THE USE OF GENERALISED DIRICHLET'S INTEGRAL IN SOLVING SOME DISTRIBUTION PROBLEMS OF STATISTICS*

BY P. V. KRISHNA IYER

(Imperial Agricultural Research Institute, New Delhi)

Received November 25, 1943

(Communicated by Rao Bahadur B. Viswa Nath, F.A.SC.)

1. INTRODUCTION

THE distribution of the mean, the standard deviation and other statistical constants of a sample drawn from a known population can be derived by various methods. Of these mention may be made of the geometrical methods due to Professor Fisher by the aid of which some of the distributions of certain statistical estimates in the case of normal populations can be determined very easily. It is also known that the use of the characteristic and moment generating functions enables us to get the distribution of estimates of *statistics* of samples belonging to normal and other populations. This has been applied in a number of cases by Irwin, Kullback and Wishart and Barttlet and others. The purpose of the present paper is to show how the generalised Dirichlet's integral enables us to evaluate the distribution of certain statistical constants of samples drawn from normal and non-normal populations. Kosambi has used this integral to get the distribution of the ratio of the generalised variances of two samples from a bivariate population.

The integral referred to is the well-known multiple integral

$$\int\int\cdot\cdot\int \mathrm{F}\left\{\left(\frac{x_1}{a_1}\right)^{p_1}+\left(\frac{x_2}{a_2}\right)^{p_2}\cdot\cdot\left(\frac{x_n}{a_n}\right)^{p_n}\right\} x_1^{l_1-1}\, x_2^{l_2-1}\cdot\cdot x_n^{l_n-1}\, dx_1 dx_2\cdot\cdot dx_n\ldots$$

$$=\frac{a_1^{l_1} a_2^{l_2}\cdot\cdot a_n^{l_n}}{p_1 p_2\cdot\cdot p_n}\,\frac{\Gamma\frac{l_1}{p_1}\,\Gamma\frac{l_2}{p_2}\cdot\cdot\Gamma\frac{l_n}{p_n}}{\Gamma\frac{l_1}{p_1}+\frac{l_1}{p_1}\cdot\cdot+\frac{l_n}{p_n}}\int_0^{\mathrm{W}} \mathrm{W}^{\frac{l_1}{p_1}+\frac{l_2}{p_2}\cdot\cdot\frac{l_n}{p_n}-1}\,\mathrm{F(W)}\,d\mathrm{W} \tag{1}$$

where $x_1, x_2, \ldots x_n$ extend to all positive values subject to the condition

$$\left(\frac{x_1}{a_1}\right)^{p_1}+\left(\frac{x_2}{a_2}\right)^{p_2}+\ldots\left(\frac{x_n}{a_n}\right)^{p_n}=\mathrm{W}, \tag{2}$$

---

* Part of this paper was read before the 28th Session of the Indian Science Congress held at Benares, January 1941, *vide Proceedings*, Pt. III, p. 11. (2) Throughout this paper m and σ^2 stand for the mean and the variance of the population, while x and v or s^2 denote the mean and the variance of the sample on the basis of the size of the sample.

It may be noted that it follows from (1) that the chance of W lying between W and W + dW subject to condition (2) is proportional to

$$W^{\frac{l_1}{p_1}+\frac{l_2}{p_2}\cdots\frac{l_n}{p_n}-1} F(W)\,dW \tag{3}$$

2. Normal Population

The probability that a sample of values $x_1, x_2 \ldots . x_n$ should fall within the elementary ranges $dx_1\, dx_2, \ldots . dx_n$ is

$$\frac{1}{(\sqrt{2\pi}\,\sigma)^n}\,\text{Exp.} - \frac{1}{2\sigma^2}\sum_1^n (x_r - m)^2 \cdot dx_1\, dx_2 \ldots . dx_n \tag{4}$$

