THE PHILIPPINE STATISTICIAN
Volume XVII, Nos. 1/2, 1968

ON THE ALGEBRA OF DISTRIBUTIONS

by

F. M. SIOSON[1]

If mathematics has been said to be the queen of the sciences, then algebra may aptly be called the soul of the queen. For, there exists no mathematical theory that does not possess an algebraic counterpart or that has no algebraization. The very notion of limit of a sequence of real numbers so fundamental in analysis is an example of an infinitary algebraic operation. An integral is a linear functional on a vector lattice over the reals or complexes of real-valued or complex-valued functions and the derivative is a linear transformation on the vector space of all differentiable functions. Any logical system is co-extensive with a corresponding algebraic system whose algebraic properties are translatable into corresponding properties of the logical system. For example, to the propositional calculus corresponds a Boolean algebra; to the functional calculus, a cylindric or polyadic algebra; and to intuitionistic logics, certain distributive lattices called Heyting algebras. Completeness of logical systems translate into certain structural properties of their algebraic systems. Geometry itself is defined as the study of properties that remain unchanged under a group (an algebraic system) of transformations acting on a set of points called space.

In this short note we shall be concerned with certain algebraizations that naturally arise in theory of probability and statistics. To completely appreciate the significance of this approach, we shall start from the very foundations of the theory itself.

[1] The author is a Mathematical Consultant of the Bureau of the Census and Statistics. The present communication was done by the author under this capacity and is part of a monograph to be published by the Bureau.

Recall (see [5], for instance) that a **probability space** is a triple (U, Σ, P) consisting of a set U (called the **sample space**), a family Σ of subsets of U closed under the operations of arbitrary unions and arbitrary intersections (called the **field of events**), and a function P: $\Sigma \rightarrow [0,1]$ (called a **probability**) such that $P(U) = 1$ and $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$ for every family of pairwise disjoint subsets A_i of U belonging to Σ. For example, if U coincides with the set R of real numbers, then a possible Σ is the family B of so-called **Borel subsets** of the real line. B is the smallest family of subsets of R containing all open, closed, half-closed, and half-open intervals in R and which is closed under the operations of countable intersections and countable unions. The probability function P on B is then simply its measure (or length). A **random variable** in a probability space (U, Σ, P) is a function X: N $\rightarrow$ R (or more correctly a function X from the probability space (U, Σ, P) into the probability space (R, B, P) such that $X^{-1}(S) \epsilon \Sigma$ for all subsets S ϵ B. Since $(-\infty, x)$ is an element of the family B, them note that $X^{-1}((-\infty, x))$ is a random event in Σ and $P(X^{-1}(-\infty, x))$ is well-defined for each real number x. We will, by convention, simply denote this probability by

$$P[X < x] = P[u \mid u \epsilon U, X(u) < x].$$

The function F: R $\rightarrow$ [0,1] defined by $F(x) = P[X < x]$ is called a **(frequency) distribution** function of the random variable or simply a **(probability) distribution**. The concept is completely characterized by the following.

DEFINITION 1. A real-valued function F of a real variable is a distribution if and only if it left-continuous, monotone non-decreasing, and such that $F(-\infty) = \lim_{x \rightarrow -\infty} F(x) = 0$ and $F(\infty) = \lim_{x \rightarrow \infty} F(x) = 1$.

A well-known way of classifying distributions subdivides the class of all distributions into three distinct types: (1) the discrete distributions that possess only a countable number of possible values, e.g. the Poisson, binomial, multinomial, negative binomial, Simpson's or triangular distributions; (2)

the absolutely continuous distributions F for which there exist non-negative continuous functions p such that

$$F(x) = \int_{-\infty}^{x} p(z)dz$$

for all real z, e.g. the normal, log-normal, Laplacian, Maxwell's, Cauchy's, Students's t, and chi-square distributions; and (3) those that are neither discrete nor continuous, e.g. distribubutions that increases only at points of set of Lebesgue measure zero like the Cantor ternary set. Distributions may also be classified according to whether they are momentless, have moments of all orders, or have moments up to a certain order.

