ON SAMPLES FROM A MULTIVARIATE NORMAL POPULATION[1]

By Solomon Kullback

1. **Introduction.** In this paper we shall discuss the distribution of certain functions calculated for samples drawn from a multivariate normal population. The method of solution is based on the theory of characteristic functions and presents further application of that theory to the distribution problem of statistics.[2]

We shall have occasion to refer to the multivariate normal population whose distribution law is given by

$$F(x) \equiv \pi^{-n/2} \mid B_{pq} \mid^{1/2} e^{-B(x-m,\, x-m)} \qquad (p, q = 1, 2, \cdots, n) \tag{1.1}$$

where $B(x - m, x - m)$ is the real, positive definite quadratic form of the $x_p - m_p$ with matrix $\| B_{pq} \|$. Here m_p is the mean in the population of the *pth* variate and $B_{pq} = \Delta_{pq}/2\sigma_p\sigma_q\Delta$ where σ_p is the standard deviation in the population of the *pth* variate; Δ is the determinant of population correlations $\rho_{pq} = \rho_{qp}$; Δ_{pq} is the co-factor of ρ_{pq} in Δ; and $\mid B_{pq} \mid$ is the determinant of the matrix $\| B_{pq} \|$.

Since the integral of (1.1) over the entire field of variation of the variables is unity, we have (using abbreviated notation)

$$\int e^{-B(x-m,\, x-m)}\, dx = \pi^{n/2} \mid B_{pq} \mid^{-1/2} \tag{1.2}$$

Equation (1.2) will be true if $\| B_{pq} \|$ is complex, provided its real part is symmetric and positive definite.[3]

The distribution of sample means of samples from the population (1.1) is independent of the distribution of the system of sample variances and covariances and is given by[4]

$$F_1(\bar{x}) \equiv \pi^{-n/2} \mid A_{pq} \mid^{1/2} e^{-A(\bar{x}-m,\, \bar{x}-m)} \tag{1.3}$$

where $A(\bar{x} - m, \bar{x} - m)$ is the real, positive definite quadratic form of the $\bar{x}_p - m_p$ with matrix $\| A_{pq} \|$. Here $\bar{x}_p = (1/N) \sum_{\alpha=1}^{N} x_{p\alpha}$ is the sample mean of the *pth*

---

[1] Presented to the American Mathematical Society, February 23, 1935.

[2] For more complete reference to the theory of characteristic functions as applied to statistics see S. Kullback, *Annals of Mathematical Statistics*, Vol. 5 (1934), pp. 263–307.

[3] J. Wishart and M. S. Bartlett, *Proc. Cambridge Phil. Soc.*, Vol. 29 (1933), pp. 260 ff.

[4] J. Wishart, *Biometrika*, Vol. 20 A (1928), pp. 32–52.

J. Wishart and M. S. Bartlett, *loc. cit.*

variate, and $A_{pq} = NB_{pq}$, where B_{pq} has been defined for equation (1.1). The distribution law of the system of sample variances and covariances is given by[5]

$$(1.4) \qquad F_2(a) \equiv \frac{|A_{pq}|^{(N-1)/2}}{\pi^{n(n-1)/4} \prod_{r=1}^{n} \Gamma(N-r)/2} e^{-A(a)} |a_{pq}|^{(N-n-2)/2}$$

where $A(a) = \sum_{p,q=1}^{n} A_{pq} a_{pq}$ and $a_{pq} = a_{qp} = (1/N) \sum_{\alpha=1}^{N} (x_{p\alpha} - \bar{x}_p)(x_{q\alpha} - \bar{x}_q)$ with A_{pq} and $\bar{x}_p$ defined as for (1.3). Since the integral of (1.4) over the entire field of variation of the a_{pq} is unity, we have[6]

$$(1.5) \qquad \int e^{-A(a)} |a_{pq}|^{(N-n-2)/2} da = \pi^{n(n-1)/4} |A_{pq}|^{(1-N)/2} \prod_{r=1}^{n} \Gamma(N-r)/2$$

Equation (1.5) will also hold if the matrix $\|A_{pq}\|$ is complex, provided its real part is symmetric and positive definite.[7]

2. **Variance.** Consider a sample of N independent items from the normal population (1.1). Let

$$(2.1) \qquad v = \sum_{p,q=1}^{n} a_{pq}$$

where a_{pq} is defined as in (1.4). From the theory of characteristic functions and (1.5), we have that the characteristic function of the distribution law of v is given by[8]

$$(2.2) \qquad \varphi(t) = \int e^{it\Sigma a_{pq}} F_2(a)\, da = |A_{pq}|^{(N-1)/2} \cdot |A_{pq} - it|^{(1-N)/2}.$$

It may be readily shown that

$$(2.3) \qquad |A_{pq} - it| = |A_{pq}| - it \sum_{p,q=1}^{n} A^{pq}$$

where A^{pq} is the co-factor of A_{pq} in $|A_{pq}|$.

