The Frequency Distribution of the Unknown Mean of a Sampled Universe

By E. C. MOLINA and R. I. WILKINSON

In drawing conclusions as to the reliability of the mean of a sample it is important that all relevant information be taken into consideration. The mathematical analysis in this paper is based on the Laplacian Bayes Theorem which implicitly comprehends the results of a sample together with the a priori knowledge available concerning the parameters of the universe.

The discussion is limited to a universe assumed to be normal but whose mean and precision constant are unknown. Several simplifying, yet quite reasonable, assumptions regarding the forms and independence of the a priori frequency distribution of the true mean and standard deviation are incorporated in the analysis so that numerical answers may more easily be deduced.

Conclusions, properly drawn, are usually quite definitely dependent upon the a priori assumptions made, and especially so in the case of small samples. A considerable space is, therefore, devoted to the solution of a problem in which the sample is only five, taking up a wide variety of these a priori assumptions. They give, in consequence, a wide range of numerical results, appearing in the form of probable errors in the mean of the sample. Each set of assumptions is briefly discussed indicating how the sampling technician may be able to make a selection consistent with his a priori knowledge of a particular problem.

EVERY observation or series of observations upon the items composing a "universe" or "population" may be regarded as constituting a sample. We may divide sampling into two broad natural classes, (1) Sampling of Attributes, and (2) Sampling of Variables. The theory of the first class concerns itself with some particular characteristic, such as the color red, which each item of the universe definitely does or does not possess, and endeavors to assign, ultimately, a numerical value to the probability that the number or proportion of the items in the universe having this characteristic lies within any given range. The second division comprehends that wide variety of problems in which each item of the universe displays to a greater or less degree the same particular quality, such as length, weight, or resistance. After having drawn a random sample of items, probability theory is called upon to assert with what likelihood certain important descriptive constants or "parameters" of the universe lie within any given ranges.

In either class the problem is legitimately attacked by means of a posteriori probability theory. This theory makes use of the two important distinct kinds of knowledge which, in varying amounts, are always at hand, namely, (1) a priori or preexisting information regarding the universe and the possible values which the unknown values of $z_{\text{min.}}$, k, and (c, a) in equation (1′) and solving for M. If, for this value of M thus found, the selected values of c and a coincide with those read respectively from Figs. 2 and 3, a point was established for the given value of $z_{\text{min.}}$ on the M, k plane. If not, sufficient trials were made until the condition given by Figs. 2 and 3 were met. The curves for $z_{\text{min.}}$ were thus determined. To obtain $I_{\text{min.}}$ it was only necessary to use the relation $I_{\text{min.}} = \dfrac{z_{\text{min.}}}{p_t}$.

parameters may assume, and (2) the actual observed value of the studied characteristic in each item of the sample. The a priori information may be meager, in some instances hardly more than the limits between which the parameters must lie, and again, from past experience a great deal may be known about the universe, such as its general form of frequency distribution, the most likely value for each of its parameters to take, and a general feeling that they will not, except in rare cases, lie outside of certain well defined ranges closely bordering their believed most likely values. When the a priori knowledge is meager, more weight must be attached to the results of the sample, but when considerable a priori information is at hand relatively less reliance should be placed in the sample; and in some rare cases it is conceivable that so much is known before the drawings are made that a particular sample, especially if small, would justifiably be disregarded entirely.

The Sampling of Attributes on the a posteriori basis for both infinite and finite universes has already been set forth in these pages at considerable length.[1] The theory of Sampling of Variables when the samples are large becomes usually a matter of assuming that some of the parameters of the sample are sufficiently close to those of the universe that no sensible error will be made in assuming them to be equal. In this case the a priori knowledge of the universe, unless far more exact than is normally found in practice, would exercise but a slight effect in the conclusions which might be drawn, and is therefore quite often properly neglected.

When, for one reason or another, some conclusions are demanded after having taken a small sized sample, it cannot safely be assumed that the sample itself adequately describes the universe, and what a priori knowledge we have must, of necessity, play an important rôle in the determination of any legitimate statements as to the constitution of the universe.

The purpose of this paper is to study in strict accordance with the theory of probability the conclusions which may be drawn concerning the true parameters of the unknown universe after a "sample of variables" of any size has been examined.

The paper is divided into the following five sections:

I. The general equation is given for the a posteriori probability

[1] "Deviation of Random Samples from Average Conditions and Significance to Traffic Men," by E. C. Molina and R. P. Crowell, January 1924. "Some General Results of Elementary Sampling Theory for Engineering Use," by P. P. Coggins, January 1928. This second paper is based on another by Mr. E. C. Molina presented before the Statistical Section of the International Mathematical Congress, held at Toronto in August 1924.

that the true mean of a sampled normal universe lies within a given range.

