A PROPERTY OF THE MULTIVARIATE t DISTRIBUTION

By Olive Jean Dunn

University of California at Los Angeles

0. Summary. The Student t distribution has the property that the distribution evaluated from $-u$ to $+u$ is an increasing function of ν, the degrees of freedom (this also applies to the distribution evaluated from $-\infty$ to $+u$). (An elementary proof of this property, due to M. Ray Mickey, is based on consideration of the ratio between the Student density functions for two consecutive values of the degrees of freedom, and makes use of the inequality [5], $(1/4\nu) - (1/360\nu^3) \leqq \log a_\nu \leqq -(1/4\nu) + (2/45\nu^3)$, where a_ν is the constant in the Student t density with ν degrees of freedom.) It is pointed out in this note that this monotonicity does not generalize to k dimensions in a usual multivariate t distribution; for k sufficiently large, the direction of the monotonicity becomes reversed.

1 Introduction. If $x = (x_1, \cdots, x_k)$ has a multivariate normal distribution with expected value $0 = (0, \cdots, 0)$, variances all equal to one, and correlation matrix $R = (\rho_{ij})$, and if νs^2 is independent of x and has a chi-square distribution with ν degrees of freedom, then $t_i = x_i/s$ are Student t variables with ν degrees of freedom, and $t = (t_1, \cdots, t_k)$ has a multivariate t distribution. See, for example, Dunnett and Sobel [1]. Its equation is

$$(1) \qquad f(t) = [\Gamma\{\tfrac{1}{2}(\nu + k)\}/(\nu\pi)^{k/2}\Gamma(\tfrac{1}{2}\nu)|R|^{\frac{1}{2}}][1 + \nu^{-1}tR^{-1}t']^{-\frac{1}{2}(\nu+k)}.$$

If $F(u)$ is defined by

$$(2) \qquad F(u) = P\{\textstyle\bigcap_{i=1}^{k} -u < t_i < u\},$$

then $F(u)$ equals the probability mass in the multivariate t distribution evaluated over a k-dimensional hypercube centered at the origin of half side u, and $F(u)$ is the distribution of the maximum of the absolute values of the k t variates. Similarly, if $G(u)$ is defined by

$$(3) \qquad G(u) = P\{\textstyle\bigcap_{i=1}^{k} -\infty < t_i < u\},$$

then $G(u)$ equals the probability mass evaluated from $-\infty$ to u in each direction and $G(u)$ is the distribution of the maximum of the k t variates. Tables of these functions (see [1], [2], [3], and [4]) can, whenever available, be used to obtain simultaneous confidence intervals with exact confidence level for k means when the correlation structure is known; $F(u)$ gives two-sided confidence intervals and $G(u)$ gives one-sided confidence intervals.

2. Theorem and proof.

Theorem. *For any given $u > 0$, and degrees of freedom $\nu_1 > \nu_2$, there exists an*

Received 17 August 1964; revised 29 October 1964.

integer K such that, for all $k \geqq K$,

$$F_{k,\nu_1}(u) < F_{k,\nu_2}(u), \quad \text{and} \quad G_{k,\nu_1}(u) < G_{k,\nu_2}(u).$$

Here $F_{k,\nu}(u)$ and $G_{k,\nu}(u)$ are $F(u)$ and $G(u)$ as defined above, with k the dimension and ν the degrees of freedom.

PROOF OF THE THEOREM.

$$(4) \qquad F_{k,\nu}(u) = P\,(\max |x_i|/s < u) = \int_0^\infty P\,(\max |x_i| < su) g_\nu(s)\, ds,$$

where $g_\nu(s)$ denotes the density function of s. Since νs^2 follows a chi-square distribution,

$$(5) \qquad g_\nu(s) = [\nu^{\nu/2}/2^{(\nu/2)-1}\Gamma(\nu/2)]s^{\nu-1} \exp\,(-\tfrac{1}{2}\,\nu s^2).$$

For $x_1, \cdots, x_k$ independent,

$$(6) \qquad \mathrm{P}\,(\max |x_i| < su) = [2\Phi(su) - 1]^k = h^k(s),$$

say, with Φ the univariate normal cdf.