We thus have for the distribution law[8] of v

$$(2.4) \qquad P(v) = (A/c)^{(N-1)/2} \cdot \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itv} (A/c - it)^{(1-N)/2} dt$$

---

[5] J. Wishart, *loc. cit.*

[6] Cf. S. S. Wilks, *Biometrika*, Vol. 24 (1932), pp. 471–494.

[7] A. E. Ingham, *Proc. Cambridge Phil. Soc.*, Vol. 29 (1933), p. 271 ff. The considerations in this paper will still hold if the condition above is imposed.

[8] S. Kullback, *loc. cit.*, p. 272.

where $A = |A_{pq}|$, $c = \sum_{p,q=1}^{n} A^{pq}$ and $A/c > 0$ since $\| A_{pq} \|$ is positive definite. By using the fact that[9]

$$(2.5) \qquad \frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} e^{\mu z} z^{-k} \, dz = \begin{cases} \mu^{k-1}/\Gamma(k), & \mu > 0 \\ 0, & \mu \leqq 0 \end{cases}$$

where $k > 0$, $a > 0$, we have

$$(2.6) \qquad P(v) = \frac{(A/c)^{(N-1)/2}}{\Gamma (N-1)/2} v^{(N-3)/2} e^{-(A/c)v}$$

3. **Ratio of variances.** If v_1 and v_2 represent the statistic v (defined in (2.1)), obtained from independent samples of N_1 and N_2 items respectively, then it may be shown that the distribution law of $w = v_1/v_2$ is given by[10]

$$(3.1) \qquad P(w) = \frac{\Gamma(N_1 + N_2 - 2)/2}{\Gamma(N_1 - 1)/2 \, \Gamma(N_2 - 1)/2} w^{(N_1-3)/2} (1 + w)^{(2-N_1-N_2)/2}.$$

If we set $w = e^{2z} n_1/n_2$, where $n_1 = N_1 - 1$ and $n_2 = N_2 - 1$ we obtain for the distribution law of z[11]

$$(3.2) \qquad P(z) = 2 \frac{\Gamma(n_1 + n_2)/2}{\Gamma n_1/2 \; \Gamma n_2/2} n_1^{n_1/2} n_2^{n_2/2} e^{n_1 z} (n_2 + n_1 e^{2z})^{-(n_1+n_2)/2}.$$

4. **Student's distribution.** Consider a sample of N independent items from the normal population (1.1). Let

$$(4.1) \qquad \mu = \sum_{p,q=1}^{n} (\bar{x}_p - m_p)(\bar{x}_q - m_q)$$

where $\bar{x}_p$ and m_p are defined as in (1.3). The characteristic function of the simultaneous distribution function of μ, defined as in (4.1) and v defined as in (2.1) is given by

$$(4.2) \qquad \varphi(t_1, t_2) = \int \exp \left\{ it_1 \sum_{p,q=1}^{n} (\bar{x}_p - m_p)(\bar{x}_q - m_q) + it_2 \sum_{p,q=1}^{n} a_{pq} \right\} F_1(\bar{x}) F_2(a) \, d\bar{x} da$$

---

[9] Cf. A. E. Ingham, *loc. cit.*

J. Wishart and M. S. Bartlett, *Proc. Cambridge Phil. Soc.*, Vol. 28 (1932), p. 455 ff.

[10] S. Kullback, note accepted for publication soon in the *Annals of Math. Statistics*.

[11] Cf. R. A. Fisher, I. *Proc. International Math. Congress, Toronto* (1924), Vol. 2, pp. 805–813.

