ASYMPTOTIC FORMULAE FOR THE DISTRIBUTION OF HOTELLING'S GENERALIZED T_0^2 STATISTIC. II

By Koichi Ito

Nanzan University, Nagoya, Japan

1. Summary. In this paper an asymptotic formula for the cumulative distribution function (c.d.f.) of Hotelling's generalized T_0^2 statistic is derived for general values of the number of dimensions in the non-null case. This result includes as a special case the previous result of the author for the null distribution of T_0^2 [5], and shows certain properties of T_0^2 which help determine the power of the test based on the statistic for moderately large samples. Both of the author's results provide an approximate complete analysis of the T_0^2 test, although the exact null and non-null distributions of T_0^2 are not available at present.

2. Introduction. Consider

$$H_0: \ \xi(p \times m) = 0(p \times m) \quad \text{against} \quad H: \ \xi \neq 0 \tag{2.1}$$

under the probability law

$$\{1/(2\pi)^{\frac{1}{2}p(m+n)} |\Sigma|^{\frac{1}{2}(m+n)}\} \exp [-\tfrac{1}{2} \operatorname{tr} \Sigma^{-1} \{(X_1 - \xi)(X_1' - \xi') + X_0 X_0'\}] \, dX_1 \, dX_0, \tag{2.2}$$

where X_1 and X_0 are $p \times m$ and $p \times n$ matrices, respectively, in which $p \leqq m$ or $>m$, but $p \leqq n$, and Σ is an unknown $p \times p$ symmetric, positive definite matrix. To test (2.1) Hotelling [3] proposed the statistic $T_0^2 \equiv m \operatorname{tr} S_1 S_0^{-1}$, where $S_1 \equiv X_1 X_1'/m$, $S_0 \equiv X_0 X_0'/n$, the prime denoting transposition of a matrix. The exact distribution of this statistic is known only when $p = 2$ and $\xi = 0$ ([3], p. 35), but for general values of p it is not available even in the null case. The author [5] gave asymptotic formulae for percentage points and for the c.d.f. of T_0^2 for general values of p when $\xi = 0$, and Siotani [9] independently obtained similar results. In the following sections we shall derive an asymptotic expansion for the c.d.f. of T_0^2 for general values of p when $\xi \neq 0$, which covers the author's previous result for $\xi = 0$ as a special case.

Let the non-zero characteristic roots of $\xi\xi'\Sigma^{-1}$ be denoted by $\lambda_1, \cdots, \lambda_q$, where q is the rank of $\xi\xi'$, $q \leqq p$, $q \leqq m$, and

$$\lambda \equiv \sum_{j=1}^{q} \lambda_j (\equiv \operatorname{tr} \xi\xi'\Sigma^{-1}). \tag{2.3}$$

It is easily seen that $\lambda \geqq 0$, with equality attained if and only if H_0 of (2.1) is true, and also that the distribution of T_0^2 when H_0 is not true involves no parameters other than p, m, n and the λ_j's. Let the non-null c.d.f. of T_0^2 be de-

Received July 25, 1959; revised February 23, 1960.

noted by

$$\Pr\{T_0^2 \leqq 2\theta \mid p, m, n; \lambda_1, \cdots, \lambda_q\} \equiv F_{p,m,n}(2\theta \mid \lambda_1, \cdots, \lambda_q). \tag{2.4}$$

Now, generalizing (4.2) of [5], we have

$$\begin{aligned} &\Pr\{m \operatorname{tr} S_1 S_0^{-1} \leqq 2\theta \mid p, m, n; \lambda_1, \cdots, \lambda_q\} \\ &\qquad = \int_R \Pr\{m \operatorname{tr} S_1 S_0^{-1} \leqq 2\theta \mid S_0; p, m, n; \lambda_1, \cdots, \lambda_q\} \Pr\{dS_0\}, \end{aligned} \tag{2.5}$$

where the first expression on the right denotes the conditional probability of the relation indicated for fixed values of the elements of S_0, and the second denotes the probability elements of S_0 which have a Wishart distribution with n degrees of freedom. The domain of integration R is over all possible values of the elements of S_0 such that S_0 is positive semi-definite. Expansion of the former about the elements of Σ in a Taylor series and integration term by term yields