Putting $\frac{x_1 + x_2 \ldots . x_n}{n} = \bar{x}$, expression (4) can be transformed into

$$\frac{n}{(\sqrt{2\pi}\,\sigma)^n}\,\text{Exp.} - \frac{1}{2\sigma^2}\left\{\sum_1^n (x_r - \bar{x})^2 + n(\bar{x} - m)^2\right\} dx_1\, dx_2 \ldots dx_{n-1}\, dx \tag{5}$$

Now it can be shown, as has been done by Fisher, that

$$\frac{1}{n}\sum_1^n (x_r - \bar{x})^2 = \frac{1}{n}\sum_1^{n-1} \xi_r^2 = v, \text{ where}$$

$$\xi_r = \frac{\sum_1^r x_r - r\,x_{r+1}}{\sqrt{r(r+1)}}.$$

$\xi_1, \xi_2 \ldots . \xi_{n-1}$ are all independent and orthogonal.

Expression (5) in view of the fact $J = \frac{1}{\sqrt{n}}$ reduces to

$$\frac{\sqrt{n}}{(\sqrt{2\pi}\,\sigma)^n}\,\text{Exp.} - \frac{1}{2\sigma^2}\sum_1^{n-1}\xi_r^2 \cdot d\xi_1\, d\xi_2 \ldots . d\xi_{n-1} \cdot \text{Exp.} - \frac{n}{2\sigma^2}(\bar{x} - m)^2\, d\bar{x}$$

using (1) we find that the chance of V lying between V and V + dV and $\bar{x}$ between $\bar{x}$ and $\bar{x} + d\bar{x}$ so that $\xi_1, \xi_2, \ldots . \xi_{n-1}$ are all positive is

$$\frac{\sqrt{n}}{(\sqrt{2\pi}\,\sigma)^n}\,\frac{(n-1)}{2^{n-1}}\,\frac{(\Gamma\frac{1}{2})^{n-1}}{\Gamma\frac{n-1}{2}}\,\text{Exp.} - \frac{nV}{2\sigma^2}\cdot V^{\frac{n-1}{2}-1}\,\text{Exp.} - \frac{n}{2\sigma^2}(\bar{x} - m)^2 \cdot dV\, d\bar{x} \tag{6}$$

Since $\xi_1, \xi_2, \ldots ., \xi_{n-1}$ can take negative values as well we should integrate for negative values also in which case (6) is to be multiplied by 2 for each integration. Therefore to get the final distribution of $\bar{x}$ and V (6) is to be multiplied by 2^{n-1}.

It follows from (6) that the distributions of V and $\bar{x}$ are given by

$$\text{A. Exp.} - \frac{nV}{2\sigma^2}\cdot V^{\frac{n-1}{2}}\, dV \tag{7}$$

and $$\text{B. Exp.} - \frac{n(\bar{x}-m)^2}{2\sigma^2}\cdot d\bar{x} \tag{8}$$

respectively, where A and B are constants which can be easily evaluated.

3. Bivariate Population

The frequency distribution in the case of a bivariate population is given by

$$\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\,\text{Exp.} - \frac{1}{2(1-\rho^2)}\left\{\frac{(x-m_1)^2}{\sigma_1^2} - \frac{2\rho(x-m_1)(y-m_2)}{\sigma_1\sigma_2}\right.$$
$$\left. + \frac{(y-m_2)^2}{\sigma_2^2}\right\}\cdot dx\,dy = \frac{1}{\sigma_1\sqrt{2\pi(1-\rho^2)}}\cdot\text{Exp.} - \frac{z^2}{2\sigma_1^2(1-\rho^2)}\cdot dz$$
$$\times\frac{1}{\sqrt{2\pi}\,\sigma_2}\,\text{Exp.} - \frac{(y-m_2)^2}{2\sigma_2^2}\cdot dy \tag{9}$$

where $$z = (x-m_1) - \frac{\rho\sigma_1}{\sigma_2}(y-m_2).$$

From (9), as is well known, it follows that the distribution is the product of two independent distributions.

