \leq x), \quad (3)$$

and hence,

$$\int_{-\infty}^{\infty} f(u)\,du = F(\infty) - F(-\infty) = 1 - 0 = 1. \quad (4)$$

Note that

$$F'(x) \equiv \frac{dF(x)}{dx} = f(x) \quad (5)$$

or

$$dF(x) = f(x)\,dx. \quad (6)$$

Also,

$$\int_{a}^{b} f(u)\,du = F(x)|_{a}^{b} = \int_{a}^{b} dF(x) = \int_{a}^{b} f(x)\,dx. \quad (7)$$

Moments and Moment Generating Functions

The moments of the variable, which aid in describing its distribution function, are defined as

$$\nu_i = \int_{-\infty}^{\infty} (x-a)^i dF(x), \quad (8)$$

where i is the order of the specific moment and a is an arbitrary point about which moments are obtained. Similarly,

$$\mu_i = \int_{-\infty}^{\infty} (x-\bar{x})^i dF(x) \quad (9)$$

defines the i-th moment about the mean.

We introduce the parameter θ (following Ref. 2, pp. 26-28) and define the moment generating function for the variable x with the aid of Eq. (6) as

$$M_x(\theta) = \int_{a}^{b} e^{\theta x} dF(x) = \int_{a}^{b} e^{\theta x} f(x)\,dx, \quad (10)$$

where the subscript is used to indicate the specific variable or distribution function being considered. This permits a power series approximation to the moment generating function. The coefficients of the terms of appropriate order agree with results obtained using Eq. (8) or Eq. (9).

The moment generating function has the property that linear functional relations are easily handled. Thus, by changing the variable in Eq. (10), with an appropriate change in the limits of integration, we get

$$M_h(\theta) = \int_{a'}^{b'} e^{\theta h(x)} dF[h(x)]. \quad (11)$$

If in Eq. (11) we let $h(x) = cH(x)$, then

$$M_{cH}(\theta) = \int_{a''}^{b''} e^{\theta cH(x)} dF[cH(x)] = M_H(c\theta). \quad (12)$$

If we let $h(x) = c + H(x)$

$$M_{c+H}(\theta) = \int_{a'''}^{b'''} e^{\theta[c+H(x)]} dF[c + H(x)]$$

$$= e^{c\theta} M_H(\theta). \quad (13)$$

Expansion of $e^{\theta x}$ in Eq. (10) as a power series permits the integration to be carried out term by term:

$$M_x(\theta) = \int_{a}^{b} f(x)\,dx + \theta \int_{a}^{b} x f(x)\,dx + \frac{\theta^2}{2!} \int_{a}^{b} x^2 f(x)\,dx + \ldots. \quad (14)$$

But Eq. (14) can be rewritten as

$$M_x(\theta) = \nu_0 + \nu_1\theta + \nu_2 \frac{\theta^2}{2!} + \nu_3 \frac{\theta^3}{3!} + \ldots, \quad (15)$$

where the coefficient of $\theta^i/i!$ is the i-th moment about the origin $a = 0$. In general, the i-th moment about the origin, as obtained from the moment generating function, is given by

$$\nu_i = \int_{a}^{b} x^i f(x)\,dx. \quad (16)$$

Note that an advantage of Eq. (15) lies in the fact that if a moment generating function can be expanded in a power series similar to Eq. (15), then the i-th moment is automatically the coefficient of the i-th term in θ.

The i-th moment about the origin can be evaluated by computing the appropriate derivative of

UNCLASSIFIED

$M_x(\theta)$ at $\theta = 0$, since repeated differentiation of Eq. (15) shows that

$$\nu_i = \left.\frac{d^i M_x(\theta)}{d\theta^i}\right|_{\theta=0}. \tag{17}$$

The case of the discrete variable is obtained by analogy. Using Eq. (16) we get

$$\nu_i = \sum_{x=a}^{b} x^i Pr(x), \tag{18}$$

where each $Pr(x)$ is the probability that the discrete variable takes on the value of x chosen in the range from a to b.