$$F_{p,m,n}(2\theta \mid \lambda_1, \cdots, \lambda_q) = \Theta \Pr\{m \operatorname{tr} S_1 \Sigma^{-1} \leqq 2\theta \mid mp; \lambda\}, \tag{2.6}$$

where Θ is the asymptotic series with respect to n of the derivative operators given by (3.12) of [5], and the second expression on the right denotes the probability of the relation indicated, where $m \operatorname{tr} S_1\Sigma^{-1}$ is distributed as a non-central χ^2 with mp degrees of freedom and deviation parameter λ of (2.3), which is known to be a function of 2θ, mp and λ only, free of n [8]. (2.6) is one of what are conveniently known as "Studentization Formulae" [10]. When $\lambda = 0$, (2.6) reduces to

$$F_{p,m,n}(2\theta \mid 0, \cdots, 0) = \Theta \Pr\{m \operatorname{tr} S_1 \Sigma^{-1} \leqq 2\theta \mid mp; 0\}, \tag{2.7}$$

and the author made use of James' method ([6], [7]) to express the right hand side of (2.7) in an asymptotic series, (4.3) of [5].

If we pursue the same approach when $\lambda > 0$, however, the algebraic labor involved will be prohibitive for obtaining an actual expression of the asymptotic series on the right of (2.6). Such being the case, we shall first obtain an asymptotic series of the characteristic function of the non-null T_0^2 distribution for large values of n, which will then be inverted to derive an asymptotic expansion of the non-null c.d.f. of T_0^2.

3. Characteristic function of the non-null T_0^2 distribution. Let the characteristic function of the non-null T_0^2 distribution be denoted by

$$\phi_{p,m,n}(t \mid \lambda_1, \cdots, \lambda_q) \equiv \int_0^\infty \exp(itT_0^2)\, dF_{p,m,n}(T_0^2 \mid \lambda_1, \cdots, \lambda_q). \tag{3.1}$$

Hsu ([4], p. 232) gave an expression for the Laplace transform of a function of non-null T_0^2, by means of which (3.1) can be expressed in the form of a multiple integral as follows:

$$\begin{aligned}\phi_{p,m,n}(t \mid \lambda_1, \cdots, \lambda_q) \\ = (n\pi)^{-\frac{1}{2}mp} \prod_{r=1}^{p} \Gamma\{\tfrac{1}{2}(n+1+m-r)\}/\Gamma\{\tfrac{1}{2}(n+1-r)\} \\ \times \int_D \left| I + \frac{A}{n} \right|^{-\frac{1}{2}(m+n)} \exp\left\{ it\,(\mathrm{tr}\,A) + \sqrt{2it} \sum_{j=1}^{q} \sqrt{\lambda_j}\, x_{jj} \right\} \prod_{r=1}^{p} \prod_{s=1}^{m} dx_{rs},\end{aligned} \tag{3.2}$$

where I is the $p \times p$ unit matrix and $A(p \times p) \equiv X(p \times m)X'(m \times p)$, the (r, s)-element of X being x_{rs}. The domain of integration D is such that $-\infty < x_{rs} < +\infty$, $r = 1, \cdots, p$; $s = 1, \cdots, m$. Now making use of the generalized Stirling formula ([2], p. 130):

$$\Gamma(y) = (2\pi)^{\frac{1}{2}} y^{y-\frac{1}{2}} \{1 + (1/12y) + O(y^{-2})\} \exp(-y) \quad \text{for } y > 0, \tag{3.3}$$

it is easily seen that

$$\begin{aligned}n^{-\frac{1}{2}m} \Gamma\{\tfrac{1}{2}(n+1+m-r)\}/\Gamma\{\tfrac{1}{2}(n+1-r)\} \\ = 2^{-\frac{1}{2}m}\{1 + (\tfrac{1}{4}m^2 - \tfrac{1}{2}mr)/n + O(n^{-2})\}.\end{aligned} \tag{3.4}$$

