METHODS OF OBTAINING PROBABILITY DISTRIBUTIONS[1]

By Burton H. Camp

The emphasis of this paper will be on method. Special results will be cited in order to illustrate the methods rather than to summarize achievement in the field; for that has been done already by Rider (1930, 1935) Irwin (1935) and Shewhart (1933) in recent surveys. The purpose is to describe and to illustrate most of the methods that have been used to determine exact probability distributions, and to show that they are all derivable from one fundamental theorem. In order to prove this unity in a simple manner, it will be desirable to omit from consideration methods which are essentially ingenious forms of counting, such as are used in sampling without replacements from finite universes, and in finding the sampling distribution of a percentile.

The general problem to be discussed may be stated as follows: N individuals $(t_1, \cdots, t_N)$ are drawn, one at a time with replacements, from a universe whose probability distribution is $\phi(t)$. A certain single valued function of the t's is formed. This is called a parameter of the sample, and is frequently also, but not necessarily, a useful estimate of the corresponding parameter of the universe. The problem is to find its probability distribution, $f(x)$. As usual, a probability distribution is a function which is required to be defined, except perhaps at a set of measure zero, throughout the infinite domain of its variables; it is nowhere negative, and its integral over its domain is unity.

Most of the more recent developments of the theory relate to a more general form of this problem. Instead of N individuals, there are N sets of n individuals in each set, and these sets are drawn respectively from $M (M \leqq N)$ universes, each of which is described by a function of n independent variables, thus:

$$\phi^{(i)}(t_1, \cdots, t_n); \; (i = 1, \cdots, M). \tag{1}$$

Instead of a single parameter there are P parameters, and each is a single valued function of the observed values of the nN individuals in the sample, thus:

$$x_i = g_i(t_1^{(1)}, \cdots, t_n^{(1)}; \cdots; t_1^{(N)}, \cdots, t_n^{(N)}); \; (i = 1, \cdots, P) \tag{2}$$

The first method to be described is fundamental and will be designated as

Theorem I. Let it be required that each g as described in (2) be not only single valued but also constant at most in a set of measure zero in the nN-way space of the t's. Then

$$\int_p f(x_1, \cdots, x_P)\, dX = \int_q \phi(t_1^{(1)}, \cdots, t_n^{(N)})\, dT \tag{I}$$

[1] Presented to the American Mathematical Society at a meeting devoted to expository papers on the theory of statistics, April 11, 1936.

where X is the space of x's and T the space of the t's, p is any measurable set of points in X, and q is the set in T for which g is in p. Often p is the P dimensional cube $(x_i + \Delta x, i = 1, \cdots P)$ at the point $(x_1, \cdots, x_p)$ and then q is the set where

$$x_i \leqq g_i \leqq x_i + \Delta x; \qquad (i = 1, \cdots, P) \tag{3}$$

and ϕ is the simultaneous distribution of the sets of t's,

$$\phi^{(1)}(t_1^{(1)}, \cdots, t_n^{(1)}) \cdots \phi^{(N)}(t_1^{(N)}, \cdots, t_n^{(N)}). \tag{4}$$

In this $\phi^{(i)}$ is the universe from which the $t^{(i)}$ set of t's is drawn. Obviously, if $N > M$, some of the $\phi^{(i)}$'s are identical, and then it is assumed that the several sets are drawn independently. Often, all of the N sets of t's are drawn from the same universe. Then $M = 1$ and all these ϕ's are identical, and (4) becomes

$$\phi = [\phi^{(1)}(t_1^{(1)}, \cdots, t_n^{(1)}] \cdots [\phi^{(1)}(t_1^{(N)}, \cdots, t_n^{(N)})].$$