Constant Probability of Bernoulli Trials

Our concern is with a discrete variable x which is binomially distributed. The binomial density function can be written (Ref. 3, p. 137) as

$$b(k,n,p) = Pr(x = k), \qquad (p \text{ is constant}). \tag{19}$$

It gives the probability that there are exactly k occurrences of the value x in n trials when the probability p of occurrence in a single trial is constant, the case of Bernoulli trials. Additionally, the binomial distribution function

$$B(k,n,p) = \sum_{x=k}^{n} b(x,n,p) = Pr(x \leq k) \tag{20}$$

gives the probability that there are k or more occurrences of x out of n trials. Furthermore, it follows that

$$B(k_1,k_2,n,p) = \sum_{x=k_1}^{n} b(x,n,p) - \sum_{x=k_2}^{n} b(x,n,p)$$

$$= Pr(k_1 \leq x \leq k_2). \tag{21}$$

Note that

$$Pr(x = k) = b(k,n,p) = \binom{n}{k} p^k q^{n-k}, \tag{22}$$

where

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}, \tag{23}$$

and that the zero-order moment is

$$\nu_0 = \sum_{k=0}^{n} b(k,n,p) = 1. \tag{24}$$

The Law of Large Numbers and the Upper Bound of Error

The number of trials or the sample size n can be estimated from the law of large numbers (Ref. 3, p. 141), which can be stated as

$$Pr\left(\left|\frac{k}{n} - p\right| < \epsilon\right) \to 1, \tag{25}$$

where k/n is the measured value of p and ϵ is the measuring error. From the law of large numbers we see that the measuring error becomes vanishingly small as n is increased without limit. It is, therefore, possible to make the measuring error as small as desired by appropriately selecting the value of n. This suggests that, by convention, we may select an upper bound of error *a priori*.

If tolerances in p are provided, count tolerances for k can be defined for Eq. (21) as follows:

$$k_1 = np_1 \text{ and } k_2 = np_2. \tag{26}$$

In general

$$p_1 \leq p \leq p_2, \tag{27}$$

i.e., the actual value of p will be a constant (for a stable process) and have a value between p_1 and p_2. The distribution of values of k for the general case is given by Eq. (19).

x-COMPONENT

Moment Generating Function

In effect we have been describing a two-level process which is permitted a long-period variation; *i.e.*, p is constant for at least the duration required to observe the result of n trials. In any sampling interval, p is expected to have a value within the limits given by Eq. (27).

We obtain describing characteristics of the distribution function of our statistical variable from the moment generating function given by Eq. (15) or Eq. (18). In Eq. (18) the function $Pr(x)$ is given by Eq. (22), and the limits $a = 0$, $b = \infty$ are of interest to us. This choice of limits assures us of considering all possible values of x.

We note that

$$\sum_{x=0}^{n} Pr(x) = \sum_{x=0}^{n} \binom{n}{x} p^x q^{n-x} = (q+p)^n, \quad (28)$$

which is the binomial theorem. As n increases beyond limit, *i.e.*, $n \to \infty$,

$$\sum_{x=0}^{n} Pr(x) \to \sum_{x=0}^{\infty} Pr(x) \quad (29)$$

and

$$\sum_{x=0}^{n} e^{\theta x} Pr(x) \to M_x(\theta). \quad (30)$$

Since there is no inconsistency in Eqs. (28) - (30), it follows that for all n the binomial moment generating function is

$$M_x(\theta) = \sum_{x=0}^{n} e^{\theta x} \binom{n}{x} p^x q^{n-x}$$

$$= \sum_{k=0}^{n} e^{\theta k} \binom{n}{k} p^k q^{n-k} = M_k(\theta)$$

$$= \sum_{k=0}^{n} \binom{n}{k} (pe^{\theta})^k q^{n-k}, \quad (31)$$

or

$$M_k(\theta) = (q + pe^{\theta})^n. \quad (32)$$

Moment Equations

Successive differentiations of Eq. (32) yield the following:

$$D^1 M_k(\theta) = npe^{\theta}(q + pe^{\theta})^{n-1} \quad (33)$$

$$D^2 M_k(\theta) = npe^{\theta}(q + pe^{\theta})^{n-1} + n(n-1)p^2 e^{2\theta}(q + pe^{\theta})^{n-2} \quad (34)$$

$$D^3 M_k(\theta) = npe^{\theta}(q + pe^{\theta})^{n-1} + 3n(n-1)p^2 e^{2\theta}(q + pe^{\theta})^{n-2} + n(n-1)(n-2)p^3 e^{3\theta}(q + pe^{\theta})^{n-3} \quad (35)$$

$$D^4 M_k(\theta) = npe^{\theta}(q + pe^{\theta})^{n-1} + 7n(n-1)p^2 e^{2\theta}(q + pe^{\theta})^{n-2} + 6n(n-1)(n-2)p^3 e^{3\theta}(q + pe^{\theta})^{n-3} + n(n-1)(n-2)(n-3)p^4 e^{4\theta}(q + pe^{\theta})^{n-4}. \quad (36)$$