R. A. Fisher, II. *Statistical Methods for Research Workers*, 4th Edition (1932), Edinburgh: Oliver and Boyd, pp. 224–227.

where F_1 and F_2 are defined as in (1.3) and (1.4) respectively. From (1.2) and (1.5) we have that

$$\varphi(t_1, t_2) = (A/c)^{N/2}(A/c - it_1)^{-1/2}(A/c - it_2)^{(1-N)/2} \tag{4.3}$$

where A and c are defined as in (2.4). The simultaneous distribution of μ and v is given by

$$P(\mu, v) = (1/2\pi)^2 \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-it_1\mu - it_2 v}\varphi(t_1, t_2)\, dt_1 dt_2 \tag{4.4}$$

which evaluated by a procedure similar to that used for (2.4) yields

$$P(\mu, v) = \frac{(A/c)^{N/2}}{\Gamma(N-1)/2\,\Gamma 1/2}\, \mu^{-1/2} e^{-\mu A/c} v^{(N-3)/2} e^{-vA/c}. \tag{4.5}$$

From (4.5) we may readily obtain the distribution of $z = \mu^{1/2}/v^{1/2}$ to be[12]

$$D(z) = 2\,\frac{\Gamma N/2}{\Gamma(N-1)/2\,\Gamma 1/2}(1 + z^2)^{-N/2}, \qquad (0 \leqq z \leqq \infty). \tag{4.6}$$

5. *k* **samples.** Suppose we have k independent samples of $N_1, N_2, \cdots, N_k$ items respectively, drawn fron the normal population defined by (1.1). Let μ_r, $(r = 1, 2, \cdots, k)$ be the statistic μ, defined by (4.1), for each of the k samples respectively; let V_r, $(r = 1, 2, \cdots, k)$ be the statistic V, defined by (2.1), for each of the k samples respectively; let μ_0 and V_0 be the values of these statistics for the sample of $N = N_1 + N_2 + \cdots + N_k$ items obtained by pooling all the samples.

It may be readily verified that

$$\mu_0 = \sum_{r=1}^{k} \mu_r N_r^2/N^2 + 2\sum_{\alpha,\beta=1}^{k} \mu_\alpha^{1/2}\mu_\beta^{1/2} N_\alpha N_\beta/N^2 \qquad (\alpha \neq \beta) \tag{5.1}$$

$$N\mu_0 + NV_0 = \sum_{r=1}^{k}(N_r\mu_r + N_r V_r) \tag{5.2}$$

$$NV_0 = \sum_{r=1}^{k}(N_r V_r + M_r\mu_r) - 2\sum_{\alpha,\beta=1}^{k} \mu_\alpha^{1/2}\mu_\beta^{1/2} N_\alpha N_\beta/N \qquad (\alpha \neq \beta) \tag{5.3}$$

where $M_r = (NN_r - N_r^2)/N$.

In view of (2.6) and (4.5), it is evident that the simultaneous distribution law of μ_r, V_r, $(r = 1, 2, \cdots, k)$ is given by

$$P(\mu)\cdot Q(v) \equiv \prod_{r=1}^{k} P(\mu_r; N_r)\, Q(V_r; N_r) \tag{5.4}$$

---

[12] Cf. "Student," *Biometrika*, Vol. 6 (1908–09), pp. 1–25.
R. A. Fisher, *Metron*, Vol. 5 (1925), pp. 90–104.
P. R. Rider, *Annals of Mathematics*, 2nd S., Vol. 31 (1930), pp. 579–582.

where

$$P(\mu_r; N_r) \equiv \frac{N_r^{1/2}}{\Gamma 1/2} (B/D)^{1/2} \mu_r^{-1/2} e^{-N_r \mu_r B/D} \tag{5.5}$$

$$Q(V_r; N_r) \equiv \frac{N_r^{(N_r-1)/2}}{\Gamma(N_r - 1)/2} (B/D)^{(N_r-1)/2} V_r^{(N_r-3)/2} e^{-N_r V_r B/D} \tag{5.6}$$

and B is the determinant $| B_{pq} |$ defined in (1.1) and $D = \sum_{p,q=1}^{n} B^{pq}$ where B^{pq} is the co-factor of B_{pq} in $| B_{pq} |$.