In the special case where there is but one parameter ($P = 1$) and but one individual in the sample ($n = N = 1$), and p is an interval, formula (I) becomes

$$\int_x^{x+\Delta x} f(x)\, dx = \int \phi dt; \tag{Ia}$$

and in the very special case where it is also true that q is an interval it becomes

$$f(x) = \phi(t) \cdot \left| \frac{dt}{dx} \right|, \tag{Ib}$$

provided also that certain derivatives (to be specified later in the proof) exist, where t is now the inverse solution of the equation,

$$x = g(t). \tag{5}$$

The proof of formula (I) is immediate, if one is willing to assume the existence of the probability distribution f; for then the left side is by definition the probability that the x's lie in p, and this is also the meaning of the right side of (I). (Ia) can be proved without assuming initially the existence of $f(x)$, for then the existence of $f(x)$ can be inferred from the existence of the right side of (Ia), because $f(x)$ may be set equal (except perhaps at a set of measure zero) to the upper right hand derivative, with respect to Δx (Δx is a variable, and x is fixed), of $\int_q \phi\, dt$, provided that one adds the condition that this derivative is nowhere infinite. The point at issue here is merely the existence of a primative for a monotone increasing function of Δx. (Ib) may be derived from (Ia) by taking the derivative of both sides with respect to Δx, if the derivatives are continuous.

Theorem I, in these various forms is used a great deal, especially in the last form (Ib). This affords one freedom to choose the most desirable function for purposes of tabulation. R. A. Fischer's z distribution, a logarithm, is an important illustration. Many authors have been interested in so choosing the function that its distribution shall be normal. They include several of the older writers, and more recently H. L. Rietz (1921, 1927), and G. A. Baker (1932, 1934). However, the theorem is of special importance in the theory, for all the other principal methods of obtaining probability distributions are essentially corollaries of it. These corollaries will be called Theorems II, III, and IV.

THEOREM II. Let $\bar{p}$ (the measure of p) and $\bar{q}$ (the measure of q) be infinitesimals of the same order and let both the oscillation of f(*i.e.* maximum f-minimum f) in p and the oscillation of ϕ in q be infinitesimals; then (I) may be written,

$$f\bar{p} = \phi\bar{q}, \tag{II}$$

where f applies to any point of p and ϕ to the corresponding point of q. This equation (II) is an approximate equation in the sense that differences of higher order than those retained are neglected. In particular, with the conditions used in formula (Ia), equation II becomes

$$f\Delta x = \phi\bar{q}.$$

The left side of (II) is an approximation to the probability sought. The right side shows that, in order to evaluate it, one need only find the volume in T space of the differential element q and multiply it by the value of ϕ in q. Formula (II) expresses the so-called geometrical method used by many authors, *e.g.*, by R. A. Fisher (1915, 1925), by Wishart (1928), and by Hotelling (1925, 1927). The chief difficulty in connection with it is in finding the volume of nN-dimensional q. In order to display the advantages and disadvantages of this method we shall pause at this point and look at a concrete example.[2]

Let two individual (t_1, t_2) be drawn independently from a normal universe and consider the simultaneous distribution $f(x, y)$ of the sum, $x = t_1 + t_2$, and product, $y = t_1t_2$, the mean of the universe being chosen as the origin. Here $N = 2$, $n = 1$, $M = 1$, and so,

$$\phi = \frac{1}{2\pi\sigma^2} e^{-\frac{1}{2\sigma^2}(t_1^2 + t_2^2)} = \frac{1}{2\pi\sigma^2} e^{-\frac{1}{2\sigma^2}(x^2 - 2xy)} \tag{6}$$

The point set q is the area lying between the two adjacent hyperbolae,

$$t_1t_2 = y, \qquad t_1t_2 = y + \Delta y,$$

and also between the two adjacent lines,

$$t_1 + t_2 = x, \qquad t_1 + t_2 = x + \Delta x,$$

where Δx and Δy are infinitesimals and are equal. This area may be computed by simple integration and is:

[2] See also C. C. Craig (1936). Craig uses another method to be explained later (formula IIIa).

$$q = \frac{2\Delta x\,\Delta y}{\sqrt{x^2 - 4y}} \qquad \text{if} \quad x^2 > 4y,$$
$$= 0 \qquad \text{if} \quad x^2 < 4y.$$

Hence II gives us immediately the desired result:

$$f(x, y)\,\Delta x \Delta y = \frac{1}{\pi\sigma^2}\, e^{-\frac{x^2-2y}{2\sigma^2}}\, \frac{1}{\sqrt{x^2 - 4y}} \cdot \Delta x \Delta y, \qquad \text{if} \quad x^2 > 4y,$$
$$= 0 \quad \text{if} \quad x^2 < 4y.$$