Taking now, a bivariate sample of n observations, let $x_1, x_2 \ldots . x_n$ represent the sample of the x-variate, and $y_1, y_2, \ldots . y_n$ be the corresponding values of the y-variate. The chance that $x_1, y_1; x_2, y_2; \ldots . x_n, y_n$; should fall within the infinitesimal ranges $dx_1, dy_1; dx_2, dy_2; \ldots . dx_n, dy_n$ is

$$\frac{1}{(2\pi)^{\frac{n}{2}}(\sigma_1\sqrt{1-\rho^2})^n}\,\text{Exp.} - \frac{\sum_1^n z_r^2}{2\sigma_1^2(1-\rho^2)}\cdot dz_1\,dz_2\ldots . dz_n \times$$
$$\frac{1}{(2\pi)^{\frac{n}{2}}\sigma_2^n}\,\text{Exp.} - \frac{\sum_1^n (y-m_2)^2}{2\sigma_2^2}\cdot dy_1\,dy_2\ldots . dy_n \tag{10}$$
$$z_r = (x_r - m_1) - \frac{\rho\sigma_1}{\sigma_2}(y_r - m_2)$$

Using (6), the second half of (10) reduces to

$$\text{C. Exp.} - \left\{\frac{nv_2 + n(\bar{y}-m_2)^2}{2\sigma_2^2}\right\}\,V_2^{\frac{n-3}{2}}\,dV_2\,d\bar{y} \tag{11}$$

C is a constant.

Turning to the first part of (10),

$\sum_1^n z_r^2 = \sum_1^n w_r^2 + n\bar{z}^2$, $w_r = (x_r - \bar{x}) - \frac{\rho\sigma_1}{\sigma_2}(y - \bar{y})$ and $\bar{z} = (\bar{x} - m_1) - \frac{\rho\sigma_1}{\sigma_2}(\bar{y} - m_2)$. Again $\sum_1^n w_r^2 = \sum_1^n t_r^2 + n(b-\beta)^2 s_2^2$, $b = \frac{rs_1}{s_2}\beta = \frac{\rho\sigma_1}{\sigma_2}$, $t_r =$

$(x_r - \bar{x}) - b(y_r - \bar{y})$, s_1^2 and s_2^2 are the variances of the sample (not on the basis of degrees of freedom but on the size.

The above substitutions reduce the first half of (10) to

$$\text{K. Exp.} - \left\{ \frac{\sum_1^n t_r^2 + n(b-\beta)^2 s_2^2 + n\bar{z}^2}{2\sigma_1^2(1-\rho^2)} \right\} \cdot dz_1\, dz_2 \ldots dz_n \quad (12)$$

$$\sum_1^n t_r^2 = nS_1^2(1-r^2) = V_{1.2}$$

The method used in the case of the normal population will enable us to show that (12) reduces to the form

$$\text{K}_1\text{. Exp.} - n\left\{ \frac{\sum_1^n t^2_r + n(b-\beta)^2 s_2^2 + n\bar{z}^2}{2\sigma_1^2(1-\rho^2)} \right\} \cdot V_{1.2}^{\frac{n-4}{2}}\, dV_{1.2}\, db\, d\bar{z}. \quad (13)$$

The distribution of $V_{1.2}$, b and $\bar{z}$ are independent of one another.

The product of (11) and (13) gives the joint distribution of $V_{1.2}$, b, s_2^2, $\bar{z}$ and $\bar{y}$. From this, the joint distribution of s_1^2, s_2^2 and r can be easily obtained by omiting the distribution portions for $\bar{z}$ and $\bar{y}$ and finding the Jacobian of

$$\left.\begin{aligned} f_1 &= s_1^2(1-r^2) - V_{1.2} \\ f_2 &= rs_1 - bs_2 \\ f_3 &= s_2^2 - V_2 \end{aligned}\right\} J = 2s_1^2$$