The evaluation of these four derivatives at the origin ($\theta = 0$, $e^{\theta} = 1$, $q + p = 1$) yields the first four moments about the origin, as follows:

$$\nu_1 = np \quad (37)$$

$$\nu_2 = np + n(n-1)p^2 = npq + (np)^2 \quad (38)$$

$$\nu_3 = np + 3n(n-1)p^2 + n(n-1)(n-2)p^3 \quad (39)$$

$$\nu_4 = np + 7n(n-1)p^2 + 6n(n-1)(n-2)p^3 + n(n-1)(n-2)(n-3)p^4. \quad (40)$$

By virtue of Eq. (13) moments about the mean can be obtained. The following relation (Ref. 2, p. 18) permits translation from the origin to the mean.

$$\mu_i = \nu_i - \nu_{i-1}\nu_1 + \frac{i(i-1)}{2}\nu_{i-2}\nu_1^2 - \ldots + (-1)^i \nu_1^i. \quad (41)$$

Using Eqs. (37) through (40) and applying Eq. (41), we obtain the following for the first four moments about the mean:

$$\mu_1 = \nu_1 - \nu_1 = 0 \quad (42)$$

$$\mu_2 = \nu_2 - \nu_1^2 = npq \quad (43)$$

$$\mu_3 = \nu_3 - 3\nu_2\nu_1 + 2\nu_1^3 = npq(q-p) \quad (44)$$

$$\mu_4 = \nu_4 - 4\nu_3\nu_1 + 6\nu_2\nu_1^2 - 3\nu_1^4 = 3(npq)^2 + npq(1 - 6pq). \quad (45)$$

It will be recalled that the first moment ν_1, called the mean or expectation, is a measure of central tendency. The second moment μ_2, the variance, is a measure of variation about the mean. The third and fourth moments, measures of the shape of the distribution function, are indicators of skewness and peakedness, respectively. It will be noted that the degree of the terms involved increases as the order of the moment is increased.

Lower Degree Measures or Moment Numbers (β, γ)

In Eqs. (43)-(45) the second, third, and fourth degrees, respectively, of p are involved. It will be

noted that the third and fourth moments are functions of the first and second moments. The second moment involves the square of the deviations from the mean; no distinction is made between deviations above or below the mean. This useful property is retained, and a first degree measure is obtained if we define

$$\sigma = \sqrt{\mu_2}, \tag{46}$$

where σ is called the standard deviation. Since the third moment involves the third power of the deviation from the mean, a new measure involving the square of the third moment will have the property of being independent of the sign of the deviation from the mean; thus, we derive two zero-degree, or third-moment numbers:

$$\begin{aligned} \beta_1 &= \mu_3^2/\mu_2^3 = (\mu_3/\sigma^3)^2; \\ \gamma_1 &= \sqrt{\beta_1} = \mu_3/\sigma^3. \end{aligned} \tag{47}$$

Similarly, we also define two new zero-degree measures, or fourth-moment numbers:

$$\beta_2 = \mu_4/\mu_2^2 = \mu_4/\sigma^4;\ \gamma_2 = \beta_2 - 3. \tag{48}$$

For a normal (Gaussian) distribution the odd moments are zero, as is also γ_2. The measure γ_2 can be used in making comparisons with the normal distribution.

We have mentioned something about the first four moments. Looking back at Eq. (14) we note that the first, or zero-degree, term has been ignored. By inspection we see that the zero-degree term gives the area of the distribution function; Eq. (24) is the normalized form of this function.

y-COMPONENT

Second Order of Sampling and the Moment Generating Function

To determine the moments it is necessary to make calculations as indicated in Eq. (18). In the binary case this amounts to getting the frequency count k of occurrence for a sample of size n. A single measurement (n trials) yields a single value of k in Eq. (31). We should expect, by virtue of Eq. (25), that

$$k/n \to p \text{ or } k \to np \tag{49}$$

within the error ϵ which has been selected as acceptable. If, however, p lies between limits as given by Eq. (27), then we can expect variations caused by tolerances in p as well as variations from the expected limits of error.

