

## ON CERTAIN DISTRIBUTION THEOREMS OF STATISTICS\*

BY SOLOMON KULLBACK

1. *Introduction.* The problems of statistics have been classified under three heads: problems of specification, problems of estimation, problems of distribution.† I intend to prove certain theorems regarding the problem of distribution in more general form than that in which they have been proved before.

By the distribution problem of statistics is meant the problem of finding the distribution function, or law of total probability, of the real, single-valued functions  $\mu_\alpha(x_1, x_2, \dots, x_n)$ , ( $\alpha = 1, 2, \dots, r$ ), given the distribution function of the real variables  $x_1, x_2, \dots, x_n$ . Many particular problems of this nature have been solved. More general solutions have been obtained, imposing certain restrictions on the functions  $\mu_\alpha(x_1, x_2, \dots, x_n)$  and the distribution function of  $x_1, x_2, \dots, x_n$ .‡

2. *Distribution Functions.* By the distribution function, or law of total probability of the real variables  $x_1, x_2, \dots, x_n$ , is meant the monotone, absolutely additive set function  $\phi(E)$ , where  $0 \leq \phi(E) \leq 1$ , and  $\phi(S) = \int_S d_{x_1 x_2 \dots x_n} \phi(E) = 1$ , where  $S$  is the whole  $x_1 x_2 \dots x_n$  space.§

---

\* Presented to the Society, April 19, 1935, under the title: *On the distribution problem of statistics.*

† R. A. Fisher, *On the mathematical foundations of theoretical statistics*, Philosophical Transactions of the Royal Society, vol. 222 A (1922), p. 309 ff.

‡ C. V. L. Charlier, *Some fundamental problems in the theory of probability*, Arkiv för Matematik, Astronomi och Fysik, vol. 8 (1912), No. 4, pp. 1–39; T. Kameda, I. *Eine Verallgemeinerung des Poissonschen Problems in der Wahrscheinlichkeitsrechnung*, Proceedings Physico-Mathematical Society, Tokyo, vol. 9 (1917–18), pp. 155 ff.; II. *Theory of generating functions and its application to the theory of probability*, Journal Faculty of Science, University of Tokyo, sec. I, vol. 1 (1925), part 1, pp. 1–62; E. L. Dodd, *The frequency law of a function of variables with given frequency laws*, Annals of Mathematics, (2), vol. 27 (1925), pp. 12–20; S. Kullback, *An application of characteristic functions to the distribution problem of statistics*, Annals of Mathematical Statistics, vol. 5 (1934), pp. 263–307.

§ See E. K. Haviland, I. *On the theory of absolutely additive distribution functions*, American Journal of Mathematics, vol. 56 (1934), pp. 625–658, particularly pp. 627–628.

3. *The Distribution Function of the Functions*  $\mu_\alpha(x_1, x_2, \dots, x_n)$ . For convenience we shall consider the case of two functions; the problem for  $r$  functions is solved by the same methods.

Let  $\mu_\alpha(x_1, x_2, \dots, x_n)$ , ( $\alpha = 1, 2$ ), be real, single-valued functions of the real variables  $x_1, x_2, \dots, x_n$ , where the distribution function of the  $x_i$ , ( $i = 1, 2, \dots, n$ ), is  $\phi(E)$ . Let the functions  $\mu_\alpha(x_1, x_2, \dots, x_n)$ , ( $\alpha = 1, 2$ ), be measurable with respect to  $\phi(E)$ , that is, for every number  $A$ , the set of those points  $P$  in which  $\mu_\alpha(P) > A$ , belongs to the definition region  $T$  of  $\phi$ .\*