If $x^2 = 4y$, $\bar{q}$ is an infinitesimal of lower order than $\bar{p} = (\Delta x)^2$, and so Theorem II does not apply. In this case we must go back to Theorem I, and from that we can learn that the probability,

$$\int_p f\,dx\,dy,$$

is an infinitesimal of the first order if $p = \Delta x\,\Delta y = (\Delta x)^2$ is of the second order. Hence it cannot be approximately represented by a finite number times $\bar{p}$. The oscillation of f in p is infinite. The form of the surface $f(x, y)$ is interesting. The ordinates rise to infinity on the contour of the parabola $x^2 = 4y$, and vanish within it. The surface is symmetrical with respect to the plane $x = 0$, but not with respect to the plane $y = 0$. However, it is clear that the total probability of any given product, y (*i.e.* the probability of this y for all possible values of x), is the same as the total probability of $-y$; hence

$$\int_{-\infty}^{\infty} f(x, y)\,dx = \int_{-\infty}^{\infty} f(x, -y)\,dx,$$

and the corresponding formulae,

$$\frac{2}{\pi\sigma^2}\, e^{\frac{y}{\sigma^2}} \int_{\sqrt{4y}}^{\infty} e^{-\frac{x^2}{2\sigma^2}} \frac{1}{\sqrt{x^2 - 4y}}\,dx \qquad (y > 0),$$

and

$$\frac{2}{\pi\sigma^2}\, e^{\frac{y}{\sigma^2}} \int_{0}^{\infty} e^{-\frac{x^2}{2\sigma^2}} \frac{1}{\sqrt{x^2 - 4y}}\,dx \qquad (y < 0),$$

must be equal; both may be reduced to the single form

$$F(y) = \frac{1}{\pi\sigma^2} \int_0^{\infty} e^{-\frac{1}{2\sigma^2}\left(t^2 + \frac{y^2}{t^2}\right)} \frac{dt}{t}, \qquad \text{if} \quad y \neq 0.$$

This is the probability distribution of y.

With this example before us, let us now reconsider the theory:

(*i*) The requirement (in II) that the oscillation of ϕ be infinitesimal in q

will be satisfied if one can show that ϕ may be expressed as a continuous function of the parameters $(x_1, x_2, \cdots, x_p)$. In our example these parameters were x and y and ϕ was so expressible (6). But if we had tried initially to find by means of (II) the distribution of the product y, independently of what values x might have, we should have been stopped at this point, because ϕ is not expressible in terms of y alone. We should also have been stopped by the requirement that $\bar{q}$ be infinitesimal of order Δy, for q would have been the space between two hyperbolas and its area for any fixed $(\Delta y > 0)$ would have been infinite. But, when thus stopped at that first point, it would have been clearly indicated to us that the distribution of y might have been found *via* the detour of finding the simultaneous distribution of both x and y, because an attempt to express ϕ in terms of y would have led to the given expression in terms of both x and y. For a similar reason R. A. Fisher (1925) was able to find the distribution of the variance by finding first the simultaneous distribution of the variance and the mean. Also, he was thus able to find the distribution of the coefficient of correlation by finding first the simultaneous distribution of all the first and second order moments.

(*ii*) A distinct advantage of this method is that q is independent of the universe ϕ, so that once found it may be used in connection with any universe which satisfies the condition that it can be expressed as a continuous function of the parameters. Thus, the distribution of the sum and product in our example may equally well be found for the universe described by the Type III curve, $Ate^{-at}(t > 0)$. For, then

$$\phi = A^2 t_1 t_2 e^{-a(t_1+t_2)} = A^2 y e^{-ax},$$

and so, using one-half of the same $\bar{q}$ as before, since now $x, y \geqq 0$,

$$\begin{aligned} f(x, y) &= A^2 y e^{-ax} \frac{2}{\sqrt{x^2 - 4y}}. && \text{if } x^2 > 4y, \\ &= 0 && \text{if } x^2 < 4y. \end{aligned}$$