We, therefore, consider a set of measurements, represented symbolically by

$$K = \{k_j\}. \tag{50}$$

Every element of the set K is obtained from the same distribution [Eq. (21)], and each has the same moment generating function [Eq. (32)]. It will be assumed that the set is finite and has N members, *i.e.*, there are N measurements taken. We have, therefore, introduced the second order of sampling, the set K. Each of the N elements k_j, is a sampling of the first order obtained after n trials. The first order of sampling sets the measuring precision [Eq. (25)]; the second order of sampling yields information about the distribution of measured values, *i.e.*, the sampling error.

We are interested in finding characteristics which describe the membership of the set K. Consider the sum of the members of the set. From the sum we obtain the average as follows:

$$\bar{k} = \frac{1}{N}\sum_{j=1}^{N} k_j = \frac{k_1 + k_2 + \ldots + k_N}{N}. \tag{51}$$

The assumption can be made that K has members which were obtained independently. The assumption of independence and the use of Eqs. (11) and (12) permit us to obtain the moment generating function of the mean of the second-order-sampling set K. We find that

$$\begin{aligned} M_{\bar{k}}(\theta) &= M_{[(k_1+k_2+\ldots+k_N)/N]}(\theta) \\ &= M_{(k_1+\ldots+k_N)}\left(\frac{\theta}{N}\right) \\ &= M_{k_1}\left(\frac{\theta}{N}\right) M_{k_2}\left(\frac{\theta}{N}\right) \ldots M_{k_N}\left(\frac{\theta}{N}\right) \\ &= \left(M_k{}^N\right)\left(\frac{\theta}{N}\right), \end{aligned} \tag{52}$$

and using Eq. (32), we finally obtain

$$M_{\bar{k}}(\theta) = (q + pe^{\theta/N})^{nN}. \tag{53}$$

From Eq. (37) and (42) we find that the mean of the universe is np. We are interested in deviations from the mean, *i.e.*, in the y component or the variable.

$$y = \bar{k} - np. \tag{54}$$

Using Eq. (13), we obtain

$$M_y(\theta) = M_{\bar{k}-np}(\theta) = e^{-np} M_{\bar{k}}(\theta)$$
$$= e^{-np} M_k{}^N \left(\frac{\theta}{N}\right). \tag{55}$$

THE COMPLETE VARIABLE $w = x + y$

Moment Generating Function

Having an expression for the moment generating function of the two components of the variable w [Eqs. (32) and (55)], we can obtain the moment generating function $M_w(\theta)$. We make use of another property of moment generating functions (Ref. 2, p. 63) which can be expressed as follows:

$$M_w(\theta) = M_{(x+y)}(\theta) = M_x(\theta) M_y(\theta), \tag{56}$$

i.e., the moment generating function of the sum of *independent* variables is equal to the product of their individual moment generating functions.

We obtain the following as the expression for the moment generating function of a sampled, binomially distributed variable:

$$M_w(\theta) = (q + pe^{\theta})^n e^{-np} (q + pe^{\theta/N})^{nN}. \tag{57}$$

Let us take one further step. We desire the moments about some arbitrary point a. Hence,

$$M_{w-a}(\theta) = e^{-(np+a)\theta} (q + pe^{\theta})^n (q + pe^{\theta/N})^{nN}. \tag{58}$$

Moment Equations

Differentiation of Eq. (58) gives

$$\frac{dM_{w-a}(\theta)}{d\theta} =$$
$$-(np+a)\, e^{-(np+a)\theta} (q + pe^{\theta})^n (q + pe^{\theta/N})^{nN}$$
$$+ npe^{-(np+a-1)\theta} (q + pe^{\theta})^{n-1} (q + pe^{\theta/N})^{nN}$$
$$+ npe^{-(np+a-1/N)} (q + pe^{\theta})^n (q + pe^{\theta/N})^{nN-1}. \tag{59}$$

The first moment is obtained by setting $\theta = 0$ in Eq. (59):

$$\nu_1(w - a) = np - a. \tag{60}$$

Differentiating Eq. (59) and setting $\theta = 0$, we obtain the second moment:

$$\nu_2(w - a) = (np - a)^2 + npq + (n/N)\, pq. \tag{61}$$

Using Eq. (41), a parallel case of Eq. (43) results:

$$\mu_2 = \nu_2 - \nu_1{}^2$$

or

$$\mu_2(w - a) = npq + (n/N)\, pq. \tag{62}$$

Differentiating once again and setting $\theta = 0$, we obtain the third moment:

$$\nu_3(w - a) = (np - a)^3$$
$$+ 3(np - a)(npq + (n/N)\, pq)$$
$$+ (1 - 2p)(npq + (n/N^2)\, pq). \tag{63}$$