Using (5.3) and (5.4), we find that the characteristic function of the simultaneous distribution law of $\varphi_r = V_r B/D$, $(r = 0, 1, \cdots, k)$ is given by

$$\varphi(t_0, t_1, \cdots, t_k) = \int e^{U(t_0)+V(t_0, t_1, \cdots, t_k)} P(u) \cdot Q(v) du dv \tag{5.7}$$

where

$$U(t_0) = (B\, it_0/D) \left\{ \sum_{r=1}^{k} \mu_r M_r/N - 2 \sum_{\alpha, \beta=1}^{k} \mu_\alpha^{1/2} \mu_\beta^{1/2} N_\alpha N_\beta/N^2 \right\}, \quad (\alpha \neq \beta)$$

and

$$V(t_0, t_1, \cdots, t_k) = (B/D) \left\{ \sum_{r=1}^{k} V_r(it_r + it_0\, N_r/N) \right\}$$

Let $\mu_r B/D = \zeta_r^2$ and $V_r B/D = \eta_r$, $(r = 1, 2, \cdots, k)$ and rewrite (5.7) as the product of $k + 1$ integrals

$$\varphi(t_0, t_1, \cdots, t_k) = I_0 I_1 \cdots I_k \tag{5.8}$$

where

$$I_0 = \frac{(N_1 N_2 \cdots N_k)^{1/2}}{\Gamma(1/2)^k} \int e^{-T(\zeta, \zeta)}\, d\zeta \tag{5.9}$$

with

$$T(\zeta, \zeta) = \sum_{r=1}^{k} \zeta_r^2 (N_r - it_0\, M_r/N) + 2\, it_0 \sum_{\alpha, \beta=1}^{k} \zeta_\alpha \zeta_\beta N_\alpha N_\beta/N^2, \quad (\alpha \neq \beta)$$

and

$$I_r = \frac{N_r^{(N_r-1)/2}}{\Gamma(N_r - 1)/2} \int_0^\infty \exp\{-\eta_r (N_r - it_0 N_r/N - it_r)\}\, \eta_r^{(N_r-3)/2}\, d\eta_r\, . \tag{5.10}$$

By employing (1.2) we find that

$$(5.11)\quad I_0 = (N_1N_2\cdots N_k)^{1/2}\begin{vmatrix} N_1 - it_0M_1/N & it_0\, N_1N_2/N^2 & \cdots & it_0\, N_1N_k/N^2 \\ it_0\, N_2N_1/N^2 & N_2 - it_0\, M_2/N & \cdots & it_0\, N_2N_k/N^2 \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ it_0\, N_kN_1/N^2 & it_0\, N_kN_2/N^2 & \cdots & N_k - it_0\, M_k/N \end{vmatrix}^{-1/2}$$

The determinant may be readily evaluated by removing the common factor N_r from the rth row (remembering the value of M_r as given in (5.3)) and applying the operations[13] (row 1 − row 2), (row 2 − row 3), $\cdots$, and then column k + column 1 + column 2 + $\cdots$ + column $k - 1$. We thus obtain

$$(5.12)\qquad I_0 = (1 - it_0/N)^{-(k-1)/2}$$

The integral in (5.10) is well-known and yields

$$(5.13)\qquad I_r = N_r^{(N_r-1)/2}\,(N_r - it_0\, N_r/N - it_r)^{-(N_r-1)/2}.$$

There thus results

$$(5.14)\quad \varphi(t_0, t_1, \cdots, t_k) = G(1 - it_0/N)^{-(k-1)/2}\prod_{\alpha=1}^{k}(N_\alpha - it_0\, N_\alpha/N - it_\alpha)^{-(N_\alpha-1)/2}$$

where $G = \prod_{\alpha=1}^{k} N_\alpha^{(N_\alpha-1)/2}$.

The simultaneous distribution law of φ_r, $(r = 0, 1, \cdots, k)$ is given by

$$(5.15)\quad P(\varphi_0, \varphi_1, \cdots, \varphi_k) = \frac{G}{(2\pi)^{k+1}} \int_{-\infty}^{\infty}\frac{e^{-it_0\varphi_0 - it_1\varphi_1 - \cdots - it_k\varphi_k}\, dt_0\, dt_1 \cdots dt_k}{(1 - it_0/N)^{(k-1)/2}\prod_{\alpha=1}^{k}(N_\alpha - it_0\, N_\alpha/N - it_\alpha)^{(N_\alpha-1)/2}}$$