From this, $F(y)$ can be found by integration (*c.f.* Kullbach, 1934)

$$F(y) = A^2 y \int_{\sqrt{4y}}^{\infty} \frac{e^{-ax}}{\sqrt{x^2 - 4y}} dx = \frac{A^2 y}{2} \int_0^{\infty} \frac{e^{-a\left(u + \frac{y}{u}\right)}}{u} du.$$

As another illustration, consider a normal universe of n intercorrelated variables in which all the total intercorrelations are equal to r (*e.g.*, the statures of n brothers) and let the sample be a single group of n (one individual for each variable).

$$\phi = \frac{1}{(2\pi)^{n/2} R} e^{-\frac{1}{2R}\left[k_1 \Sigma t_i^2 + k_2 \sum_{i \neq j} t_i t_j\right]},$$

where $R = (1 - r)^{n-1}[1 - (n - 1)r]$, $k_1 = (1 - r)^{n-2}[1 - (n - 2)r]$, and $k_2 = -r(1 - r)^{n-2}$. Suppose one wishes to find the simultaneous distribution of the variance x and the mean y for such samples.[3] Since for Student's problem Fisher has found the value of q for this x and y to be

$$\bar{q} = cx^{\frac{n-3}{2}} \Delta x \Delta y,$$

their distribution $f(x, y)$ for this universe may be written down immediately. In terms of x and y the bracket in the exponent of ϕ is $y^2(k_1 n - k_2 n + k_2 n^2) + xn(k_1 - k_2)$, and so $f(x, y)$ is the product of $\bar{q}$ and this form of ϕ:

$$f(x, y) = K e^E x^{\frac{n-3}{2}}, \qquad E = -\frac{1}{2R}[(k_1 n - k_2 n + k_2 n^2)y^2 - n(k_1 - k_2)x].$$

(*iii*) Another attribute of this method is that it sometimes lends itself to easy extensions from a simple case where there is only one restriction ($N - 1$ degrees of freedom) to similar cases when there are more restrictions. Thus R. A. Fisher (1924) proceeded from the variance of a sample from a single universe to the variance from a set of universes, as required in the theory of analysis of variance; and thus also (1915) he had proceeded from the distribution of r to that of multiple R; and Hotelling (1927) showed how these distributions could be obtained when the values of each variate were themselves intercorrelated (as in a time series) and not merely correlated with values of the other variates.

THEOREM III. Now let us consider again the fundamental form (I). For convenience let $nN = m$. If the conditions will not permit us to write the right side in the form in (II), it is still possible that we may be able to find that $(m + 1)$-dimensional volume by some other method. In particular, whenever it is possible to iterate the integral once we have the formula:

$$\text{(III)} \qquad \int_p f dX = \int_{T'} dT' \int_{q_m} \phi \, dt_m,$$

where q_m is the section of q by t_m space at the point $(t_1, \cdots, t_{m-1})$ of T' space, T' space being the space of the $(t_1, \cdots, t_{m-1})$ coordinates. With added conditions one may deduce from (III), for the case where there is but a single parameter x, the approximate equation:

$$\text{(IIIa)} \qquad f\,dx = dx \int_{T'} dT' \cdot \phi(t_1, \cdots, t_m) \frac{dt_m}{dx},$$

in which t_m is supposed to have been expressed in terms of the other coordinates by solving the equation $x = g(t_1, \cdots, t_m)$. It is an approximate equation in the same sense as (II) was. Sufficient conditions for this change in the left side of (III) have already been mentioned in discussing (II). The propriety of making the corresponding change in the right hand side may be left for determination when the form of ϕ is given. It will perhaps be sufficient here to point out that our earlier example illustrates both the case where this change

[3] A special case of a more general problem solved first by R. A. Fisher.

is permissible and where it is not. For, let it be required to find the distribution $f(y)$ of the product $y = t_1t_2$ without reference to the sum, $t_1 + t_2$. Formula (III) yields

$$\text{(7)} \qquad \int_y^{y+\Delta y} f(y)\, dy = 2\int_0^\infty dt_1 \int_{y/t_1}^{(y+\Delta y)/t_1} dt_2 \cdot \frac{1}{2\pi\sigma^2} e^{-\frac{1}{2\sigma^2}\left(t_1^2 + t_2^2\right)}$$