II. Certain mild restrictions are placed on the general equation of (I) to facilitate its use in practice.

III. The selection of a priori frequency functions in practice is discussed.

IV. A typical example is selected and solved for various a priori existence probability distributions with a discussion of the ranges of errors.

V. Conclusions.

I. The General A Posteriori Equation

It is common, unless information is known to the contrary, to assume that the universe from which the sample is to be made is composed of an infinite number of items all having a particular characteristic whose numerical value from item to item follows the normal frequency law. In the remainder of this paper we shall limit ourselves to a discussion involving only this type of universe. The problem may now be precisely stated:

A set of n observations has been made on a variable quantity drawn from a universe wherein the normal law of errors

$$\left(\frac{h}{\pi}\right)^{1/2} e^{-h(x-m)^2}, \qquad h = 1/2\sigma^2$$

is satisfied but the values of the mean and the precision constant, or standard deviation, are unknown; *before* the observations were made the probability in favor of the simultaneous existence of the inequalities

$$m < \text{mean} < m + dm \tag{1}$$

$$h < \text{precision constant} < h + dh \tag{2}$$

was some function of m and h, say $W(m, h)dmdh$; what is the probability that *after* the observations were made the unknown mean satisfies the inequality (1)?

Let $x_1, x_2 \cdots x_n$ be the values for x given by the n observations. Set

$$n\bar{x} = \sum_1^n x_i, \qquad ns^2 = \sum_1^n (x_i - \bar{x})^2.$$

Now if m and h were known the probability that a set of n observations, *not yet* made, would give values $x_1, x_2, \cdots x_n$ would be

$$\left(\frac{h}{\pi}\right)^{(1/2)n} e^{-h\Sigma(x_i-m)^2} dx_1 dx_2 \cdots dx_n. \tag{3}$$

Therefore, by the Laplacian generalization of the Bayes formula, the a posteriori probability that

$$m < \text{mean} < m + dm$$

is (cancelling factors which do not involve m or h)

$$P(m)dm = \frac{dm \int_0^\infty W(m,\, h) h^{(1/2)n} e^{-h\Sigma(x_i - m)^2} dh}{\int_{-\infty}^{\infty} dm \int_0^\infty W(m,\, h) h^{(1/2)n} e^{-h\Sigma(x_i - m)^2} dh}$$ [2]

$$= A dm \int_0^\infty W(m,\, h) h^{(1/2)n} e^{-h\Sigma(x_i - m)^2} dh, \qquad (4)$$

where A is a constant such that

$$\int_{-\infty}^{\infty} P(m) dm = 1.$$

II. Introduction of Restrictions on General Equation

We are now confronted by a difficulty inherent to a posteriori probability problems. What do we know as to the form of the a priori existence probability function $W(m, h)$? If in a specific practical problem the form of $W(m, h)$ is unknown, no conclusions can be drawn from the set of observations *unless some assumptions are made* and then the weight assignable to the conclusions drawn is a delicate question depending on the reasonableness of the assumptions.[3]

The analysis and results given below are based on assumptions which the writers believe will be found justifiable in many problems of practical interest.

A first assumption which suggests itself is that m and h are independent a priori so that we may write

$$W(m,\, h) = W_1(m) W_2(h). \qquad (5)$$

On this assumption

$$P(m)dm = A W_1(m) dm \int_0^\infty W_2(h) h^{(1/2)n} e^{-h\Sigma(x_i - m)^2} dh. \qquad (6)$$

As a second step toward tentative solutions assume that

$$W_2(h) = K h^{(1/2)c} e^{-ah}, \qquad (7)$$

[2] See Poincare: "Calcul des Probabilites"; 2d edition; articles 178 and 179.

[3] In this connection see italicized paragraph, page 266, "Probability and Its Engineering Uses," T. C. Fry, 1928.

where K, c and a are constants. This, by means of the change of variable

$$y = h[a + \sum (x_i - m)^2]$$

and throwing the definite integral

$$\int_0^\infty y^{(1/2)(n+c)} e^{-y} dy$$

in with the constant A, reduces (6) to

$$P(m)dm = A' W_1(m)[a + \sum (x_i - m)^2]^{-(1/2)(n+2+c)} dm. \qquad (8)$$

We are still confronted with the a priori existence probability function $W_1(m)$.