Integrating successively with respect to $t_k, t_{k-1}, \cdots, t_1$ and applying (2.5) we have

$$(5.16)\quad P(\varphi_0, \varphi_1, \cdots, \varphi_k) = G \exp\left\{-\sum_{\alpha=1}^{k} N_\alpha\varphi_\alpha\right\}\prod_{\alpha=1}^{k}\frac{\varphi_\alpha^{(N_\alpha-3)/2}}{\Gamma(N_\alpha - 1)/2}\cdot\frac{1}{2\pi} \int_{-\infty}^{\infty}\frac{e^{-it_0\left(\varphi_0 - \frac{N_1}{N}\varphi_1 - \cdots - \frac{N_k}{N}\varphi_k\right)}}{(1 - it_0/N)^{(k-1)/2}}\, dt_0$$

[13] Cf. A. C. Aitken, *Quarterly Journal Math.*, Vol. 2 (1931), pp. 130–135.

and finally

$$P(\varphi_0, \varphi_1, \cdots, \varphi_k) = GN^{(k-1)/2}e^{-N\varphi_0} \tag{5.17}$$
$$\frac{(\varphi_0 - N_1\varphi_1/N - \cdots - N_k\varphi_k/N)^{(k-3)/2}}{\Gamma(k-1)/2} \prod_{\alpha=1}^{k} \frac{\varphi_\alpha^{(N_\alpha-3)/2}}{\Gamma(N_\alpha - 1/2)}.$$

If we apply to (5.17) the transformation

$$\begin{cases} \varphi_0 = \varphi_0 \\ \varphi_r = N\zeta_r\varphi_0/N_r \qquad (r = 1, 2, \cdots, k) \end{cases} \tag{5.18}$$

and integrate out φ_0, we obtain for the simultaneous distribution law of $\zeta_r = N_r\varphi_r/N\varphi_0 = N_rV_r/NV_0$

$$D(\zeta_1, \zeta_2, \cdots, \zeta_k) = \frac{\Gamma(N-1)/2}{\Gamma(k-1)/2}(1 - \zeta_1 - \zeta_2 - \cdots - \zeta_k)^{(k-3)/2} \tag{5.19}$$
$$\prod_{\alpha=1}^{k} \frac{\zeta_\alpha^{(N_\alpha-3)/2}}{\Gamma(N_\alpha - 1)/2}$$

where the limits of variation in (5.19) are[14]

$$\begin{cases} 0 \leqq \zeta_1 \leqq 1 \\ 0 \leqq \zeta_r \leqq 1 - \zeta_1 - \zeta_2 - \cdots - \zeta_{r-1}, \qquad (r = 2, 3, \cdots, k) \end{cases} \tag{5.20}$$

6. **Correlation ratio.** Let $\zeta = \log(1 - \zeta_1 - \zeta_2 - \cdots - \zeta_k)$ where the ζ_r, $(r = 1, 2, \cdots, k)$ are defined and distributed as in (5.19). The characteristic function of the distribution law of ζ is given by

$$\varphi(t) = \frac{\Gamma(N-1)/2}{\Gamma(k-1)/2}\int (1 - \zeta_1 - \zeta_2 - \cdots - \zeta_k)^{(k+2it-3)/2} \tag{6.1}$$
$$\prod_{\alpha=1}^{k} \frac{\zeta_\alpha^{(N_\alpha-3)/2}}{\Gamma(N_\alpha - 1)/2} d\zeta_\alpha$$

where the limits of variation are given by (5.20). The integral in (6.1) is readily evaluated as a Dirichlet integral,[15] and we obtain

$$\varphi(t) = \frac{\Gamma(N-1)/2\ \Gamma(k-1+2it)/2}{\Gamma(k-1)/2\ \Gamma(N-1+2it)/2}. \tag{6.2}$$

---

[14] Cf. J. Neyman and E. S. Pearson, I. *Bulletin de l'Académie Polonaise des Sciences et des Lettres, Série A, Sciences Mathématiques*, 1931, pp. 460–481.

[15] E. Goursat-E. R. Hedrick, *Mathematical Analysis*, Vol. I (1904) (Ginn and Co., N. Y.), p. 308.