This is valid for every value of y including $y = 0$. If $y \neq 0$, we may change the right hand side as in (IIIa) and obtain as the probability that y is in the interval $(y, y + \Delta y)$:

$$\text{(8)} \qquad \int_y^{y+\Delta y} f(y)\, dy = \frac{\Delta y}{\pi\sigma^2} \int_0^\infty \frac{1}{t_1} e^{-\frac{1}{2\sigma^2}\left(t_1^2 + \frac{y^2}{t_1^2}\right)} dt_1 + \epsilon,$$

where ϵ is a differential of higher order than Δy. This may be proved by computing the difference between the value of (7) when t_2 has constantly the value $(y + \Delta y)/t_1$ and when it has constantly the value y/t_1. If $y = 0$ this change in the right side of (7) is not valid; it is easily seen that in this case the integral on the right of (8) is infinite. It may be shown, however, in this case that

$$\text{(9)} \qquad \int_0^{\Delta y} f(y)\, dy = \tfrac{1}{4} - \frac{1}{2\pi} \int_1^\infty \frac{e^{-\frac{x\Delta y}{\sigma^2}}\, dx}{x\sqrt{x^2 - 1}},$$

and that this is an infinitesimal, and that it is of order as small as one.

Many authors think of (IIIa) as the fundamental formula in the theory of probability distributions. One of the simplest and earliest applications of it was to establish the so-called reproductive property of the normal law: that the sum of two variates is distributed normally if each is distributed normally. Jackson (1935) has used it to establish a similar property for two Type III distributions which have the same exponent of e. Usually this integral is difficult to evaluate when $N > 2$ because of the unsymmetrical form into which it is cast, but when $N = 2$ and there is but one parameter (IIIa) it is perhaps the most convenient of all the formulae.

Theorem IV. An exceedingly useful formula is obtainable from (I) in the following manner. Let $\theta(x_1, \cdots, x_P; \alpha_1, \cdots, \alpha_Q)$ be a finite single valued function of the old parameters (x) and of some new parameters (α). Subject to general conditions to be stated we may write:

$$\text{(IV)} \qquad \int_X \theta f\, dX = \int_T \theta' \phi\, dT,$$

an identity with respect to each α, where θ' is the result of substituting (2) for the x's in θ.

Since this theorem has not been proved in this general form, an outline of the proof will be given. Sufficient conditions are:

(*a*) All the integrals involved shall exist.

(*b*) If p is limited (in the sense that it lies within a finite hypersphere), so is q, and conversely.

Proof. Let X_0 be a limited p set and T_0 the corresponding q set such that both (*c*) and (*d*) hold ($\epsilon > 0$):

$$\left| \int_{X_0} f\theta \, dX - \int_X f\theta \, dX \right| < \epsilon, \tag{c}$$

$$\left| \int_{T_0} \phi\theta' \, dT - \int_T \phi\theta' \, dT \right| < \epsilon. \tag{d}$$

It is easy to see that such an X_0 and a corresponding T_0 do exist, as follows: Let X_0' be a limited set for which (*c*) is true, and for which it will remain true no matter what points are added to X_0'. Similarly, let T_0' be a limited set for which (*d*) is true and for which it will remain true, no matter what points are added to T_0'. Presumably X_0' and T_0' do not correspond to each other, but we may now let X_0 be the totality of all the points of X_0' and of all those points of X corresponding to T_0', and let T_0 be the totality of all the points of T_0' and of all those points of T corresponding to X_0'. Then X_0 and T_0 do correspond to each other and have the desired properties (*c*) and (*d*). Now, since θ is finite, it is limited in X_0. Let

$$| \theta | < H \text{ in } X_0. \tag{e}$$

Divide the interval $(-H, H)$ into s equal subintervals of length h, thus defining in X_0 according to Lebesgue the measurable sets,

p_i ($i = 1, \cdots, s$), and corresponding q_i sets in T_0:

$$\begin{cases} Os\ \theta \leqq h \text{ in } p_i, \\ Os\ \theta' \leqq h \text{ in } q_i. \end{cases} \tag{f}$$

Choose arbitrarily any point of p_i and let k_i be the corresponding value of θ. Then let

$$\bar{\theta} = k_i \text{ in } p_i \ (i = 1, \cdots, s), \text{ and similarly let}$$
$$\bar{\theta}' = k_i \text{ in } q_i \ (i = 1, \cdots, s).$$