For any fixed u and $\nu_1 > \nu_2$, the difference

$$(7) \qquad F_{k,\nu_1}(u) - F_{k,\nu_2}(u) = \int_0^\infty h^k(s)[g_{\nu_1}(s) - g_{\nu_2}(s)]\, ds.$$

For large values of s, $g_\nu(s)$ behaves like $\exp\,[-\frac{1}{2}\,\nu s^2]$, and so the quantity inside the brackets in (7) behaves like $\exp\,[-\frac{1}{2}\,\nu_1 s^2] - \exp\,[-\frac{1}{2}\,\nu_2 s^2]$ and so is negative. For s larger than some value c_1, which depends only on ν_1 and ν_2, then, the bracketed quantity is negative. Denoting the integrand in (7) by $i(s)$, the right hand integral of (7) may therefore be broken into four parts, as follows:

$$(8) \qquad F_{k,\nu_1}(u) - F_{k,\nu_2}(u) = \int_{s<c_1, i(s)>0} + \int_{s<c_1, i(s)<0} + \int_{c_1}^{c_2} + \int_{c_2}^\infty .$$

Here $c_2 > c_1$ and the quantities under the integral signs are the same as in (7). Over the range of integration in the last three of these four integrals, the integrand is negative. Omitting the two middle integrals, and remembering that $h(s)$ is an increasing function of s, we have (recalling the above definition of $i(s)$),

$$(9) \qquad F_{k,\nu_1}(u) - F_{k,\nu_2}(u) < h^k(c_1) \int_{s<c_1, i(s)>0} [g_{\nu_1}(s) - g_{\nu_2}(s)]\, ds + h^k(c_2) \int_{c_2}^\infty [g_{\nu_1}(s) - g_{\nu_2}(s)]\, ds$$
$$= h^k(c_1)\{\int_{s<c_1, i(s)>0} [g_{\nu_1}(s) - g_{\nu_2}(s)]ds + [h(c_2)/h(c_1)]^k \int_{c_2}^\infty [g_{\nu_1}(s) - g_{\nu_2}(s)]\, ds\}.$$

From (9), since the ratio $h(c_2)/h(c_1)$ is a constant larger than 1, and since the two integrals are simply two constants, the first positive and the second negative, we may find K such that for any $k > K$ the entire expression is negative.

The proof for $G_{k,\nu}(u)$ is identical and so is omitted.

3. Discussion. The theorem covers the case of all correlations equal to zero. When all correlations are equal to one, the distribution is the same as the univariate Student t distribution, so that for all dimensions, $F(u)$ and $G(u)$ are monotonically increasing functions of ν. Other correlation matrices may be considered in some sense to lie between these two extremes. It is not too surprising, then, in looking at tables for various correlation matrices, to find the reversal in direction occurring at higher dimensions than in the case of the identity matrix. In various unpublished tables of $F(u)$ the change is found to occur at a dimension where $F(u)$ is approximately .25 or .30.

The author finds it amusing to realize that for certain confidence levels, k simultaneous confidence intervals for k means would have expected lengths shorter if the variances are estimated from the data than the lengths of the intervals formed if the variances are known and the normal tables used. This seems to be of no practical importance, for the overall confidence level must be lower than is usually preferred—say .25 or below!

In spite of its lack of practical importance, it seems worthy of a note, if only to illustrate that one must be wary as the dimensions increase.

REFERENCES

[1] Dunnett, C. W. and Sobel, M. (1955). Approximations to the probability integral and certain percentage points of a multivariate analogue of Student's t-distribution. *Biometrika* **42** 258–260.

[2] Gupta, S. S. and Sobel, M. (1957). On a statistic which arises in selection and ranking problems. *Ann. Math. Statist.* **28** 957–967.

[3] Halperin, Max, Greenhouse, S. W., Cornfield, Jerome, and Zalokar, Julia. (1955). Tables of percentage points for the Studentized maximum absolute deviate in normal samples. *J. Amer. Statist. Assoc.* **50** 185–195.

[4] Pillai, K. C. S., and Ramachandran, K. V. (1954). Distribution of a Studentized order statistic. *Ann. Math. Statist.* **25** 569.

[5] Wallace, David L. (1959). Bounds on normal approximations to Student's and chi-square distributions. *Ann. Math. Statist.* **30** 1130.