For convenience we shall write  $\mu_\alpha(x)$  for  $\mu_\alpha(x_1, x_2, \dots, x_n)$ ,  $t\mu(x)$  for  $t_1\mu_1(x_1, x_2, \dots, x_n) + t_2\mu_2(x_1, x_2, \dots, x_n)$  and  $d_x\phi(E)$  for  $d_{x_1x_2\dots x_n}\phi(E)$ . Accordingly, since  $e^{it\mu(x)}$  is measurable with respect to  $\phi(E)$  and *eo ipso* summable with respect to  $\phi(E)$ ,† the integral  $\int_R e^{it\mu(x)} d_x\phi(E)$ , where  $i = \sqrt{-1}$ , exists for all values of  $t_1$  and  $t_2$ , and is Radon's generalization of the Lebesgue-Stieltjes integral.‡ Since  $\phi(E)$  is of finite total variation, the limit of the above integral exists as  $R \rightarrow S$ , where  $R$  is an arbitrarily large rectangle and  $S$  is the entire  $x_1x_2 \dots x_n$ -space, and we write

$$(1) \quad \Lambda(t_1, t_2; \mu_\alpha, \phi) = \int_S e^{it\mu(x)} d_x\phi(E).$$

We must now define the appropriate distribution function in the  $\mu_1\mu_2$ -space. Let  $E'$  be any set in the  $\mu_1\mu_2$ -space contained in an arbitrarily large rectangle, and let  $E$  be the corresponding set in the  $x_1x_2 \dots x_n$ -space.

We then define the set function  $\Phi$  by  $\Phi(E') = \phi(E)$ . Since  $\phi(E)$  is a distribution function and  $\mu_\alpha(x)$ , ( $\alpha = 1, 2$ ), are single-valued functions of  $x_1, x_2, \dots, x_n$ , it follows that  $0 \leq \Phi(E') \leq 1$  and that  $\Phi$  is absolutely additive. We shall show below that  $\Phi(S') = 1$ , where  $S'$  is the whole  $\mu_1\mu_2$ -space. Thus  $\Phi(E')$  is the appropriate distribution function in the  $\mu_1\mu_2$ -space.

\* J. Radon, *Theorie und Anwendungen der absolut additiven Mengenfunktionen*, Sitzungsberichte der Mathematischen-Naturwissenschaftlichen Klasse der Kaiserlichen Akademie zu Wien, vol. 122 (1913), pp. 1295–1438, particularly pp. 1299 and 1325. E. Hopf, *On causality, statistics and probability*, Journal of Mathematics and Physics, Massachusetts Institute of Technology, vol. 13 (1934), pp. 51–102, particularly §1. J. L. Doob, *Probability and statistics*, Transactions of this Society, vol. 36 (1934), pp. 759–775.

† J. Radon, loc. cit., p. 1326.

‡ J. Radon, loc. cit., pp. 1322–1332.

4. *The Characteristic Function of the Distribution Function of  $\mu_\alpha(x)$ .*\* We shall now show that

$$(2) \quad \int_S e^{it\mu(x)} d_x \phi(E) = \int_{S'} e^{it\mu} d_\mu \Phi(E'),$$

where  $t\mu \equiv t_1\mu_1 + t_2\mu_2$  and  $d_\mu \Phi(E') \equiv d_{\mu_1\mu_2} \Phi(E')$ .

Separate the real and imaginary parts of (2) and consider, for example, the approximating sums†

$$S_n = \sum_{k=-\infty}^{\infty} y_k^{(n)} \phi(E_k^{(n)}),$$

where  $E_k^{(n)}$  is the set of points  $(x_1, x_2, \dots, x_n)$  for which

$$y_k^{(n)} \leq \cos(t\mu(x)) < y_{k+1}^{(n)}, \quad \text{and} \quad S_n' = \sum_{k=-\infty}^{\infty} y_k^{(n)} \Phi(E_k'^{(n)}),$$

where  $E_k'^{(n)}$  is the set of points  $(\mu_1, \mu_2)$  for which  $y_k^{(n)} \leq \cos(t\mu) < y_{k+1}^{(n)}$ . The set  $E_k'^{(n)}$  belongs to the domain of definition of  $\phi$  and  $E_k'^{(n)}$  to that of  $\Phi$ . The equality of the approximating sums follows from the fact that in accordance with the definition in §3,  $\phi(E_k^{(n)}) = \Phi(E_k'^{(n)})$ . A similar result follows for the imaginary part, and thus (2) is established.