Then

$$\int_{X_0} \bar{\theta} f \, dX = \sum_i k_i \int_{p_i} f \, dX,$$

and

$$\int_{T_0} \bar{\theta}'\phi \, dT = \sum_i k_i \int_{q_i} \phi \, dt.$$

Since by (I)

$$\int_{p_i} f \, dX = \int_{q_i} \phi \, dT,$$

$$\int_{X_0} \bar{\theta} f \, dX = \int_{T_0} \bar{\theta}'\phi \, dT. \tag{g}$$

Now

$$\left| \int_{X_0} (\bar{\theta} - \theta) f \, dX \right| \leqq \int_{X_0} | \bar{\theta} - \theta | f \, dX \leqq h \int_{X_0} f \, dX,$$

and

$$\left| \int_{T_0} (\bar{\theta}' - \theta') \, dX \right| \leqq h \int_{T_0} \phi \, dX.$$

So, as h approaches zero both sides of (g) approach limits and their limits are equal:

$$\int_{X_0} \theta f \, dX = \int_{T_0} \theta' \phi \, dT.$$

Hence by (c) and (d) the integrals

$$\int_X \theta f \, dx, \qquad \int_T \theta' \phi \, dT,$$

differ at most by 2ϵ, and so, being independent of ϵ they do not differ at all.

In order to determine the form of f from (IV) one must first evaluate the right side,

$$\int_T \theta\phi \, dt = \psi(\alpha_1, \cdots, \alpha_q);$$

and then solve the integral equation,

$$\int_X \theta f \, dX = \psi. \tag{10}$$

It is the solution of this equation that usually presents the most difficulty. Particular forms of θ that are being used are

$$\theta = e^{\alpha_1 x + \cdots + \alpha_P x_P}, \tag{11}$$

in which case ψ is said to be the "characteristic function" or "moment generating function"; and

$$\theta = x_1^{\alpha_1} \cdots x_P^{\alpha_P}, \tag{12}$$

in which case ψ is a "moment function" or "moment" of f. Other forms might be used. For example, a very convenient method of demonstrating the correctness of the usual formula for the simultaneous distribution of the correlation (x), means (y, z), and variances (u, v), in samples from a normal bivariate universe is by the use of

$$\theta = e^{\alpha_1 (u^2 + v^2 + y^2 + z^2) + \alpha_2 (uvx + yz)}$$

This method of finding f is not a final determination of the probability function desired until it has been shown that the solution is unique, a serious problem in itself; it is one of those which Professor Shohat may consider.[4] There are three methods of solving the integral equation (10):

(*i*) The first might be called guessing. Though unscientific, it is in fact often effective. Especially is it available if the distribution has already been surmised but not demonstrated. Thus, it was open to Student (1908) when he correctly surmised the distribution of the variance. Similarly it was open to Soper (1913) when he incorrectly surmised the distribution of r.

(*ii*) Papers by Romanovsky (1925) and Wilks (1932) have shown how the problem of solving the integral equation may be shifted to the problem of solving a partial differential equation, but this in turn may involve the solution of another equally difficult integral equation in the process or determining the arbitrary function.

(*iii*) If each α be replaced by an imaginary βi and one uses a Fourier transform, one arrives at a set of formulae which are most important. For the case where there is but one x and one β, they may be written:

$$\int_{-\infty}^{\infty} e^{i\beta x} f(x)\, dx = \int_T e^{i\beta g} \phi\, dT = \psi(\beta). \tag{13}$$

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\beta x} \psi(\beta)\, d\beta. \tag{14}$$

Dodd (1925) has given an equivalent set of formulae involving only real variables. It is easy to prove that both sets may be changed to the single formula,

$$f(x) = \frac{1}{\pi} \int_T \phi\, dt \int_0^{\infty} \cos \beta(x - g)\, d\beta. \tag{15}$$

Kullbach (1936) has established the validity of the formulae corresponding to (13) and (14) for the general case of $(P + Q)$ parameters. Wishart and Bartlett (1933) used the general forms to find the distribution of the generalized product moment in samples from an n-dimensional normal system.