The distribution law of ζ is given by

$$(6.3) \qquad P(\zeta) = \frac{\Gamma(N-1)/2}{\Gamma(k-1)/2} \cdot \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-it\zeta} \frac{\Gamma(k-1+2it)/2}{\Gamma(N-1+2it)/2} \, dt.$$

Now it may be shown that[16]

$$(6.4) \qquad \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-it\zeta} \frac{\Gamma(k-1+2it)/2}{\Gamma(N-1+2it)/2} \, dt = \frac{e^{\zeta(k-1)/2}(1-e^{\zeta})^{(N-k-2)/2}}{\Gamma(N-k)/2}$$

so that

$$(6.5) \qquad P(\zeta) = \frac{\Gamma(N-1)/2}{\Gamma(k-1)/2\,\Gamma(N-k)/2} e^{\zeta(k-1)/2}(1-e^{\zeta})^{(N-k-2)/2}.$$

If we set $e^{\zeta} = \eta^2$, then we obtain for the distribution[17] of η^2

$$(6.6) \qquad D(\eta^2) = \frac{\Gamma(N-1)/2}{\Gamma(k-1)/2\,\Gamma(N-k)/2} (\eta^2)^{(k-3)/2}(1-\eta^2)^{(N-k-2)/2}.$$

From its definition we have that

$$(6.7) \qquad \eta^2 = (NV_0 - N_1V_1 - \cdots - N_kV_k)/NV_0$$

which reduces to

$$(6.8) \qquad \eta^2 = (N_1W_1 + N_2W_2 + \cdots + N_kW_k)/NV_0$$

where $W_\alpha = \sum_{p,q=1}^{n} (\bar{x}_{p\alpha} - \bar{x}_{p0})(\bar{x}_{q\alpha} - \bar{x}_{q0})$ with $\bar{x}_{p\alpha}$ the sample mean of the *pth* variate in the *αth* sample and $\bar{x}_{p0}$ the sample mean of the *pth* variate in the sample formed by pooling all the samples.[18]

In a similar manner, we have that the distribution law of $\eta_\alpha^2 = \zeta_\alpha$, $(\alpha = 1, 2, \cdots, k)$ is given by

$$(6.9) \qquad D(\eta_\alpha^2) = \frac{\Gamma(N-1/2)}{\Gamma(N_\alpha-1)/2\,\Gamma(N-N_\alpha/2)} (\eta_\alpha^2)^{(N_\alpha-3)/2}(1-\eta_\alpha^2)^{(N-N_\alpha-2)/2}$$

It may be of interest to point out another derivation for the distribution of $h^2 = 1 - \eta^2$. Let

$$(6.10) \qquad \begin{cases} \theta = (B/D)(N_1V_1 + N_2V_2 + \cdots + N_kV_k) \\ \theta_0 = (B/D)NV_0 \end{cases}$$

---

[16] Whittaker and Watson, *Modern Analysis*, 2nd Ed., pp. 283, 333.

[17] Cf. R. A. Fisher, *loc. cit.*, I.

H. Hotelling, *Proc. National Academy of Sciences*, Vol. XI (1925), pp. 657–662.

[18] Cf. S. S. Wilks, *loc. cit.*, p. 482.

The characteristic function of the simultaneous distribution law of θ and θ_0 is immediately derivable from (5.14) by replacing t_0 by Nt_0 and t_r by $N_r t$ $(r = 1, 2, \cdots, k)$. There results

$$\varphi(t, t_0) = (1 - it_0)^{-(k-1)/2}(1 - it_0 - it)^{-(N-k)/2}. \tag{6.11}$$

By a procedure similar to that already used we find that the simultaneous distribution law of θ and θ_0 is given by

$$P(\theta, \theta_0) = \frac{\theta^{(N-k-2)/2}(\theta_0 - \theta)^{(k-3)/2}e^{-\theta_0}}{\Gamma(N-k)/2\,\Gamma(k-1)/2}. \tag{6.12}$$

By applying to (6.12) the transformation $\theta = \theta_0 h^2$, $\theta_0 = \theta_0$ and integrating out the value of θ_0, we find for the distribution law of h^2

$$D(h^2) = \frac{\Gamma(N-1)/2}{\Gamma(N-k)/2\,\Gamma(k-1)/2}(h^2)^{(N-k-2)/2}(1 - h^2)^{(k-3)/2}. \tag{6.13}$$

From (6.12) and (6.10) it may be shown that the following estimates of variance all have the same expected value[19]

$$\begin{cases} \dfrac{N_1V_1 + N_2V_2 + \cdots + N_kV_k}{N-k} \\[2ex] \dfrac{NV_0}{N-1} \\[2ex] \dfrac{N_1W_1 + N_2W_2 + \cdots + N_kW_k}{k-1} \end{cases} \tag{6.14}$$