Using Eq. (41), a parallel case of Eq. (44) results:

$$\mu_3 = \nu_3 - 3\nu_2\nu_1 + 2\nu_1{}^3$$

or

$$\mu_3(w - a) = (1 - 2p)(npq + (n/N^2)\, pq). \tag{64}$$

Differentiating for the fourth time and setting $\theta = 0$, we obtain

$$\nu_4(w - a) = (np - a)^4$$
$$+ 6(np - a)^2 (npq + (n/N)\, pq)$$
$$+ 4(np - a)(npq + (n/N^2)\, pq)(1 - 2p)$$
$$+ 3\,(npq + n/Npq)^2$$
$$+ (npq + (n/N^3)pq)(1 - 6pq). \tag{65}$$

Using Eq. (41), a parallel case of Eq. (45) results:

$$\mu_4 = \nu_4 - 4\nu_3\nu_1 + 6\nu_2\nu_1{}^2 - 3\nu_1{}^4$$

or

$$\mu_4(w - a) = 3(npq + (n/N)\, pq)^2$$

$$+ (1-6pq)(npq + (n/N^3)\, pq). \qquad (66)$$

Using the above derived expressions for the moments in Eqs. (47) and (48), we obtain the following moment numbers:

$$\beta_1 = \frac{(1-2p)^2(N^2+1)^2}{npqN(N+1)^3} \qquad (67)$$

$$\gamma_1 = \frac{(1-2p)(N^2+1)}{\sqrt{npqN(N+1)^3}} \qquad (68)$$

$$\beta_2 = 3 + \frac{1-6pq}{pq}\,\frac{N^3+1}{nN(N+1)^2} \qquad (69)$$

$$\gamma_2 = \frac{1-6pq}{pq}\,\frac{N^3+1}{nN(N+1)^2}. \qquad (70)$$

Upper Bound of Error

The DeMoivre-Laplace limit theorem (Ref. 3, pp. 168-173) can be stated as follows:

$$Pr(a \leq k \leq b) \sim \Phi(x_b) - \Phi(x_a), \qquad (71)$$

where

$$\Phi(x_b) = \int_{-\infty}^{x_b} \varphi(u)\,du.$$

Let us first derive a result for the x-component of the variable w; then we shall consider w.

Let

$$a = n(p-\epsilon),\ b = n(p+\epsilon),$$

$$x_a = \frac{a}{\sigma} = \frac{n(p-\epsilon)}{\sqrt{(npq)}},$$

$$x_b = \frac{b}{\sigma} = \frac{n(p+\epsilon)}{\sqrt{(npq)}}.$$

Hence,

$$Pr\ (n(p-\epsilon) \leq k \leq n(p+\epsilon))$$

$$\sim \Phi\left(\frac{np+n\epsilon}{\sqrt{(npq)}}\right) - \Phi\left(\frac{np-n\epsilon}{\sqrt{(npq)}}\right),$$

or

$$Pr\ (np - n\epsilon - np \leq k - np \leq np + n\epsilon - np)$$

$$\sim \Phi\left(\frac{n\epsilon}{\sqrt{(npq)}}\right) - \Phi\left(-\frac{n\epsilon}{\sqrt{(npq)}}\right) =$$

$$Pr\ (-n\epsilon \leq k - np \leq n\epsilon)$$

$$\sim \Phi\left(\epsilon\sqrt{\frac{n}{pq}}\right) - \Phi\left(-\epsilon\sqrt{\frac{n}{pq}}\right),$$

or, finally,

$$Pr\left(\left|\frac{k}{n} - p\right| \leq \epsilon\right) \sim \Phi\left(\epsilon\sqrt{\frac{n}{pq}}\right) - \Phi\left(-\epsilon\sqrt{\frac{n}{pq}}\right). \qquad (72)$$

Note that Eq. (72) leads to Eq. (25) as $n \to \infty$.