When the solution of the integral equations of (IV) cannot be found, one has to put up with the semi-invariants or with the moments of f. Formulae (IV) and (11) yield the semi-invariants, (IV) and (12) the moments about the given origin, and from either of these one may obtain the moments about the mean point. These methods are old but they are still important. Time does not permit me to discuss them, because it would not be proper to close this paper without some reference to limit methods.

Limit Methods. It is well known that the distribution of means of samples taken from almost[5] any universe approaches the normal law as a limit as N becomes infinite. This theorem is subject to great generalizations, as is indicated in papers of A. Liapounoff (1901), S. Bernstein (1926), Romanovsky

[4] In a later paper at the same symposium.

[5] There are exceptions. *E. g.*, means of samples taken from the universe $a/\pi(a + t^2)$ have a distribution identical with the universe itself.

(1929, 1930) and C. C. Craig (1932). Subject to very general conditions it has been shown that: If the characteristic function of one probability distribution contains a parameter and approaches as a limit, uniformly in every finite domain of its variables, the characteristic function of another probability distribution; then the first distribution approaches as a limit the second distribution. Hence S. Bernstein and Romanovsky have shown that: If the universe is an n-way correlation solid of a certain very general type, then the n means obtained by a selection of a sample of N sets of variates, $x_i = \frac{1}{N}(t_{i_1} + \cdots + t_{iN})$, $(i = 1, \cdots, n)$, have a distribution which approaches as a limit a normal correlation solid as N becomes infinite. A similar theorem has been established also in the interesting case of Romanovsky's "belonging coefficients", which include K. Pearson's coefficient of racial likeness. Also, by the method of maximum likelihood, Hotelling (1930) has proved that under certain general conditions all optimum estimates of the parameters of a frequency distribution have a joint distribution approaching the normal as N becomes infinite. The validity of the method of maximum likelihood when used for this purpose has been established by J. L. Doob (1934).

Finally, one may note an apparently new limit theorem of another type. Its general nature will be obvious from the following application:

Let a sample of N be drawn from the universe,

$$\begin{aligned} \phi &= Ae^{-at^{2\lambda}}, \quad \text{if} \quad t > 0, \\ &= 0 \qquad\quad\ \ \text{if} \quad t \leqq 0. \end{aligned}$$

It is readily proved, by means of (IV), that the distribution $f(x)$ of the parameter,

$$x = (t_1^{2\lambda} + \cdots + t_N^{2\lambda})^{1/N}$$

is a curve of the form,

$$\begin{aligned} f(x) &= Bx^{N-1}\, e^{-x^{2\lambda}} \text{ where } x > 0, \\ &= 0 \qquad\qquad\quad \text{elsewhere.} \end{aligned}$$

Now let λ become infinite. The universe approaches as a limit the rectangle:

$$\begin{aligned} \Phi &= A \text{ where } 0 \leqq t < 1, \\ &= 0 \qquad\quad \text{elsewhere.} \end{aligned}$$

The parameter x approaches as a limit X, where X = maximum t_i. The distribution $f(x)$ approaches as a limit the new distribution,

$$\begin{aligned} F(X) &= NX^{N-1} \text{ where } 0 < |X| < 1, \\ &= 0 \qquad\qquad\quad \text{elsewhere.} \end{aligned}$$

Hence we have proved in a new way, what was already known: that the distribution of the greatest variate obtained by sampling from a rectangular universe is of the form $F(X)$.

The limit theorem implicit in this illustration can be established in sufficient generality, but I do not yet know whether it has other applications of value.

REFERENCES

G. A. BAKER, Transformations of bimodal distributions, Annals of Mathematical Statistics, vol. 1 (1932), pp. 334–344.

G. A. BAKER, Transformation of non-normal frequency distributions into normal distributions, Annals of Mathematical Statistics, vol. 5 (1934), pp. 113–123.

S. BERNSTEIN, Sur l'extension du théorème limite du calcul des probabilités aux sommes de quantités dependantes, Mathematische Annalen, vol. 97 (1926), pp. 1–59.