If in (2) we set  $t_1 = t_2 = 0$ , we have

$$\int_{S'} d_\mu \Phi(E') = \int_S d_x \phi(E) = 1,$$

that is,  $\Phi(S') = 1$ , as was mentioned above.

5. *Application of an Inversion Formula.* By E. K. Haviland's inversion formula for Fourier-Stieltjes transforms,‡ we have

$$(3) \quad g(\nu_1, \nu_2) = \frac{1}{(2\pi)^2} \left( \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \right) \frac{(e^{-it_1\xi_1} - 1)(e^{-it_2\xi_2} - 1)}{(it_1)(it_2)} e^{-it_1\nu_1 - it_2\nu_2} \cdot \Lambda(t_1, t_2; \Phi) dt_1 dt_2,$$

where

\* P. Lévy, *Calcul des Probabilités*, 1925, p. 161.

† J. Radon, loc. cit., pp. 1326-1328.

‡ E. K. Haviland, II. *On the inversion formula for Fourier-Stieltjes transforms in more than one dimension*, American Journal of Mathematics, vol. 57 (1935), pp. 94-100.

$$g(\nu_1, \nu_2) = \Phi(Q + P_{\nu_1\nu_2}) = G(\nu_1 + \xi_1, \nu_2 + \xi_2) - G(\nu_1 + \xi_1, \nu_2) \\ - G(\nu_1, \nu_2 + \xi_2) + G(\nu_1, \nu_2),$$

where  $Q$  is  $(0 \leq \omega_1 < \xi_1; 0 \leq \omega_2 < \xi_2)$ , and  $Q + P_{\nu_1\nu_2}$  is a non-singular rectangle\* of  $\Phi$ , where also  $G(\mu_1, \mu_2)$  is the point function corresponding to  $\Phi(E')$ ,

$$\Lambda(t_1, t_2; \Phi) = \int_{S'} e^{it\mu} d_\mu \Phi(E'),$$

and the double integral in (3) is a Cauchy principal value, as the parentheses indicate. We thus have the following theorem

**THEOREM.** *If  $\mu_\alpha(x)$ , ( $\alpha = 1, 2$ ), are real, single-valued functions which are measurable with respect to the monotone, absolutely additive set function  $\phi(E)$  which is the distribution function of the real variables  $x_1, x_2, \dots, x_n$ , then (a) the characteristic function of the distribution function of  $\mu_1, \mu_2$  is given by*

$$\Lambda(t_1, t_2; \mu_\alpha, \phi) = \int_S e^{it\mu(x)} d_x \phi(E),$$

and (b) the distribution function of  $\mu_1, \mu_2$  is essentially† determined by setting  $\Phi(Q + P_{\nu_1\nu_2})$  equal to the expression on the right-hand side of (3), with  $\Lambda(t_1, t_2; \mu_\alpha, \phi)$  in place of  $\Lambda(t_1, t_2; \Phi)$ .

If  $\phi(E)$  is absolutely continuous and  $\Lambda(t_1, t_2; \mu_\alpha, \phi)$  is absolutely integrable, then‡

$$(4) \quad h(\mu) = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-it\mu} \Lambda(t_1, t_2) dt_1 dt_2,$$

where

$$\Lambda(t_1, t_2) = \int_{-\infty}^{\infty} e^{it\mu(x)} f(x) dx, \quad dx = dx_1 dx_2 \cdots dx_n,$$

and where  $f(x) \equiv f(x_1, x_2, \dots, x_n)$  and  $h(\mu) \equiv h(\mu_1, \mu_2)$  are the respective probability densities in the  $x_1 x_2 \cdots x_n$  space and  $\mu_1 \mu_2$ -space.§

WASHINGTON, D. C.

\* E. K. Haviland, loc. cit., I, p. 627-628.

† See E. K. Haviland, loc. cit., II, p. 99.

‡ See E. K. Haviland, loc. cit., II, p. 100.

§ See the references on distribution functions cited in the Introduction.