If X and Y are two arbitrary random variables in the same probability space, then the function X + Y defined by the equation $(X + Y)(u) = X(u) + Y(u)$ for all u is also a random variable in the same space. When X has the distribution F and Y the distribution G, then X + Y has the distribution F*G given by the Lebesgue-Stialtjes integral

$$(F * G)(x) = \int_{-\infty}^{\infty} F(x - z)dG(z).$$

In case the integrand is continuous, the above integral coincides with the more familiar Riemann-Stieltjes integral. In any case it is easy to show that the operation * between two distributions satisfies the following properties:

(a) $(F * G) * H = F * (G * H)$,

(b) $F * G = G * F$,

for all distributions F, G, and H. The function E defined by

$$E(x) = \begin{cases} 0 & \text{for all } x \leqq 0, \\ 1 & \text{for all } x > 0, \end{cases}$$

is clearly also a distribution and

(c) $E * F = F = F * E$ for all distributions F.

If $\mathscr{D}$ denotes the family of all distributions in the probability space (R, B, P), then statement (a) says ($\mathscr{D}$, *) is a semigroup which by (b) is commutative and possesses an identity element E by (c). Clearly, the set $\mathscr{X}$ of all random variables over (R, B, P) is also a commutative semigroup with an identity element. If ω: $\mathscr{X} \rightarrow \mathscr{D}$ is the function that associates with each random variable its distribution, then obviously this function is both one-to-one and onto and, moreover, $\omega(x + Y) = \omega(X) * \omega(Y)$. This function ω is then called an isomorphism between the two semigroups. Algebraically speaking, we say then that the two semigroups are essentially the same since every property (algebraic) of one is also a property of the other.

The study of the operation * (called convolution) between two distributions is one of the main tasks of probability theory and mathematical statistics. The various limit theorems and laws of large numbers in statistics are in fact algebraic theorems on the nature of this operation when it is repeated indefinitely many times. The literature in probability theory is replete with theorems concerning this operation. Algebraically, the familiar theorem of Cramer simply states that if the convolution of two distributions is normal then each one of its (convoluted) factors is also normal. The analogous result for Poisson distributions is known as Raikov's theorem. Another result by Khintchine states that every distribution is the (convoluted) product (not necessarily unique) of indecomposable distributions and infinitely divisible distributions. A distribution F is said to be infinitely divisible if and only for each integer n, $F = G * G * \ldots * G$ (n times). The Poisson and normal distributions, for instance, are infinitely divisible. Any infinite convolution of distributions, if it exists, is infinitely divisible.

Another concept quite fundamental in the study of distributions is that of a characteristic function.

DEFINITION 2. A complex-valued function f of a real variable is called a **characteristic function** if and only if it is continuous, $f(0) = 1$, and for any integer n, any set of real numbers $r_1, r_2, \ldots, r_n$, and complex numbers $z_1, z_2, \ldots, z_n$, one has

$$\sum_{k=1}^{n} \sum_{j=1}^{n} f(r_k - r_j)\bar{z}_j z_k \geqq 0.$$

A well-known result in probability and now be stated algebraically as follows:

THEOREM 1. The family $\mathscr{C}$ of all characteristic functions constitute a semigroup under the operation of pointwise multiplication (denoted by.) isomorphic with the semigroup $\mathscr{D}$ of distributions under the function ω: $\mathscr{D} \rightarrow \mathscr{C}$ with $\omega(F) = f$ such that

$$f(t) = \int_{-\infty}^{\infty} \exp\,(itx)\ dF(x).$$

The main result of this paper is formulated in the next.

THEOREM 2. Although the semigroup $\mathscr{D}$ of all distributions over the probability space (R, B, P) does not satisfy the cancellation laws, it is the disjoint union of subsets each of which a semigroup with cancellation law under the same operation in $\mathscr{D}$.

The proof of the above theorem will utilize the following two Lemmas.

LEMMA A. The commutative semigroup $\mathscr{D}$ of one-dimensional distributions over the reals does not satisfy the cancellation laws.

Proof. This result was first shown by B. V. Gnedenko [1] and A. Ya. Khintchine [2] who proved the existence of two distinct distributions F and G such that

$$F * F = F * G.$$

Their examples were, however, rather complicated and the explicit form of one of their distributions has never been exhibited (see page 83 of reference [3]). We propose to exhibit here another such pair which are not only simpler but also more explicit.