7. Distribution of variances. Let

$$\begin{cases} \theta_r = N_rV_rB/D & (r = 1, 2, \cdots, k) \\ \theta_0 = NV_0B/D \\ \theta = (B/D)(N_1V_1 + N_2V_2 + \cdots + N_kV_k) \end{cases} \tag{7.1}$$

where the right members of (7.1) are defined as in section 5. It is evident that the characteristic function of the simultaneous distribution law of θ, θ_0, θ_r, $(r = 1, 2, \cdots, k-1)$ is derivable from (5.14) by replacing t_0 by Nt_0, t_r by $N_r(t_r + t)$, $(r = 1, 2, \cdots, k-1)$ and t_k by N_kt. Thus

$$\begin{aligned} \varphi(t, t_0, t_1, \cdots, t_{k-1}) &= (1 - it_0)^{-(k-1)/2} \\ &\quad (1 - it_0 - it)^{-(N_k-1)/2}\prod_{\alpha=1}^{k-1}(1 - it_0 - it_\alpha - it)^{(1-N_\alpha)/2}. \end{aligned} \tag{7.2}$$

[19] Cf. J. Neyman and E. S. Pearson, II. *Biometrika*, Vol. 20A (1928), pp. 273–274. S. Kullback, *Annals of Mathematical Statistics*, Vol. 6 (1935), pp. 76–77.

By proceeding as in section 5 we arrive at the result that the simultaneous distribution law of θ, θ_0, θ_r, $(r = 1, 2, \cdots, k - 1)$ is given by

$$P(\theta, \theta_0, \theta_r) = \frac{e^{-\theta_0} (\theta_0 - \theta)^{(k-3)/2} (\theta - \theta_1 - \theta_2 - \cdots - \theta_{k-1})^{(N_k-3)/2}}{\Gamma(k - 1)/2 \, \Gamma(N_k - 1)/2} \prod_{\alpha=1}^{k-1} \frac{\theta_\alpha^{(N_\alpha - 3)/2}}{\Gamma(N_\alpha - 1)/2} \tag{7.3}$$

where $\theta_0 \geqq \theta$, $\theta \geqq \theta_1 + \theta_2 + \cdots + \theta_{k-1}$.

By integrating out the variable θ_0 from (7.3) we have for the simultaneous distribution law of θ, θ_r, $(r = 1, 2, \cdots, k - 1)$

$$D(\theta, \theta_r) = \frac{e^{-\theta} (\theta - \theta_1 - \theta_2 - \cdots - \theta_{k-1})^{(N_k-3)/2}}{\Gamma(N_k - 1)/2} \prod_{\alpha=1}^{k-1} \frac{\theta_\alpha^{(N_\alpha-3)/2}}{\Gamma(N_\alpha - 1)/2} \tag{7.4}$$

A procedure similar to that used to derive (5.19) yields for the simultaneous distribution law of

$$\psi_r = \theta_r/\theta \qquad (r = 1, 2, \cdots, k - 1) \tag{7.5}$$

$$P(\psi_1, \psi_2, \cdots, \psi_{k-1}) = \frac{\Gamma(N - k)/2}{\Gamma(N_k - 1)/2} (1 - \psi_1 - \psi_2 - \cdots - \psi_{k-1})^{(N_k-3)/2} \prod_{\alpha=1}^{k-1} \frac{\psi_\alpha^{(N_\alpha-3)/2}}{\Gamma(N_\alpha - 1)/2} \tag{7.6}$$

where the limits of variation in (7.6) are[20]

$$\begin{cases} 0 \leqq \psi_1 \leqq 1 \\ 0 \leqq \psi_r \leqq 1 - \psi_1 - \psi_2 - \cdots - \psi_{r-1}, \qquad (r = 2, \cdots, k - 1). \end{cases} \tag{7.7}$$

In a manner similar to the derivation of (6.6) we find the distribution law of $h_\alpha^2 = \psi_\alpha$, $(\alpha = 1, 2, \cdots, k - 1)$, $h_k^2 = 1 - \psi_1 - \psi_2 - \cdots - \psi_{k-1}$ to be

$$D(h_\alpha^2) = \frac{\Gamma(N - k)/2}{\Gamma(N_\alpha - 1)/2 \, \Gamma(N - k - N_\alpha + 1)/2} (h_\alpha^2)^{(N_\alpha-3)/2} (1 - h_\alpha^2)^{(N-k-N_\alpha-1)/2}, \qquad (\alpha = 1, 2, \cdots, k). \tag{7.8}$$