Hence the constant in front of the integral in (3.2) becomes

$$\begin{aligned}(n\pi)^{-\frac{1}{2}mp} \prod_{r=1}^{p} \Gamma\{\tfrac{1}{2}(n+1+m-r)\}/\Gamma\{\tfrac{1}{2}(n+1-r)\} \\ = (2\pi)^{-\frac{1}{2}mp}\{1 + mp(m-p-1)/4n + O(n^{-2})\}.\end{aligned} \tag{3.5}$$

On the other hand,

$$\begin{aligned}\left| I + \frac{A}{n} \right|^{-\frac{1}{2}(m+n)} = [1 + \{(\mathrm{tr}_1\,A)^2 - 2\,\mathrm{tr}_2\,A - 2m\,\mathrm{tr}_1\,A\}/4n + 0(n^{-2})] \\ \times \exp(-\tfrac{1}{2}\,\mathrm{tr}_1\,A),\end{aligned} \tag{3.6}$$

where $\mathrm{tr}_k\, Q$ stands for the sum of all $k \times k$ minors formed by the intersection of any k rows of Q with k columns bearing the same number, Q being a square matrix. $\mathrm{tr}_1 Q$, which is the sum of all principal diagonal elements of Q, is $tr\ Q$. Hence (3.2) becomes

$$\begin{aligned}\phi_{p,m,n}(t \mid \lambda_1, \cdots, \lambda_q) = (2\pi)^{-\frac{1}{2}mp}\{1 + mp(m-p-1)/4n + O(n^{-2})\} \\ \times \int_D [1 + \{(\mathrm{tr}\,A)^2 - 2\,\mathrm{tr}_2\,A - 2m\,\mathrm{tr}\,A\}/4n + O(n^{-2})] \\ \times \exp\left\{ -\tfrac{1}{2}(1 - 2it)\,\mathrm{tr}\,A + \sqrt{2it} \sum_{j=1}^{q} \sqrt{\lambda_j}\, x_{jj} \right\} \prod_{r=1}^{p} \prod_{s=1}^{m} dx_{rs}.\end{aligned} \tag{3.7}$$

Now, after some transformations ([4], p. 227), the probability law of

$$X_1(p \times m)$$

given in (2.2) becomes

$$(3.8)\qquad (2\pi)^{-\frac{1}{2}mp} \exp(-\tfrac{1}{2}\lambda) \exp\left(-\tfrac{1}{2}\operatorname{tr} A + \sum_{j=1}^{q} \sqrt{\lambda_j}\, x_{jj}\right) \prod_{r=1}^{p} \prod_{s=1}^{m} dx_{rs},$$

where λ and λ_j's are given in (2.3), and x_{rs}'s and A are the same as used in (3.2). With respect to this probability law it is easily shown that

$$E(1) = 1,$$

$$E\{\exp g(x_{rs}\text{'s}, it, \lambda_1, \cdots, \lambda_q)\} = (1-2it)^{-\frac{1}{2}mp} \exp\{-\tfrac{1}{2}\lambda + [it\lambda/(1-2it)]\},$$

$$E\{(\operatorname{tr} A)\exp g(x_{rs}\text{'s}, it, \lambda_1, \cdots, \lambda_q)\} = (1-2it)^{-\frac{1}{2}(mp+2)} \cdot \{mp + [2it\lambda/(1-2it)]\} \exp\{-\tfrac{1}{2}\lambda + [it\lambda/(1-2it)]\},$$

$$E\{(\operatorname{tr} A)^2 \exp g(x_{rs}\text{'s}, it, \lambda_1, \cdots, \lambda_q)\} = (1-2it)^{-\frac{1}{2}(mp+4)}$$

$$(3.9)\qquad \cdot\left\{mp(mp+4) + (mp+2)\frac{4it\lambda}{1-2it} + \left(\frac{2it\lambda}{1-2it}\right)^2\right\} \times \exp\left(-\tfrac{1}{2}\lambda + \frac{it\lambda}{1-2it}\right),$$

$$E\{(\operatorname{tr}_2 A)\exp g(x_{rs}\text{'s}, it, \lambda_1, \cdots, \lambda_q)\} = (1-2it)^{-\frac{1}{2}(mp+4)}$$