C. C. CRAIG, On the composition of dependent elementary errors, Annals of Mathematics, vol. 33 (1932), pp. 184–206.

C. C. CRAIG, On the frequency function of xy, Annals of Mathematical Statistics, vol. 7 (1936), pp. 1–15.

E. L. DODD, The frequency law of a function of variables with given frequency laws, Annals of Mathematics, vol. 27 (1925), pp. 12–20.

J. L. DOOB, Probability and statistics, Transactions of the American Mathematical Society, vol. 36 (1934), pp. 759–775.

R. A. FISHER, Frequency distributions of the values of the correlation coefficient in samples from an indefinitely large population, Biometrika, vol. 10 (1915), pp. 507–521.

R. A. FISHER, On a distribution yielding the error function of several well known statistics, Proceedings of the International Mathematical Congress, Toronto, 1924, vol. 2, pp. 805–813.

R. A. FISHER, Applications of "Student's" distribution, Metron, vol. 5, no. 3, (1925), pp. 90–104.

H. HOTELLING, The distribution of correlation ratios calculated from random data, Proceedings of the National Academy of Sciences, vol. 11, no. 10 (1925), pp. 657–662.

H. HOTELLING, An application of analysis situs to statistics. Bulletin of the American Mathematical Society, vol. 33 (1927), pp. 467–476.

H. HOTELLING, The consistency and ultimate distribution of optimum statistics, Transactions of the American Mathematical Society, vol. 32 (1930), pp. 847–859.

J. O. IRWIN, Recent advances in mathematical statistics, Journal of the Royal Statistical Society, vol. 98, part 1 (1935), pp. 88–92.

D. JACKSON, Mathematical principles in the theory of small samples, American Mathematical Monthly, vol. 42 (1935), pp. 344–364.

S. KULLBACH, An application of characteristic functions to the distribution problem of statistics, Annals of Mathematical Statistics, vol. 5 (1934), pp. 263–307.

S. KULLBACH, On certain distribution theorems of statistics, Bulletin of the American Mathematical Society, vol. 42 (1936), pp. 407–410.

A. LIAPOUNOFF, Nouvelle forme du théorème sur la limite de probabilité, Memoires de l'Académie de St. Pétersbourg (8), vol. 11 (1901).

P. R. RIDER, A survey of the theory of small samples, Annals of Mathematics, vol. 31 (1930), pp. 577–628.

P. R. RIDER, Recent progress in statistical method, Journal of the American Statistical Association, vol. 30 (1935), pp. 58–88.

H. L. RIETZ, Frequency distributions obtained by certain transformations of normally distributed variates, Annals of Mathematics (2) vol. 23 (1921–22), pp. 292–300.

H. L. RIETZ, On certain properties of frequency distributions of the powers and roots of the variates of a given distribution, Proceedings of the National Academy of Sciences, vol. 13 (1927), pp. 817–820.

U. ROMANOVSKY, On the moments of the standard deviations and of correlation coefficients in samples from a normal population, Metron, Vol. 5(4) (1925), pp. 1–45.

U. ROMANOVSKY, Sur une extension du théorème de A. Liapounoff sur la limite de probabilité, Bulletin de l'Academie des Sciences de l'U. S. S. R. (1929), pp. 209–225.

U. ROMANOVSKY, On the moments of means of functions of one or more random variables, Metron, vol. 8(1–2) (1930), pp. 251–291.

W. A. SHEWHART, Annual Survey of Statistical Technique: Sample Theory, Econometrica, vol. 1 (1933), pp. 225–237.

H. E. SOPER, On the probable error of the correlation coefficient to a second approximation, Biometrika, Vol. 9 (1913), pp. 91–115.

STUDENT, The probable error of the mean, Biometrika, vol. 6 (1908), pp. 1–25.

S. S. WILKS, On the distributions of statistics in samples from a normal population of two variables with matched sampling of one variable, Metron, vol. 9(3–4) (1932), pp. 87–126.

J. WISHART, The generalized product moment distribution in samples from a normal multivariate population, Biometrika, vol. 20A, (1928), pp. 32–52.

J. WISHART AND M. S. BARTLETT, The generalized product moment distribution in a normal system, Proceedings of the Cambridge Philosophical Society, vol. 29 (1933), pp. 260–270.

WESLEYAN UNIVERSITY.