A plausible form, suggested by the well known "Student" [4] distribution of the ratio $(\bar{x} - m)/s$ for a set of observations *to be* made from a normal universe of *known* mean and standard deviation, is

$$W_1(m) = A_1[1 + B(M - m)^2]^{-(1/2)N}, \qquad (9)$$

where M is the value of m which is a priori most probable, N and B are positive constants while the equation

$$\int_{-\infty}^{\infty} W_1(m) dm = 1$$

gives

$$A_1 = \frac{B^{1/2}\Gamma(\frac{1}{2}N)}{\pi^{1/2}\Gamma[\frac{1}{2}(N - 1)]}.$$

With this assumed form and noting that

$$\sum (x_i - m)^2 = ns^2 + n(\bar{x} - m)^2$$

equation (8) gives

$$P(m)dm = A''[1 + B(M - m)^2]^{-(1/2)N} \times \left[1 + \left(\frac{ns^2}{a + ns^2}\right)\left(\frac{\bar{x} - m}{s}\right)^2\right]^{-(1/2)(n+2+c)} dm, \qquad (10)$$

the integral of $P(m)dm$ between plus and minus infinity determining A''.

Recapitulating: formula (10) gives us the a posteriori frequency distribution for m in terms of the observed data and the arbitrary constants a, c, N, B, M which have entered into the problem in

[4] The writers are aware of the fact that the "Student" frequency function has been put forward in more than one place as the solution for an a posteriori problem. But it should be noted that the various *deductions* of this function which have been given by "Student" and others are entirely a priori.

consequence of the three assumptions made regarding the form of the a priori existence probability function $W(m, h)$; the three assumptions being embodied in equations (5), (7) and (9).

III. Practical Selection of A Priori Frequency Distributions

In equation (10) we have first to assign a numerical value to each of the five constants a, c, N, B, M, before the probability $P(m)$ can be evaluated for any desired range of m. Obviously, in actual practise, the selection of their values is extremely important and too much care cannot be exercised in an attempt to satisfy the engineering judgment that all of the a priori information at hand has been nicely comprehended.

In an endeavor to reduce the number of constants to which we must assign values we shall consider first the a priori function

$$W_2(h) = Kh^{(1/2)c}e^{-ah}.$$

Setting

$$h = c/2a \tag{11}$$

makes $W_2(h)$ a maximum. On the other hand, the value of h which would make the observed set of values of x most probable is given by the equation

$$\frac{1}{h} = \frac{2\sum(x_i - m)^2}{n},$$

or, if m be set equal to $\bar{x}$, we obtain the simpler equation

$$h = 1/2s^2. \tag{12}$$

Upon eliminating h from (11) and (12),[5]

$$a = cs^2. \tag{13}$$

In Fig. 3 are shown four frequency curves of $W_2(h)$. Curve *I* is plotted for $c = 3$ according to equation (13), and to illustrate the wide possibility of forms, curves *II* and *III* have been constructed, keeping $c = 3$, after changing equation (13) to

$$a = \frac{cs^2}{1 - .1s^2} \quad \text{and} \quad a = \frac{cs^2}{1 + .1s^2},$$

respectively. Curve IV again satisfies equation (13) but has c increased from 3 to 6.

[5] It should be carefully noted that there is no necessary relation between the *a priori most probable value of h* and the *value of h which would make the observed event most probable.* The elimination of h between (11) and (12) is justified solely by the practical consideration that a tentative relation between a and c will reduce by one the number of arbitrary constants to which numerical values must be assigned.

For every assumption of a and c in the a priori distribution of h there is, of course, a corresponding a priori distribution, $\phi(\sigma)$, of the standard deviation. Here

$$\phi(\sigma)d\sigma = K'\sigma^{-(c+3)}e^{-(1/2)a/\sigma^2}d\sigma$$

and

$$K' = \frac{a^{(1/2)c+1}}{2^{(1/2)c}\Gamma(\frac{1}{2}c+1)}.$$

The distributions of σ corresponding to each of the four frequency curves of h in Fig. 3 are shown in Fig. 4 with similar designations.

In many cases, too, it is obvious that very little is known concerning the shape and the parameters of the mean's a priori distribution beyond that it is generally unimodal and quite likely to be fairly symmetrical about its most probable value; a mathematical expression of this has been set up in equation (9). In this circumstance we may not introduce serious restrictions if we make two further assumptions which greatly simplify (9).