In the case of the variable w, which has a variance $npq\ [1 + (1/N)]$, let the upper bound of error be δ. Then,

$$Pr\left(\left|\frac{w}{n} - p\right| \leq \delta\right) \sim \Phi\left(\frac{\delta}{1+(1/N)}\sqrt{\frac{n}{pq}}\right)$$

$$-\Phi\left(-\frac{\delta}{1+(1/N)}\sqrt{\frac{n}{pq}}\right). \qquad (73)$$

The right-hand side of Eqs. (72) and (73) is of the form

$$\Phi(t) - \Phi(-t). \qquad (74)$$

Now

$$\Phi(-t) = 1 - \Phi(t); \qquad (75)$$

so

$$\Phi(t) - \Phi(-t) = 2\Phi(t) - 1. \qquad (76)$$

If we assign a confidence limit CL as follows

$$Pr\left(\left|\frac{w}{n} - p\right| \leq \delta\right) \equiv CL, \qquad (77)$$

then

$$\Phi(t) = \frac{1+CL}{2}, \qquad (78)$$

from which t in Eqs. (72) and (73) can be evaluated.

GRAPHICAL RESULTS

The effect of the two orders of sampling on the values of the moments can be shown graphically. By an appropriate choice of scales, it is possible to present the results numerically in a general way. Figures 1 through 4 show the graphs

UNCLASSIFIED

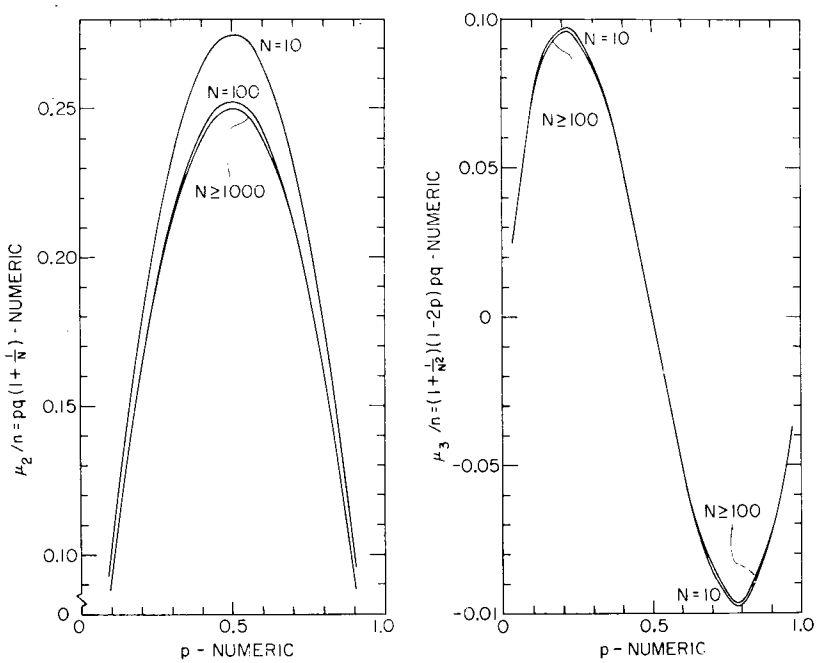

Fig. 1—The effect of two orders of sampling on the second (μ_2) and third (μ_3) moments about the mean. The generalized ordinate scales must be multiplied by n to obtain the moments. In the moment equations, shown on the ordinate axes, n refers to the size of the first sampling; N refers to the size of the second sampling.