From the distribution law in (7.3) we readily obtain that the characteristic function of the distribution law of $\gamma_\alpha^2 = \log (\theta_\alpha/(\theta_0 - \theta)$ is given by

$$\varphi(t) = \frac{\Gamma(N_\alpha - 1 + 2it)/2 \, \Gamma(k - 1 - 2it)/2}{\Gamma(N_\alpha - 1)/2 \, \Gamma(k - 1)/2} \qquad (\alpha = 1, 2, \cdots, k). \tag{7.9}$$

---

[20] Cf. J. Neyman and E. S. Pearson, *loc. cit.*, I.

We thus have that the distribution law of γ_α^2 is given by

$$(7.10)\quad P(\gamma_\alpha^2) = \frac{1}{\Gamma(N_\alpha - 1)/2\,\Gamma(k-1)/2} \cdot \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-it\gamma_\alpha^2}\,\Gamma(N_\alpha - 1 + 2\,it)/2\,\Gamma(k - 1 - 2\,it)/2\,dt.$$

The integral in (7.10) is known,[21] and there results

$$(7.11)\quad P(\gamma_\alpha^2) = \frac{\Gamma(N_\alpha + k - 2)/2}{\Gamma(N_\alpha - 1)/2\,\Gamma(k-1)/2}\, e^{\gamma_\alpha^2 (N_\alpha - 1)/2} \left(1 + e^{\gamma_\alpha^2}\right)^{-(N_\alpha + k - 2)/2}$$

If we set $e^{\gamma_\alpha^2} = \theta_\alpha/(\theta_0 - \theta) = \lambda_\alpha^2$ we have for the distribution of λ_α^2

$$(7.12)\quad D(\lambda_\alpha^2) = \frac{\Gamma(N_\alpha + k - 2)/2}{\Gamma(N_\alpha - 1)/2\,\Gamma(k-1)/2}\,(\lambda_\alpha^2)^{(N_\alpha - 3)/2}\,(1 + \lambda_\alpha^2)^{-(N_\alpha + k - 2)/2}$$

An extension of the procedure used to obtain (7.9) yields as the characteristic function of the simultaneous distribution of $\gamma_1^2, \gamma_2^2, \cdots, \gamma_k^2$

$$(7.13)\quad \varphi(t_1, t_2, \cdots, t_k) = \frac{\Gamma(k - 1 - 2it_1 - 2it_2 - \cdots - 2it_k)/2}{\Gamma(k-1)/2} \prod_{\alpha=1}^{k} \frac{\Gamma(N_\alpha - 1 + 2it_\alpha)/2}{\Gamma(N_\alpha - 1)/2}$$

Successive application of the method used to evaluate (7.10) yields as the simultaneous distribution law of the γ_α^2

$$(7.14)\quad P(\gamma_1^2, \gamma_2^2, \cdots, \gamma_k^2) = \frac{\Gamma(N-1)/2}{\Gamma(k-1)/2}\,(1 + e^{\gamma_1^2} + \cdots + e^{\gamma_k^2})^{-(N-1)/2} \prod_{\alpha=1}^{k} \frac{e^{\gamma_\alpha^2 (N_\alpha - 1)/2}}{\Gamma(N_\alpha - 1)/2}.$$

The simultaneous distribution of the λ_α^2 defined as in (7.12) is given by

$$(7.15)\quad D(\lambda_1^2, \lambda_2^2, \cdots, \lambda_k^2) = \frac{\Gamma(N-1)/2}{\Gamma(k-1)/2}\,(1 + \lambda_1^2 + \lambda_2^2 + \cdots + \lambda_k^2)^{-(N-1)/2} \prod_{\alpha=1}^{k} \frac{(\lambda_\alpha^2)^{(N_\alpha - 3)/2}}{\Gamma(N_\alpha - 1)/2}.$$

---

[21] Whittaker and Watson, *loc. cit.*, pp. 283, 383.

8. **Conclusion.** In this paper we have presented further instances of the applicability of the theory of characteristic functions to the distribution problem of statistics. In a subsequent paper the author hopes to illustrate the application of the results here developed to specific numerical problems.

WASHINGTON, D. C.