The first is that we set $M = \bar{x}$ which says that, a priori, the most probable value of the unknown mean was the same as that which was later calculated as the mean in the sample.[6] It is admitted that the chance of exactly fixing on $M = \bar{x}$ from a priori information is very small, yet if our knowledge is so slight that we must introduce some guesswork here, the selection of the value of $\bar{x}$ at least has the advantage of being a *possible* one which M might assume and, except in rare cases, it will not be greatly distant from the true m in that particular lot. The logical difficulty here also may be minimized by selecting a form of $W_1(m)$ of such flatness that over a considerable range of values in the neighborhood of $\bar{x}$ the existence probability does not take on widely differing magnitudes.[7]

The second assumption can more readily be allowed, and consists in empirically defining

$$B = \frac{n}{a + ns^2}.$$

This removes a degree of freedom from the $W_1(m)$ function but, as far as its form is concerned, except in special cases, the one variable, N, may serve quite well in characterizing the pre-existing information. As is clearly shown in Fig. 1, the increase of N indicates a greater

[6] While it does not matter in this particular problem, the authors wish to carefully distinguish, at least in thought, between an "observed" parameter and a parameter calculated from individual observations.

[7] The setting of $M = \bar{x}$, it should be noted, has no effect if all values of the mean are made a priori equally likely by setting $N = 0$ (that is, $W_1(m) = A_1$).

certainty in the investigator's mind that the true value of m lies closer and closer to the assumed most probable figure, M.

With these two assumptions incorporated in equation (10) we may now write

$$P(m)dm = f(t)dt = A'''(1 + t^2)^{-(1/2)T}dt, \tag{10'}$$

in which

$$t^2 = \left(\frac{\bar{x} - m}{s}\right)^2 \left(\frac{ns^2}{a + ns^2}\right) = B(\bar{x} - m)^2, \tag{14}$$

$$T = n + 2 + c + N,$$

$$A''' = \frac{1}{\int_{-\infty}^{\infty} (1 + t^2)^{-(1/2)T}dt} = \frac{\Gamma(\frac{1}{2}T)}{\pi^{1/2}\Gamma[\frac{1}{2}(T - 1)]}.$$

The formula (10′) is a "Student"[8] frequency form with the arguments n and $\frac{\bar{x} - m}{s}$ replaced by $n + 2 + c + N$ and $\frac{\bar{x} - m}{s(1 + a/ns^2)^{1/2}}$ respectively.

Fig. 2 shows curves plotted for ranges of t such that

$$A''' \int_{-t}^{+t} (1 + t^2)^{-(1/2)T}dt = .50,\ .80,\ .90, \text{ and } .9973,$$[9]

and the errors in the mean corresponding to any of these probabilities, after determining t, may be found by evaluating $\bar{x} - m$ in equation (14).

IV. Solution of a Typical Example

Five samples of retardation coils rated at 47 ohms are taken from a large lot, and careful measurements show them to have resistances of 46.30, 44.40, 47.72, 50.50, and 45.58 ohms respectively. We are asked to determine the probable and 99.73 per cent errors of the average of these resistances, assuming that the samples have been drawn from a normal universe.

The average of these five values is $\bar{x} = 46.90$ ohms and their standard deviation about this average is found to be $s = 2.097$.

From the preceding discussion it is evident that as many answers to this problem may be obtained as there are assumptions made regarding, in general, the a priori distributions of the mean and

[8] Student: "The Probable Error of a Mean," Biometrika, Vol. VI, No. 1, March 1908.

[9] Student: "New Tables for Testing the Significance of Observations," Metron, Vol. V, No. 3, I–XII–1925. Tables I and II, pages 114–118, for values of $n' = 2$ to 21.

precision constant, and in our particular analysis the five constants found in equation (10). In Table I and Fig. 1 we tabulate and portray graphically twenty-one complete solutions of the example based upon as many sets of values given to these constants. A wide

TABLE I

$\bar{x} = 46.90$ $n = 5$ $s = 2.097$

No.	Parameters in Existence Probability Distributions							$T = n + 2 + c + N$	A Posteriori Probability Values of Errors in Observed Mean	
	Mean				Precision Constant				Probable or 50% Error	99.73% Error
	M	N	A_1	B	c	a	K			
1	46.9	0	—	.2274	−3	0	—	4	.923	11.16
2	46.9	0	—	.2274	−2	0	—	5	.776	6.94
3	46.9	0	—	.2274	0	0	—	7	.623	4.23
4	46.9	0	—	.1421	3	13.19*	475.5	10	.626	3.61
5	46.9	0	—	.1098	3	23.55†	2,024	10	.712	4.10
6	46.9	0	—	.1605	3	9.163‡	191.2	10	.589	3.39
7	46.9	0	—	.1034	6	26.38*	80,770	13	.625	3.39
8	46.9	1	—	.2274	0	0	—	8	.566	3.59
9	46.9	1	—	.1421	3	13.19*	475.5	11	.587	3.33
10	46.9	2	.1518	.2274	0	0	—	9	.524	3.17
11	46.9	2	.1200	.1421	3	13.19*	475.5	12	.558	3.08
12	46.9	2	.1055	.1098	3	23.55†	2,024	12	.635	3.50
13	46.9	2	.1275	.1605	3	9.163‡	191.2	12	.525	2.90
14	46.9	2	.1023	.1034	6	26.38*	80,770	15	.575	3.01
15	46.9	4	.3036	.2274	0	0	—	11	.464	2.63
16	46.9	4	.2400	.1421	3	13.19*	475.5	14	.511	2.71
17	46.9	4	.2110	.1098	3	23.55†	2,024	14	.581	3.09
18	46.9	4	.2551	.1605	3	9.163‡	191.2	14	.480	2.55
19	46.9	4	.2047	.1034	6	26.38*	80,770	17	.537	2.76
20	41.0	2	.1200	.1421	3	13.19*	475.5	—	.665§	3.87§
21	49.0	2	.1200	.1421	3	13.19*	475.5	—	.635§	3.65§