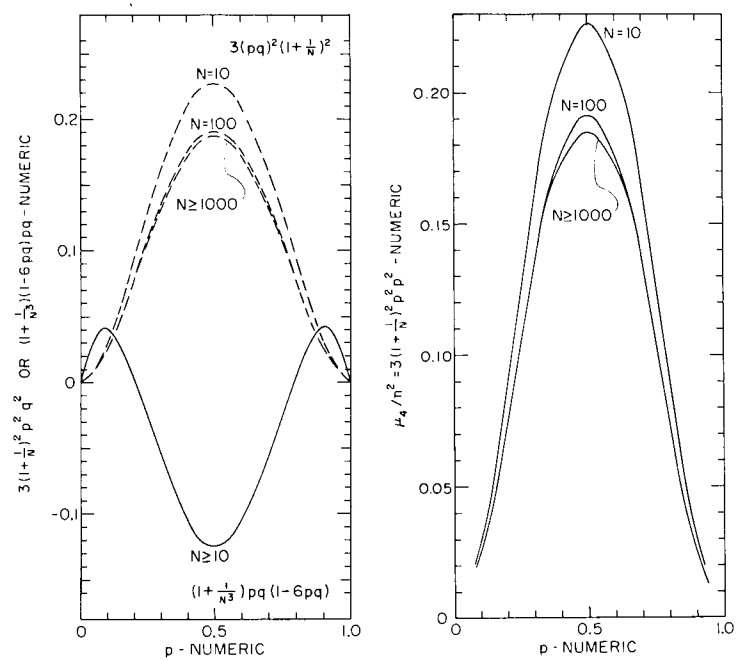

Fig. 2—The effect of two orders of sampling on the fourth (μ_4) moment about the mean. The left-hand figure shows the individual components which make up the expression for the fourth moment; the right-hand figure shows their combined result for large values of n. The generalized scale on the component figure must be multiplied by n for the one component and by n^2 for the other component. In the combined graph an approximately generalized scale is used, but it applies only when n is large. This scale must be multiplied by n^2 to obtain the moment. The symbol n refers to the first sample size; N refers to the second sample size.

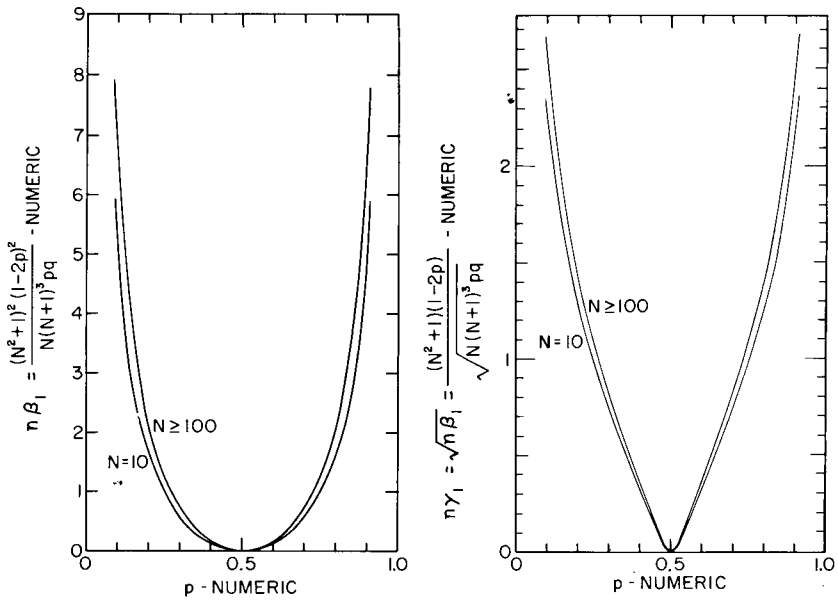

Fig. 3—The effect of two orders of sampling on the third moment numbers β_1 and γ_1. The generalized scales must be multiplied by n^{-1} for β_1 and by $n^{-1/2}$ for γ_1. The symbol n refers to the first sample size; N refers to the second sample size.

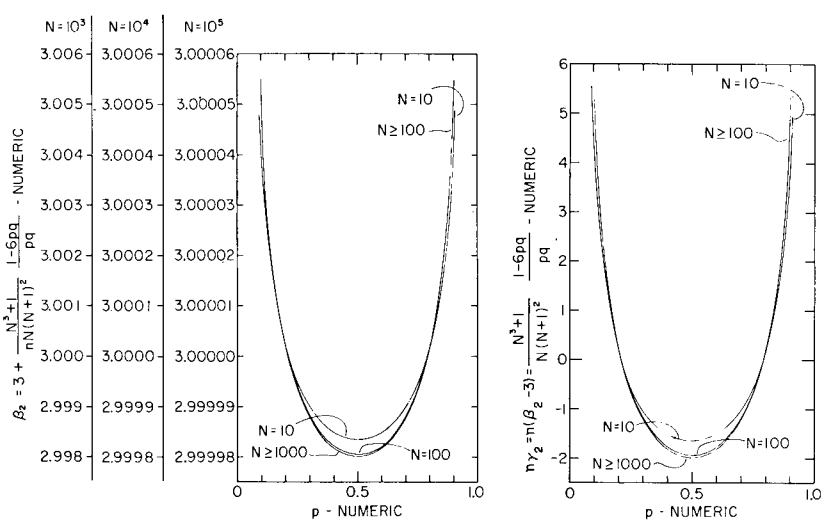

Fig. 4—The effect of two orders of sampling on the fourth moment numbers β_2 and γ_2. The scale for β_2 is not generalized; that for $n\gamma_2$ is. The scale for $n\gamma_2$ must be multiplied by n^{-1} to obtain γ_2. The symbol n refers to the first sample size; N refers to the second sample size.