Substituting this, the distribution of s_1^2, s_2^2 and r reduces to

$$\text{L. Exp.} - \frac{n}{1-\rho^2}\left\{ \frac{s_1^2}{\sigma_1^2} - \frac{2\rho r s_1 s_2}{\sigma_1\sigma_2} + \frac{s_2^2}{\sigma_2^2} \right\} \cdot s_1^{n-2}\, s_2^{n-2}\, (1-r^2)^{\frac{n-4}{2}}\, ds_1\, ds_2^2\, dr \quad (14)$$

By taking the distributions of $V_{1.2}$, $\bar{z}$ and $\bar{y}$ it is possible to obtain the distribution of generalized Student's Z.

$$\frac{\bar{z}^2}{\sigma_1^2(1-\rho^2)} + \frac{(\bar{y}-m_2)^2}{\sigma_2^2} = \frac{1}{1-\rho^2}\left\{ \frac{(\bar{x}-m_1)^2}{\sigma_1^2} - \frac{2\rho(\bar{x}-m_1)(\bar{y}-m_2)}{\sigma_1\sigma_2} + \frac{(\bar{y}-m_2)^2}{\sigma_2^2} \right\}$$

The joint distribution of $V_{1.2}$, $\bar{z}$ and $\bar{y}$ is

$$\text{M. Exp.} - \left\{ \frac{(nV_{1.2} + n\bar{z}^2)}{2\sigma_1^2(1-\rho^2)} + \frac{n(\bar{y}-m_2)^2}{2\sigma_2^2} \right\} \cdot V_{1.2}^{\frac{n-4}{2}}\, dV_{1.2}\, d\bar{z}\, d\bar{y} \quad (14)$$

Assuming $\frac{\bar{z}^2}{\sigma_1^2(1-\rho^2)} + \frac{(\bar{y}-m_2)^2}{\sigma_2^2} = \frac{T^2}{\sigma_1^2(1-\rho^2)}$, and using Dirichlet's integral,

(14) reduces to $\text{M}_1\text{. Exp.} - \frac{n(V_{1.2}+T^2)}{2\sigma_1^2(1-\rho^2)} \cdot V_{1.2}^{\frac{n-4}{2}}\, T\, dV_{1.2}\, dT \quad (15)$

Putting $Z^2 = \frac{T^2}{V_{1.2}}$, (15) becomes

$$M_1. \quad \text{Exp.} - \frac{nV_{1.2}}{2\sigma_1^2(1-\rho^2)} (1+Z^2) \cdot V_{1.2}^{\frac{n-2}{2}} Z dV_{1.2}\, dZ \qquad (16)$$

Integrating from 0 to ∞ for $V_{1.2}$, (16) becomes

$$M_2. \quad \frac{Z\, dZ}{(1+Z^2)^{\frac{n}{2}}}$$ which ultimately is the same as that obtained by Hotelling.

It can be seen that

$$Z^2 = \frac{\sigma_1^2}{s_1^2(1-r^2)} \left\{ \frac{(\bar{x}-m_1)^2}{\sigma_1^2} - \frac{2\rho(\bar{x}-m_1)(\bar{y}-m_2)}{\sigma_1\sigma_2} + \frac{(\bar{y}-m_2)^2}{\sigma_2^2} \right\}$$

The real values of σ_1^2, σ_2^2 and ρ are not known; hence in the above test σ_1, σ_2 and ρ are to be estimated from the sample.

From the above discussion it appears that the simplest way of testing the difference between two bivariate samples is to test the significance of the difference between the following by the t test.

1. The difference between the regression coefficients.

2. The difference between the values of $(\bar{x}-m_1) - \frac{\rho\sigma_1}{\sigma_2}(\bar{y}-m_2)$ on the basis of the combined variance $s_1^2(1-r^2)$; for $\rho\frac{\sigma_1}{\sigma_2}$ the common value for the two samples together can be used.

3. The difference between $(\bar{y}-m_2)$ by using the variance s_2^2.

The assumptions made above will certainly affect the validity of the tests to a certain extent. But as the assumptions are in regard to the ratio $\frac{\sigma_1}{\sigma_2}$, $\frac{\sigma_1^2}{\sigma_2^2}$ and ρ, they are not likely to affect the tests to any appreciable extent.