* $a = cs^2$.

† $a = \dfrac{cs^2}{1 - .1s^2}$.

‡ $a = \dfrac{cs^2}{1 + .1s^2}$.

§ Errors determined by planimeter method.

variety of a priori conditions is assumed giving, in consequence, widely varying probable and 99.73 per cent errors.

The a priori frequency functions, $\phi(\sigma)$, for the standard deviation in the first seven cases are shown in broken lines superposed upon the distributions of precision constants in Fig. 1. The scales of h and σ are not to be confused, the attempt being only to represent the *form* of the $\phi(\sigma)$ frequency curves.

(*a*) If we wish to be very conservative we might select values for the unknown constants which would make all values of m and σ

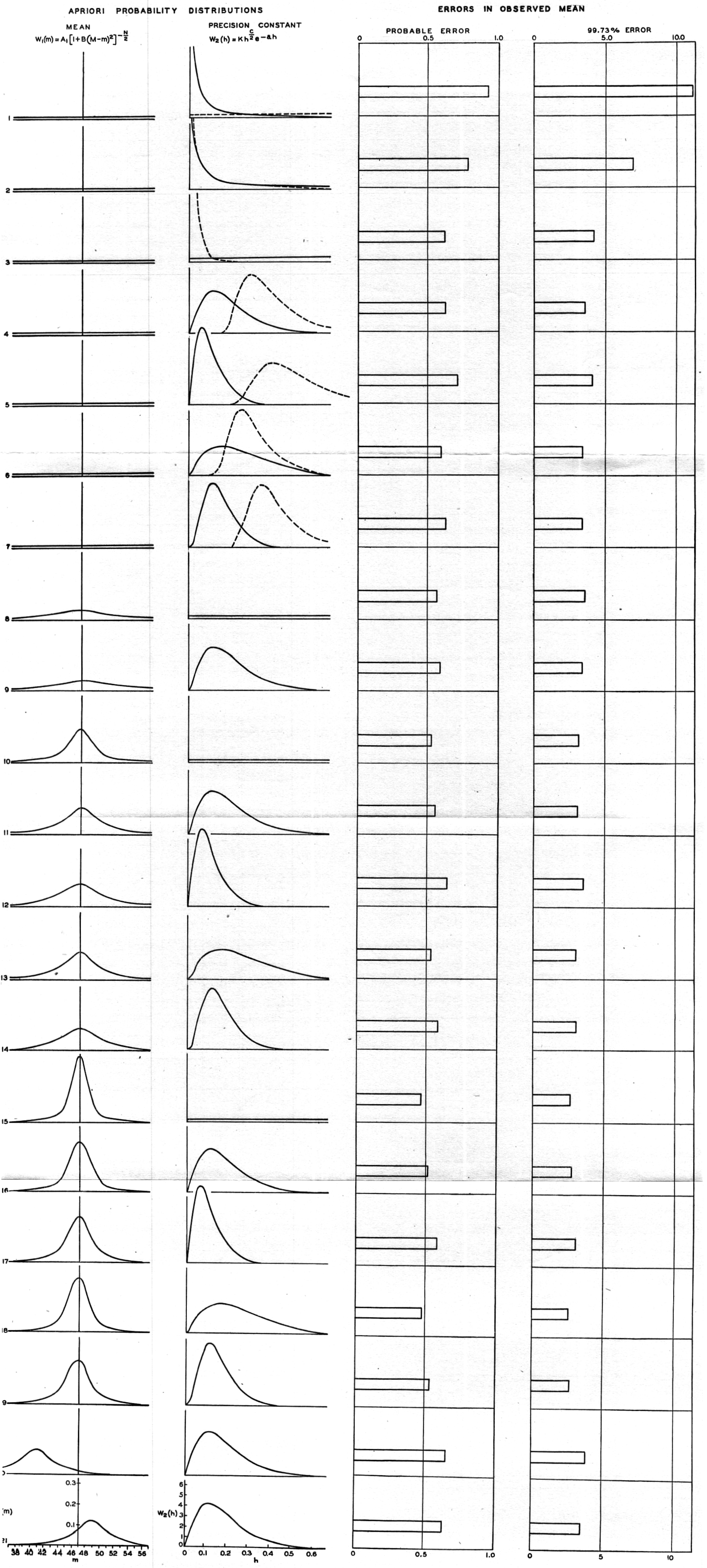
APRIORI PROBABILITY DISTRIBUTIONS
MEAN
$W_1(m) = A_1\left[1+B(M-m)^2\right]^{-\frac{N}{2}}$
PRECISION CONSTANT
$W_2(h) = Kh^{\frac{C}{2}}e^{-ah}$
ERRORS IN OBSERVED MEAN
PROBABLE ERROR
99.73% ERROR
m
h
$W_2(h)$