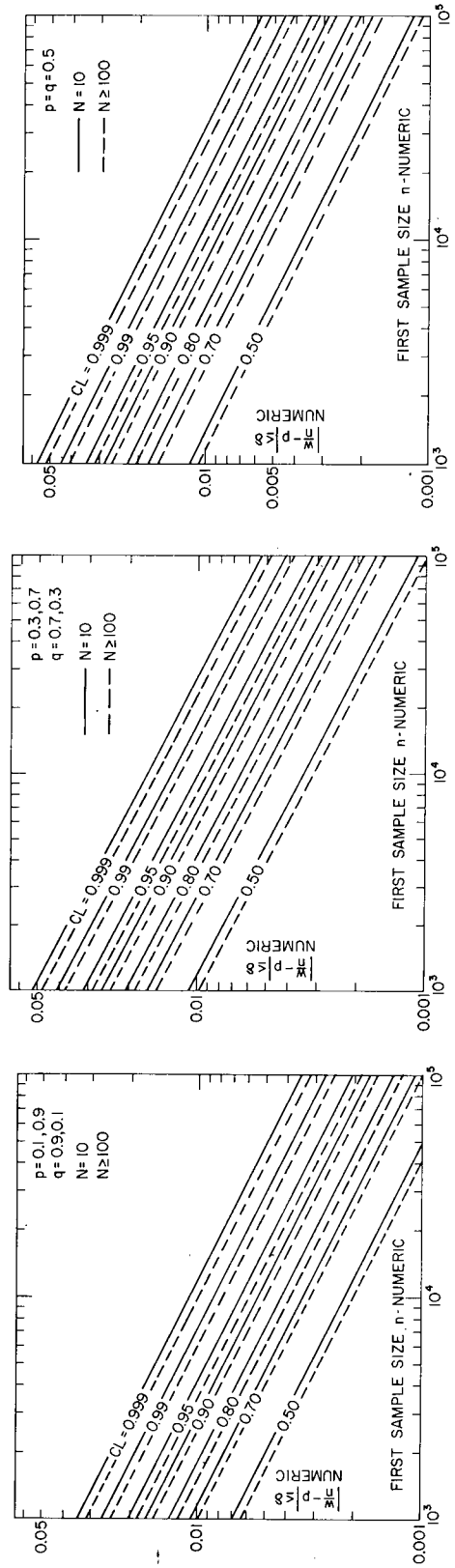

Fig. 5 — The expected measuring error $|(w/n) - p|$ as a function of the first sample size n. An upper bound of error δ is selected by choosing a confidence limit CL, where $CL \equiv Pr(|(w/n) - p| \leq \delta)$. The symbol N refers to the second sample size.

UNCLASSIFIED

of μ_2 and μ_3; μ_4; β_1 and γ_1; and β_2 and γ_2, respectively, as functions of p. The effect of both orders of sampling is clearly shown. It is interesting to note that μ_4 has two components, one of which is essentially independent of the second sample size; however, this component has little or no effect when combined with the other component.

For the case where p is measured, an upper bound of error can be estimated from Fig. 5 when the first (n) and second (N) sample sizes are known. Alternatively, for a given *a priori* error which can be permitted, the necessary sample sizes n and N can be selected.

While the results are indicative of ideal expectations, the derived theoretical results must be used with care. This is so because early assumptions were made [Eqs. (25), (29), and (71)] which involve large sample sizes.

ACKNOWLEDGMENT

Various parts of the paper reflect the fruitfulness of discussions with Mr. C. V. Parker and Dr. P. A. Crafton of the Security Systems and Avigation Branch, Electronics Division, and with Mr. A. Ziffer of the Statistical Analysis Branch of the Applied Mathematics Staff of the Naval Research Laboratory.

REFERENCES

1. Kendall, M.G., "The Advanced Theory of Statistics, Vol. 1," 3rd ed., London Griffin, 1947
2. Hoel, P.G., "Introduction to Mathematical Statistics," New York:Wiley, 1947
3. Feller, W., "An Introduction to Probability Theory and Its Application, Vol. 1," 2nd ed., New York:Wiley, 1957