$$\cdot\left\{\tfrac{1}{2}mp(m-1)(p-1) + (m-1)(p-1)\frac{2it\lambda}{1-2it} + \left(\frac{2it}{1-2it}\right)^2 \sum_{1=j<k}^{q} \lambda_j\lambda_k\right\} \times \exp\left(-\tfrac{1}{2}\lambda + \frac{it\lambda}{1-2it}\right),$$

where E stands for the expectation with respect to (3.8), and

$$g(x_{rs}\text{'s}, it, \lambda_1, \cdots, \lambda_q) = it(\operatorname{tr} A) + (\sqrt{2it} - 1)\sum_{j=1}^{q} \sqrt{\lambda_j}\, x_{jj}.$$

Substitution of (3.9) in (3.7) to carry out integration term by term yields

$$\phi_{p,m,n}(t \mid \lambda_1, \cdots, \lambda_q) = (1-2it)^{-\frac{1}{2}mp}\left[1 + \left\{mp(m-p-1) - 2m\left(mp + \frac{2it\lambda}{1-2it}\right)(1-2it)^{-1} + \left[mp(m+p+1) + (m+p+1)\frac{4it\lambda}{1-2it} + \left(\frac{2it}{1-2it}\right)^2 \sum_{j=1}^{q} \lambda_j^2\right](1-2it)^{-2}\right\}\Big/ 4n + O(n^{-2})\right] \exp\left(\frac{it\lambda}{1-2it}\right), \qquad (3.10)$$

which is an asymptotic expansion of the characteristic function of the non-null T_0^2 distribution for large values of n.

4. Asymptotic expansion of the non-null c.d.f. of T_0^2. Let $F(x)$ be the c.d.f. of a statistic, and $\phi(t)$ be its characteristic function. Then it is well known (e.g., see [10], p. 638) that if F and all its derivatives vanish at the extremes of the range of x and exist for all x in that range, then, by integration by parts,

$$(-it)^r\phi(t)$$

is the characteristic function of the rth derivative of $F(x)$, or $F^{(r)}(x)$, i.e.,

$$\int_{-\infty}^{\infty} \exp(itx)\, dF^{(r)}(x) = (-it)^r \phi(t). \tag{4.1}$$

Now if we denote the c.d.f. of a non-central χ^2 with 2ρ degrees of freedom and deviation parameter λ by $F_{2\rho}(2\theta \mid \lambda)$, its rth derivative by $F_{2\rho}^{(r)}(2\theta \mid \lambda)$, and their characteristic functions by $\phi_{2\rho}(t \mid \lambda)$ and $(-it)^r\phi_{2\rho}(t \mid \lambda)$, respectively, then these functions satisfy the above conditions, and hence (3.10) can be written as

$$\begin{aligned}\phi_{p,m,n}(t \mid \lambda_1, \cdots, \lambda_q) = \phi_{mp}(t \mid \lambda) &+ \{mp(m-p-1)\phi_{mp}(t \mid \lambda) \\ &- 2m^2p\phi_{mp+2}(t \mid \lambda) + mp(m+p+1)\phi_{mp+4}(t \mid \lambda) \\ &+ 4m\lambda(-it)\phi_{mp+4}(t \mid \lambda) - 4(m+p+1)\lambda(-it)\phi_{mp+6}(t \mid \lambda) \\ &+ 4\sum_{j=1}^{q} \lambda_j^2(-it)^2\phi_{mp+8}(t \mid \lambda)\}/4n + O(n^{-2}).\end{aligned} \tag{4.2}$$

Inversion of (4.2) gives

$$\begin{aligned}F_{p,m,n}(2\theta \mid \lambda_1, \cdots, \lambda_q) = F_{mp}(2\theta \mid \lambda) &+ \{mp(m-p-1)F_{mp}(2\theta \mid \lambda) \\ &- 2m^2pF_{mp+2}(2\theta \mid \lambda) + mp(m+p+1)F_{mp+4}(2\theta \mid \lambda) \\ &+ 4m\lambda F_{mp+4}^{(1)}(2\theta \mid \lambda) - 4(m+p+1)\lambda F_{mp+6}^{(1)}(2\theta \mid \lambda) \\ &+ 4\sum_{j=1}^{q} \lambda_j^2 F_{mp+8}^{(2)}(2\theta \mid \lambda)\}/4n + O(n^{-2}).\end{aligned} \tag{4.3}$$