FIG. 1.

equally likely, that is, $N = 0$, $c = -3$, $a = 0$.[10] Here the precision constant's a priori distribution is decidedly exponential and we might predict the large probable and 99.73 per cent errors in the observed average which actually result.

Case 1 in Table I and Fig. 1 presents the problem in its entirety with the resultant errors tabulated as well as shown graphically.

(*b*) The engineer's knowledge, however, in all probability, is not so limited as in (*a*) above, at least regarding the precision constant

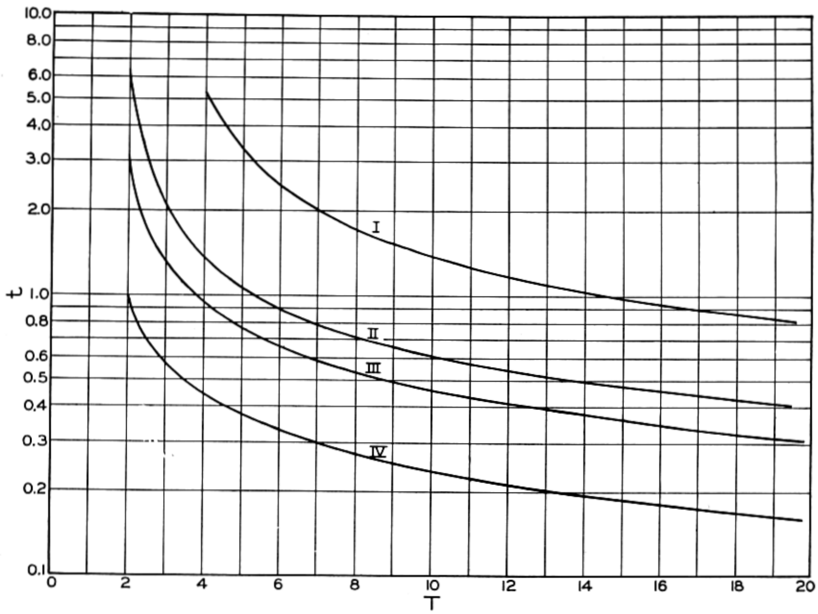

Fig. 2—Errors of averages of samples of size n.
I—99.73 per cent Error.
II—90.00 per cent Error.
III—80.00 per cent Error.
IV—50.00 per cent Error.

Note: Abscissa: $T = n + 2 + c + N$.
Ordinate: t = The Product of the Error of the Average and the Square Root of B.

(or the standard deviation). He knows that extremely small values of the precision constant are less likely than larger ones, and to some extent we picture the transition from (*a*) to this impression in the Cases Nos. 2 and 3 which as before may be found completely portrayed

[10] The formula for $P(m)$ resulting from a substitution of these constants in equation (10′) reproduces the result obtained by Drs. J. Neyman and E. S. Pearson for all values of the a priori function $W'(m, \sigma)$ equally likely: Biometrika, Vol. XX*A*, Parts I and II, July 1928; "On the Use and Interpretation of Certain Test Criteria for Purposes of Statistical Inference," page 196, equation XXXV.

in Table I and Fig. 1. Case No. 2, it is interesting to note, is the familiar "Student" formula; Case No. 3's outstanding characteristic is that all values of h are a priori equally likely. The errors in the mean have here been greatly reduced by merely changing the existence probability distribution of the precision constant.

(*c*) Again, the experienced analyst is quite likely to assume willingly that the distribution of the precision constant (and likewise the standard deviation) is of a unimodal form having its maximum value not greatly distant from the figure determined in the sample. Cases Nos. 4 to 7 inclusive typify this kind of assumption while, at the same time, all values of the true mean are held a priori, equally likely.