4. Multivariate Normal Population

Following Yule and Kendall, the distribution in the case of the multivariate population of n-variates, $x_1, x_2, \ldots . x_n$ is

$$\frac{1}{(2\pi)^{\frac{n}{2}}\sqrt{\omega}} \text{Exp.} - \frac{\phi^2}{2},\ dx_1\, dx_2 \ldots . dx_n,$$

$$\phi^2 = \frac{1}{\omega}\{\omega_{11}x_1^2 + \omega_{22}x_2^2 + \ldots . 2\omega_{12}x_1x_2 + \ldots . 2\omega_{n,n-1}, x_n x_{n-1}\}$$

$$\omega=\sigma_1^2\sigma_2^2\ldots\sigma_n^2\begin{vmatrix}1 & \rho_{12} & \ldots & \rho_{1n}\\ \rho_{21} & 1 & \ldots & \rho_{2n}\\ \ldots & \ldots & \ldots & \ldots\\ \ldots & \ldots & \ldots & \ldots\\ \rho_{n1} & \rho_{n2} & \ldots & 1\end{vmatrix}$$

$\phi^2=\dfrac{x_1^2}{\sigma_1^2}+\dfrac{x^2_{2.1}}{\sigma^2_{2.1}}+\ldots\dfrac{x^2_{n.12..n-1}}{\sigma^2_{n.12..n-1}}$, $x_1, x_2, \ldots x_n$ are the deviations from the respective means and $x_{2.1}=x_2-\beta_{21}x_1$; $x_{3.12}=x_3-\beta_{32.1}x_2-\beta_{31.2}x_1$; and so on. β_{21}, $\beta_{32.1}$, etc., are the partial regression coefficients.

By proceeding exactly on the same lines as in the case of two variables, the joint distribution can be easily established, for $x_1, x_{2.1}, \ldots x_{n.123\ldots n}$ are all independently distributed.

We will now show how the distribution of the generalized Z can be obtained.

$$\frac{\bar{x}_1^2}{\sigma_1^2}+\frac{\bar{x}_{2.1}^2}{\sigma_{2.1}^2}+\ldots\frac{\bar{x}^2_{n.12..n-1}}{\sigma^2_{n.12..n-1}}=\frac{1}{\omega}\left\{\omega_{11}\bar{x}_1^2+\omega_{22}\bar{x}_2^2+\ldots 2\omega_{nn-1}\bar{x}_n\bar{x}_{n-1}\right\}=\text{Y} \quad (16)$$

$\sigma^1_{n.12..n-1}=\dfrac{\omega}{\omega_{nn}}\sigma_n^2$. The distribution of $\text{V}_{n.12..n-1}=s^2_{n.12,n-1}$ is

$$\text{M. Exp.}-\frac{\text{N V}_{n.12..n-1}}{2\sigma^2_{n.12..n-1}}\cdot \text{V}_{n.12\ldots n-1}^{\frac{\text{N}-n-2}{2}}\,d\text{V}_{n.12..n-1} \quad (17)$$

By using Dirichlet's integral it can be shown that the distribution of Y is

$$\text{P. Exp.}-\frac{\text{NY}}{2}\cdot\text{Y}^{\frac{n}{2}-1}\,d\text{Y} \quad (18)$$

Taking the joint distributions of (17) and (18) we get

$$\text{R. Exp.}-\frac{\text{N}}{2}\left\{\frac{\text{V}_{n.12..n-1}}{\sigma^2_{n.12..n-1}}+\text{Y}\right\}\cdot\text{V}_{n.12..n-1}^{\frac{\text{N}-n-2}{2}}\,\text{Y}^{\frac{n}{2}-1}\,d\text{V}_{n.12..n-1}\,d\text{Y} \quad (19)$$