Hence we have proved

THEOREM: *The non-null* c.d.f. *of the statistic* $T_0^2 \equiv m \operatorname{tr} S_1S_0^{-1}$, *where* S_1 *and* S_0 *are subject to the probability law* (2.2), *has an asymptotic expansion* (4.3) *for large values of* n.

When $\lambda = 0$, (4.3) becomes an asymptotic expansion of the null c.d.f. of T_0^2, i.e.,

$$\begin{aligned}F_{p,m,n}(2\theta \mid 0, \cdots, 0) &= F_{mp}(2\theta \mid 0) \\ &+ mp\{(m-p-1)F_{mp}(2\theta \mid 0) - 2mF_{mp+2}(2\theta \mid 0) \\ &+ (m+p+1)F_{mp+4}(2\theta \mid 0)\}/4n + O(n^{-2}),\end{aligned} \tag{4.4}$$

which is equivalent to (4.3) of [5]. In order to obtain several more terms in the above expansion, more systematic algebraic work would be necessary.

5. Remarks. To test (2.1) under (2.2) there have been proposed for use many test statistics which are functions of the characteristic roots of $mS_1S_0^{-1}$, and Hotelling's T_0^2 statistic is one of them, being the sum of the characteristic roots. Unfortunately there is no theory at present to tell how to choose among different test statistics when we want power against certain kinds of alternatives ([1], p. 224), because far too little is known about the relative merits of these test statistics. In this situation the formulae (4.3) and (4.4) together with (3.33) of [5], which gives an asymptotic formula for the percentage point of the null T_0^2 distribution, provide an approximate complete analysis of the test of significance based on T_0^2 for large values of n. (4.3) shows that the non-null c.d.f. of T_0^2 is related exclusively to $\sum_{j=1}^{q} \lambda_j$ and $\sum_{1=j<k}^{q} \lambda_j\lambda_k$ for large values of n, but it is conjectured that for small values of n the exact non-null c.d.f. of T_0^2 involves all the symmetric functions of the λ_j's. If results similar to (4.3) are obtained for other test statistics, they may throw some light to the question of choosing one to be used consistently in carrying out the actual tests of significance.

The author wishes to express his indebtedness to Prof. W. Kruskal and the referee for their comments and suggestions which led to this revision of the original work.

REFERENCES

[1] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*, John Wiley & Sons, New York, 1958.

[2] Harold Cramér, *Mathematical Methods of Statistics*, Princeton University Press, Princeton, N. J., 1951.

[3] Harold Hotelling, "A generalized T-test and measure of multivariate dispersion," *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, 1951, pp. 23–41.

[4] P. L. Hsu, "On generalized analysis of variance," *Biometrika*, Vol. 31 (1940), pp. 221–237.

[5] Koichi Ito, "Asymptotic formulae for the distribution of Hotelling's generalized T_0^2 statistic," *Ann. Math. Stat.*, Vol. 27 (1956), pp. 1091–1105.

[6] G. S. James, "The comparison of several groups of observations when the ratios of the population variances are unknown," *Biometrika*, Vol. 38 (1951), pp. 324–329.

[7] G. S. James, "Tests of linear hypotheses in univariate and multivariate analysis when the ratios of the population variances are unknown," *Biometrika*, Vol. 41 (1954), pp. 19–43.

[8] P. B. Patnaik, "The non-central χ^2- and F-distributions and their applications," *Biometrika*, Vol. 36 (1949), pp. 202–232.

[9] Minoru Siotani, "On the distribution of the Hotelling's T^2-statistics," *Ann. Inst. Stat. Math.*, Vol. 8 (1956), pp. 1–14.

[10] David L. Wallace, "Asymptotic approximations to distributions," *Ann. Math. Stat.*, Vol. 29 (1958), pp. 635–654.