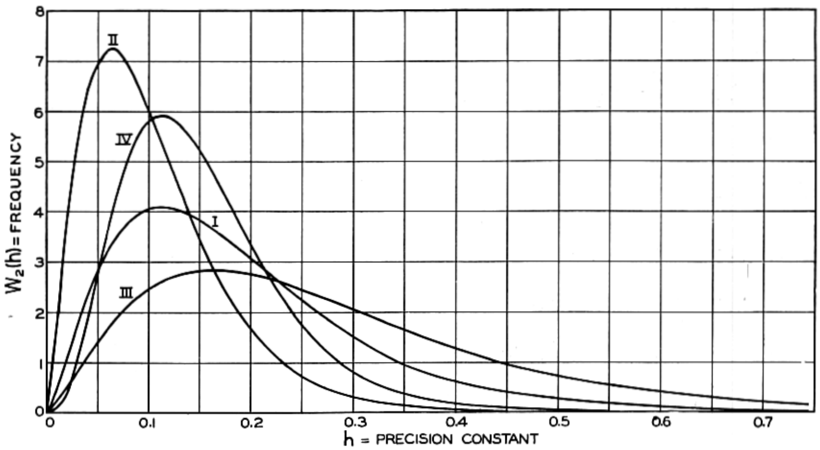

Fig. 3—Typical a priori frequency distributions of the precision constant.

$$W_2(h) = Kh^{(1/2)c}e^{-ah}.$$

$$\text{I}—c = 3,\ a = cs^2 = 13.1922.$$

$$\text{II}—c = 3,\ a = \frac{cs^2}{1 - .1s^2} = 23.5467\cdot$$

$$\text{III}—c = 3,\ a = \frac{cs^2}{1 + .1s^2} = 9.1629\cdot$$

$$\text{IV}—c = 6,\ a = cs^2 = 26.3845.$$

The constants for Cases Nos. 4 and 7 have been so selected as to bring the modal value of h at that found from the sample, that is, that value of h has been made most likely a priori which will make the probability of occurrence of the particular value $(1/2s^2)$ calculated from the observations, a maximum. Case No. 7 is a considerably more peaked distribution than Case No. 4 indicating more faith in the modal figure selected as being close to the true value. Cases Nos. 5 and 6 illustrate how the mode of the $W_2(h)$ function which always lies at $h = c/2a$ may be shifted either down or up and the extent of modification in the resulting errors which may be expected.

The four frequency distributions of h just considered are the same as those shown in more detail in Fig. 3; the corresponding distributions of the standard deviation are found detailed on Fig. 4.

(*d*) If the interpreter of the data is closely familiar with the sampled product and has been observing similar lots for some time he may have a reasonably good idea as to the value of the general average

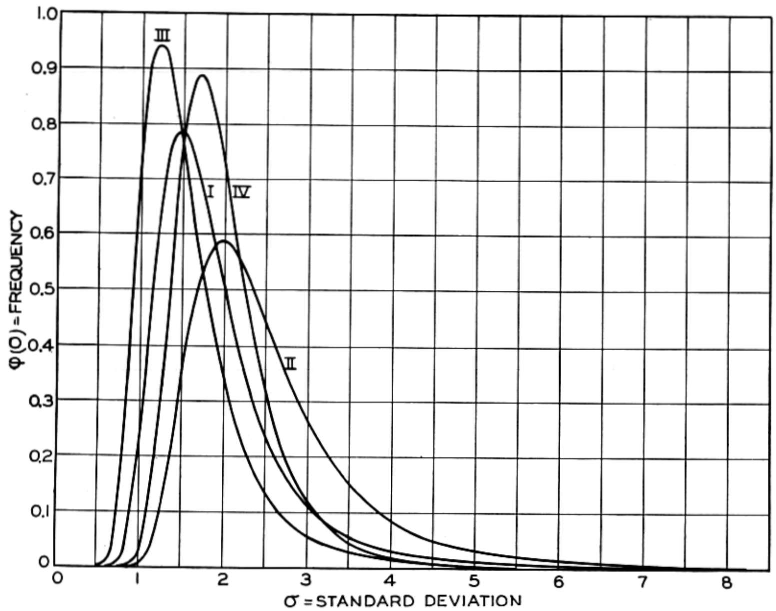

Fig. 4—Typical a priori frequency distributions of the standard deviation.