Let F be the absolutely continuous distribution function defined by the density function

$$p(x) = \frac{1}{4\pi} \left[\frac{2 \sin \pi x}{x} + \frac{\sin \pi(x+1)}{x+1} + \frac{\sin \pi(x-1)}{x-1} \right]$$

and G be the discrete distribution function such that

$$G(x) = \begin{cases} 0 & \text{for} \quad x \leqq -1, \\ 1/4 & \text{for} \quad -1 < x \leqq 0, \\ 3/4 & \text{for} \quad 0 < x \leqq 1, \\ 1 & \text{for} \quad x > 1. \end{cases}$$

By direct computation, it can be shown though rather laboriously that $F * F = F * G$. Of course, by Theorem 1, this can also be shown via characteristic functions. The characteristic function g of G is given by

$$g(t) = \int_{-\infty}^{\infty} \exp(itx) dG(x) = \tfrac{1}{2} + \tfrac{1}{4} (\exp(it) + \exp(-it)) =$$

$$\tfrac{1}{2}(1 + \cos t) = \cos^2 (t/2).$$

That of F is computed as follows:

$$f(t) = \int_{-\infty}^{\infty} \exp(itx) p(x) dx = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(itx) \sin \pi x \, dx/x$$

$$+ \frac{1}{4\pi} \int_{-\infty}^{\infty} \exp(itx) \sin \pi(x+1) \, dx/(x+1) +$$

$$\frac{1}{4\pi} \int_{-\infty}^{\infty} \exp(itx) \sin \pi(x-1) dx/(x-1.$$

By making the substitution $y = \pi(x + 1)$ in the middle integral and $y = \pi(x - 1)$ in the last integral, and $y = \pi x$ in the first, we get $f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(ity/\pi) \sin y \, dy/y +$

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(ity/\pi)[(\exp it + \exp -it)/2][\sin y/y]dy$$

$$= \frac{1}{2\pi}(1 + \cos t) \int_{-\infty}^{\infty} \cos(ty/\pi) [\sin y/yldy$$

$$= \begin{cases} \cos^2(t/2) & \text{for } |t| \leqq \pi, \\ 0 & \text{for } |t| > \pi. \end{cases}$$

Clearly, $f(t)f(t) = f(t)g(t)$ for all t. Hence by Theorem 1 we have $F * F = F * G$ with F/G. For the evaluation of the above integral, refer to any good table of integrals.

LEMMA B. For any pair of distributions F and G in $\mathcal{D}$. if $F * F = F * G = G * G$, then $F = G$.

Proof. We again use Theorem 1. Let f and g correspond respectively to F and G under the isomorphism between $\mathcal{D}$ and $\mathcal{C}$. Suppose now that we have $F * F = F * G = G * G$, but $F \neq G$ or in terms of the above isomorphism we have $f^2 = fg = g^2$ but $f \neq g$. From this last inequality, then there exists a real number r such that $f(r) \neq g(r)$ or $f(r) - g(r) \neq 0$. Thus since $f^2 - fg = 0$ and $fg - g^2 = 0$, then

$$f(r)[f(r) - g(r)] = 0 \text{ and } [f(r) - g(r)]g(r) = 0.$$

Whence $f(r) = 0$ and $g(r) = 0$ and therefore $f(r) = g(r)$, a contradiction! The Lemma therefore follows.

Proof of Theorem 2. The proof of Theorem 2 should follow by now from Lemmas A and B by applying a result [4] of Stefan Schwarz. Due to the relative unavailability of this paper, however, we shall reproduce a complete proof here.

The first part of the theorem follows from Lemma A. To prove the main and last part, we now use Lemma B. Let θ be a relation on $\mathscr{S}$such that $(F, G) \in \theta$ if and only if there exist natural numbers m and n such that $F^m = G^n$, where $F^m = F * F * \ldots * F$ (m times) and similarly for G^n.