Putting $\dfrac{\text{Y}\sigma^2_{n.12\ldots n-1}}{\text{V}_{n.12\ldots n-1}}=\text{Z}^2$, (19) becomes

$$\text{R}_1\text{. Exp.}-\frac{\text{N V}_{n.12..n-1}}{2}(1+\text{Z}^2)\cdot\text{V}_{n.12\ldots n-1}^{\frac{\text{N}-2}{2}}\,\text{Z}^{n-1}\,d\text{V}_{n.12..n-1}\,d\text{Z}. \quad (20)$$

Integrating for $\text{V}_{n.12..n-1}$ from 0 to ∞, (20) reduces to

$$\text{Q}\,\frac{\text{Z}^{n-1}}{(1+\text{Z}^2)^{\frac{\text{N}}{2}}}\,d\text{Z} \quad (21)$$

N represents the size of the sample.

It can be seen that $Z^2 = \dfrac{\sigma^2_{n.12..n-1}}{\omega \cdot s^2_{n.12..n-1}} \left\{\omega_{11}\bar{x}_1^2 + \omega_{22}\bar{x}_2^2 + ..2\omega_{nn-1}\bar{x}_n\bar{x}_{n-1}\right\}$

$$= \frac{\omega'_{nn}\,\sigma_n^2}{\omega_{nn}\,s_n^2}\frac{1}{\omega'}\left\{\omega_{11}\,\bar{x}_1^2 + \omega_{22}\,\bar{x}_2^2 + \ldots .2\omega_{nn-1}\,\bar{x}_n\,\bar{x}_{n-1}\right\} \tag{22}$$

If in (22) we substitute the sample values for σ_1, σ_2, etc.,

$Z^2 = \dfrac{1}{\omega'}\left\{\omega_{11}'\bar{x}_1^2 + \omega_{22}'\bar{x}_2^2 + \ \ldots\ \ 2\omega'_{nn-1}\,\bar{x}_n\bar{x}_{n-1}\right\}$, ω', ω_{11}', etc. are with reference to the sample.

As in the case of the bivariate samples, the simplest way of testing the difference between two samples is to test the significance of the following differences by the t test.

(1) The difference between the various regression coefficients.

(2) The difference between the values of x_1, $x_{2.1}$, $x_{3.12}$, $\ldots x_{n.12\ldots n-1}$; by using the common value of the regression coefficient. The remarks in regard to assumptions made in connection with the bivariate population is true in this case also.

If there are more than two samples, the test can be done by doing the analysis of variance for b's, x_1's, $x_{2.1}$'s .... $x_{n.12\ldots n-1}$'s.

5. Distribution of Means of Samples from Type III Population

The expression corresponding to (4) is

$$\frac{1}{(\Gamma p)^n}\,\text{Exp.} - \overset{n}{\Sigma}, x_r\,(x_1 x_2 \ldots . x_n)^{p-1}\,dx_1 dx_2 \ldots . dx_n \tag{23}$$

Put $\dfrac{x_1 + x_2 \ldots . x_n}{n} = \bar{x}$ and let x_1, $x_2 \ldots . x_n$ take all positive values satisfying the above condition. From (1) it can be easily seen that the chance of $\bar{x}$ lying between $\bar{x}$ and $x + d\bar{x}$ is.

$$\frac{(\Gamma p)^n n^{np}}{(\Gamma p)^n\,\Gamma np} \cdot \text{Exp.} - n\bar{x} \cdot \bar{x}^{np-1}\,d\bar{x}$$

$$= \frac{n}{\Gamma np} \cdot \text{Exp.} - n\bar{x} \cdot (n\bar{x}^{n|0-1})\,d\bar{x} \tag{24}$$