$$\phi(\sigma) = K'\sigma^{-(c+3)}e^{-(1/2)a/\sigma^2}.$$
$$\text{I}—c = 3,\ a = cs^2 = 13.1922.$$
$$\text{II}—c = 3,\ a = \frac{cs^2}{1 - .1s^2} = 23.5467.$$
$$\text{III}—c = 3,\ a = \frac{cs^2}{1 + .1s^2} = 9.1629.$$
$$\text{IV}—c = 6,\ a = cs^2 = 26.3845.$$

of items produced under these same essential conditions. In Cases Nos. 8 to 19, inclusive, use is made of this knowledge on the assumption that $\bar{x}$, the mean of the sample, turns out to be so nearly equal to M, the most likely a priori value of the true mean m, that we may safely call them identical. Three values of N, regulating the spread of the $W_1(m)$ distribution to conform to the observer's best judgment of the true circumstances have been associated with the same sequence of a priori assumptions regarding the precision constant as were presented in Cases Nos. 3 to 7.

The errors found in Cases Nos. 8 to 19 on the various combinations of a priori frequency curves lie in a fairly narrow band distinctly below those determined from the more conservative assumptions underlying Cases Nos. 1, 2 and 3. This well illustrates the importance of carefully surveying and as far as possible completely utilizing the knowledge available before the sample has been made.

(*e*) Finally, cases are bound to occur in which the engineer can quite definitely say that some value of M other than $\bar{x}$ is a priori most probable; this situation is encountered in Cases Nos. 20 and 21. These are identical with Case No. 11 except that in Case No. 20, M has been reduced about 6 ohms and in Case No. 21 raised about

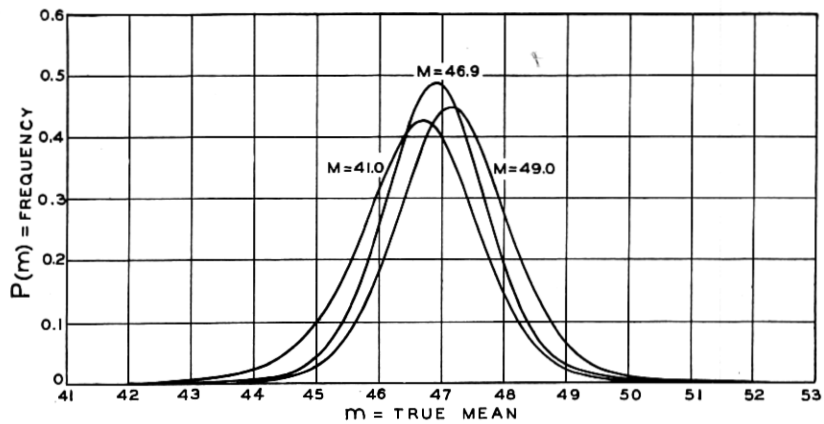

Fig. 5—Typical a posteriori frequency distributions of the unknown mean.

$$P(m) = A''[1 + B(M - m)^2]^{-(1/2)N}\left[1 + \left(\frac{ns^2}{a + ns^2}\right)\left(\frac{\bar{x} - m}{s}\right)^2\right]^{-(1/2)[n+2+c]}.$$

2 ohms. The errors are somewhat increased by these changes in M, as, of course, we should have predicted. Comparisons such as this should help the investigator to decide whether or not his previously selected figure for M is sufficiently close to $\bar{x}$ that they may safely be equated.

In the event that it is decided that M may not be set equal to $\bar{x}$, in any particular problem, as in Cases Nos. 20 and 21, the symmetrical "Student" form of distribution for $P(m)$, (except when $N = 0$) no longer occurs. This is clear from an inspection of Fig. 5 which shows the three cases plotted on the same scale.

It is suggested then, since the integral of $P(m)dm$ here may become difficult to handle, that recourse be had to the use of a planimeter on the distribution plotted from equation (10) on rectangular co-ordinate paper. In this way may be determined within what range, equidistant above and below $\bar{x}$, lies the proportion of the total area corresponding to the desired probability.[11]

V. Conclusions

We have presented a general equation for the probability that the true mean of a sampled normal universe lies within a given range, incorporating the kind of knowledge the investigator may be expected to have before the sample was made as well as the information directly presented by the individual observations themselves. It cannot be overemphasized that the problem by its very nature is indefinite since it would be a rare instance indeed to find a mathematical expression which would completely and exactly summarize the a priori knowledge, impressions and beliefs in the mind of any person confronted with its solution. All that can be found is, at best, an approximate probability based upon certain assumptions we are willing to make in order to arrive at a numerical result. And only by utilizing as far as possible all of the available knowledge will the most nearly correct probability values ascertainable be realized.

[11] On certain test cases of "Student" distributions, the error in planimeter readings averaged about one-half of one per cent, and in no case exceeded one per cent.