Clearly $(F, F) \in \theta$ and $(F, G) \in \theta$ implies $(G, F) \in \theta$, that is, θ is both reflexive and symmetric. It is also transitive. For, suppose $(F, G) \in \theta$ and $(G, H) \in \theta$ so that $F^m = G^n$ and $G^r = H^s$ for some natural numbers m, n, r, and s. Then $F^{mr} = G^{nr} = (G^r)^n = (H^s)^n = H^{ns}$. Whence $(F, H) \in \theta$ and θ is an equivalence relation. Thus, $\mathscr{S}$is subdivided into the disjoint equivalence classes by the relation θ.

Let C be any one of these equivalence classes. If $F, G \in C$, then clearly $(F, G) \in \theta$. To show that C is a semigroup, it will then suffice to show that $F * G$ also belong to C. Since $(F, G) \in \theta$, then $F^m = G^n$ for some integers m and n. Then $(F * G)^n = F * G * \ldots * F * G$ (n times) $= F^n * G^n = F^n * F^m = F^{n+m}$. Whence $(F{*}G, F) \in \theta$ and $F * G \in C$.

We shall next show that C has the cancellation laws. Suppose $F, G, H \in C$ and $F * H = G * H$. Then for some natural numbers m, n, r, and s we have

$$F^m = H^n \quad \text{and} \quad G^r = H^s.$$

Note that $F^{m+1} = F * F^m = F * H^n = F * H * H^{n-1} = G * H * H^{n-1} = G * H^n = G * F^m = F^m * G$. This shows that there are numbers m such that $F^{m+1} = F^m * G$. Since any set of natural numbers has a smallest member, let p be the smallest number such that $F^{p+1} = F^p * G$. We divide the proof into two cases.

Case I. Suppose p is even. Then $(F^{p/2} * G)^2 = F^p * G * G = F^{+p1} * G= F * F^p * G = F * F^{p+1} = F^{p+2} = (F^{p/2+1})^2$. Also $(F^{p/2} * G) = F^{p+1} * G = (F^{p/2} * G){*}(F^{p/2\ +1})$ so that $(F^{p/2} * G)^2 = (F^{p/2} * G) * (F^{p/2+1}) = (F^{p/2+1})^2$. By Lemma B, then $F^{p/2+1} = F^{p/2} * G$, which is a contradiction of the choice of p. Whence

Case II: p is odd. In this case then $p+1 = q$ is even and since $F^{p+1} = F^{p} * G$, then $F^{q+1} = F^{q} * G$ where q is even. Repeating the same argument in Case I, we then obtain $F^{q/2+1} = F^{q/2} * G$. By virtue of the choice of p, then $q/2 \geqq p$ or $(p + 1)/2 \geqq p$. This means that $p + 1 \geqq 2p$ or $p \leqq 1$. Since in any case $p \geqq 1$, then $p = 1$. This means that we have exactly shown that $F * G = F^2$. Analogously, by symmetry, we should be able to show that $F * G = G^2$. Whence $F^2 = F * G = G^2$. Whence by Lemma B, we get $F = G$. The cancellation law therefore holds in C.

The final result follows, since C is arbitrary.

REFERENCES

[1]. GNEDENKO, B. V. "On characteristic functions", *Bulletin of the Mathematical Society of Moscow State University,* Volume 1, Number 5 (1937) pp. 17-18 (in Russian).

[2]. KHINTCHINE, A. Ya. "On a criterion for characteristic functions", *Bulletin of the Mathematical Society of Moscow State University,* Volume 1, Number 5 (1937) p. 18.

[3]. LINNIK, Yu. V. *Decomposition of Probability Distributions.* Edinburgh-London: Oliver and Boyd, 1964. First English Edition.

[4]. SCHWARZ, Stefan. "Semigroups satisfying some weakened forms of the cancellation laws", *Mat. Fyz. Casopis Sloven. Akad. Vied,* Volume 6 (1956) pp. 149-158 (In Slovak, English summary).

[5]. SIOSON, F. M. "Matrix-valued probability theory", *The Philippine Statistician,* Volume 16 (1967) pp. 1-9.

[6]. SIOSON, F. M. *The Algebra of Distribution Functions.* (to appear). Manila: The Bureau of Printing, 1968. (A Monograph of the Bureau of the Census and Statistics.)