6. Distribution of Means of Samples from Rectangular Populations

(*a*) *Samples of size two.*—The chance of two observations lying between x_1, $x_1 + dx_1$; and x_2, $x_2 + dx_2$ is

$$dx_1\,dx_2 \tag{25}$$

If $\frac{x_1 + x_2}{2} = \bar{x}$, the probability of $\bar{x}$ lying between $\bar{x}$ and $\bar{x} + d\bar{x}$ is

$$4\bar{x}d\bar{x} \tag{26}$$

Since x_1 and x_2 can take values ranging between 0 and 1 only, it is evident that (26) will hold good only so long as $\bar{x} \leqslant \frac{1}{2}$. When $\bar{x}$ is greater than half, one of the observed values, say x_1, can be greater than one which is not included in the original distribution, *i.e.*, the range of the original distribution is only from 0 to 1 and this condition is satisfied so long as $\bar{x} \leqslant \frac{1}{2}$ and therefore $4\bar{x}d\bar{x}$ represents the distribution of the mean only up to $x \leqslant \frac{1}{2}$.

Now consider the values of x_1 and x_2 when $\bar{x} > \frac{1}{2}$, x_1 can lie anywhere between 0 and 1 and x_2 between 1 and 2. Evidently the value of x_2 does not come within our population values and therefore that portion of x contributed by the above values of x_1 and x_2 is to be subtracted from $4\bar{x}\ d\bar{x}$. Similarly taking the alternative case of x_2 lying between 0 and 1 and x_1 between 1 and 2 we find that this portion also is to be subtracted from $4\bar{x}\ d\bar{x}$.

It follows from the above discussion that the distribution of the mean in the region $x, \frac{1}{2}$ is given by

$$\begin{aligned} & 4\bar{x}\ d\bar{x} - {}_2c_1 \left[\frac{d}{d\bar{x}} \int_1^{2\bar{x}} \int_0^{2\bar{x}-x_1} dx_1\ dx_2\right] d\bar{x} \\ & = 4\bar{x}\ d\bar{x} - 8\left(\bar{x} - \tfrac{1}{2}\right) d\bar{x} \\ & = 4\left(1 - \bar{x}\right) d\bar{x} \end{aligned} \tag{27}$$

(*b*) *Samples of size n.*—By proceeding on the same lines as in the case of two samples it can be shown that the distribution of the mean of samples of size n consists of different arcs over different sections 0 to $\frac{1}{n}$; $\frac{1}{n}$ to $\frac{2}{n}$;..... $\frac{n-1}{n}$ to 1 and the equation over the region $\frac{r}{n}$ to $\frac{r+1}{n}$ is

$$\frac{n^n}{\Gamma n}\left\{\sum_0^r (-1)^r\ {}_nc_r \left(\bar{x} - \frac{r}{n}\right)^{n-1}\right\} d\bar{x}. \tag{28}$$

It may be mentioned that the above expression is identical with those obtained by Hall and Irwin by the use of geometric and characteristic function methods.

The same method enables us to obtain the distribution of the means of samples in the case of Types II, $K_1 \{x(1-x)\}^{p-1}\ dx$ and Type I curves, $k_2\, x^{p-1}\ (1-x)^{q-1}\ dx$.

SUMMARY

It has been shown how the well-known Dirichlet's multiple integral can be utilised to derive the distribution of (1) the means and the standard

deviations of samples from normal and multinormal curves; (2) the means of samples from Types I and III curves.

References

Fisher, R. A. .. *Biometrika*, 1915, **10**, 507.

——— .. *Metron*, 1926, **5**, 90.

Hall, P. .. *Biometrika*, 1927, **19**, 240.

Hotelling .. *Ann. Math. Stat.*, 1932, **2**, 360.

Irwin, J. O. .. *Biometrika*, 1927, **19**, 225.

——— .. *Metron*, 1930, **8**, 51.

Kosambi, D. D. .. *Curr. Sci.*, 1941, **10**, 191.

Kullback, S. .. *Ann. Math. Stat.*, 1934, **5**, 263.

——— .. *Ibid.*, 1935, **6**, 202.

Wishart, J., and Bartlett, M. S. *Proc. Camb. Phil. Soc.*, 1932, **28**, 455.

——— .. *Ibid.*, 1933, **29**, 260.

Udny Yule, G., and Kendall, M. G. *An Introduction to the Theory of Statistics*, 1940, Griffin & Co., Ltd., London.