THE SAMPLING DISTRIBUTIONS OF STATISTICS DERIVED FROM THE MULTIVARIATE t-DISTRIBUTION

By E. A. Cornish*

[*Manuscript received November* 30, 1954]

Summary

The sampling distributions of the more important statistical derivates from the multivariate t-distribution are established.

I. Introduction

A multivariate generalization of Student's t-distribution has been considered by Dunnett and Sobel (1954) in connexion with certain multiple decision problems concerned with the ranking, according to their mean values, of normal populations having a common unknown variance. These authors studied the probability integral of the bivariate population in detail, giving exact and asymptotic expressions, and tables for certain special cases. The particular applications to which they refer have been discussed by Bechhofer, Dunnett, and Sobel (1954). The same general distribution was derived, and its principal properties established independently by the author (Cornish 1954), when considering the pretreatment to be given to certain types of replicated experimental observations, before applying tests of normality. This distribution possesses properties which make it suitable as a basis for exact tests of significance in various problems, and Dunnett and Sobel have taken the first step towards its use in practice by providing tables of the probability integral. In this paper, we shall be concerned with sampling distributions of statistics derived from the multivariate t-distribution. The general sampling distribution of the means and sums of squares and products of the variates is first established, and from it, the sampling distributions of the more important statistical derivates are obtained.

II. General Distribution of Means and Sums of Squares and Products

Suppose $x_1, x_2, \ldots, x_p$ are distributed in a non-singular multivariate normal distribution having a null vector of means and variance-covariance matrix $\sigma^2\mathbf{R}$. In the class of cases to be considered, the symmetric correlation matrix $\mathbf{R}=[\rho_{ij}]$ is known, but the variance σ^2 is unknown. If s^2 is an estimate of σ^2, based on ν degrees of freedom and distributed independently of $x_1, x_2, \ldots, x_p$, then, as indicated above, it has been shown that the variates

$$t_i=x_i/s, \quad i=1, 2, \ldots, p, \quad \ldots\ldots\ldots\ldots\ (1)$$

* Division of Mathematical Statistics, C.S.I.R.O., Adelaide.

have the distribution

$$\frac{|\mathbf{R}|^{-\frac{1}{2}}\Gamma\frac{1}{2}(\nu+p)}{(\pi\nu)^{\frac{1}{2}p}\Gamma\frac{1}{2}\nu}(1+\mathbf{t}'\mathbf{R}^{-1}\mathbf{t}/\nu)^{-\frac{1}{2}(\nu+p)}d\mathbf{t}, \quad \ldots\ldots\ldots\ldots \quad (2)$$

where $\mathbf{t}'$ is the row vector $[t_1\, t_2 \ldots t_p]$.

If n samples are taken from the multivariate normal distribution yielding the variate values

$$x_{ik}, \qquad i=1, 2, \ldots, p, \quad k=1, 2, \ldots, n,$$

the probability of obtaining the corresponding sets of t-variates, defined by

$$t_{ik}=x_{ik}/s, \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (3)$$

is

$$\left\{\frac{|\mathbf{R}|^{-\frac{1}{2}}\Gamma\frac{1}{2}(\nu+p)}{(\pi\nu)^{\frac{1}{2}p}\Gamma\frac{1}{2}\nu}\right\}^n \prod_{k=1}^{n}(1+\mathbf{t}_k'\mathbf{R}^{-1}\mathbf{t}_k/\nu)^{-\frac{1}{2}(\nu+p)}d\mathbf{t}_k. \quad \ldots\ldots \quad (4)$$

The means and sums of squares and products of deviations of these t-variates are defined respectively by the relations

$$\left.\begin{aligned} \bar{t}_i &= \frac{1}{n}\sum_{k=1}^{n} t_{ik}, && i=1, 2, \ldots, p, \\ T_{ij} &= \sum_{k=1}^{n}(t_{ik}-\bar{t}_i)(t_{jk}-\bar{t}_j), && i, j=1, 2, \ldots, p. \end{aligned}\right\} \quad \ldots\ldots \quad (5)$$

Formally, the distribution function of the new variates given by the relations (5) can be derived from the distribution (4), but it is much more simply obtained from the known corresponding result for normally distributed quantities.

The sample means of the normal variates are defined by

$$\bar{x}_i=\frac{1}{n}\sum_{k=1}^{n} x_{ik}, \qquad i=1, 2, \ldots, p.$$

Denoting the matrix

$$[x_{ik}-\bar{x}_i], \qquad i=1, 2, \ldots, p, \quad k=1, 2, \ldots, n,$$

by $\mathbf{X}$, and writing

$$\mathbf{X}\mathbf{X}'=\mathbf{C}=[C_{ij}],$$

the distribution of the means and sums of squares and products (Wishart 1928) is

$$\frac{\left|\frac{\sigma^2}{n}\mathbf{R}\right|^{-\frac{1}{2}}}{(2\pi)^{\frac{1}{2}p}}\exp\left\{-\tfrac{1}{2}\bar{\mathbf{x}}'\left(\frac{\sigma^2}{n}\mathbf{R}\right)^{-1}\bar{\mathbf{x}}\right\}d\bar{\mathbf{x}}$$

$$\times\frac{|\sigma^2\mathbf{R}|^{-\frac{1}{2}(n-1)}|\mathbf{C}|^{\frac{1}{2}(n-p-2)}\exp\left\{-\frac{1}{2\sigma^2}\mathrm{Tr}\,(\mathbf{R}^{-1}\mathbf{C})\right\}}{2^{\frac{1}{2}p(n-1)}\pi^{\frac{1}{4}p(p-1)}\prod_{i=1}^{p}\Gamma\frac{1}{2}(n-i)}\prod_{i\leqslant j}dC_{ij}, \quad \ldots \quad (6)$$

where $\bar{\mathbf{x}}'$ is the row vector $[\bar{x}_1\, \bar{x}_2 \ldots \bar{x}_p]$.

From (3) and (5)

$$\left.\begin{aligned} \bar{t}_i &= \bar{x}_i/s, \\ T_{ij} &= C_{ij}/s^2, \end{aligned}\right\} \qquad (7)$$

and the jacobian of (7) is $s^{p(p+2)}$. Substitute in (6), using the relations (7), multiply by the distribution of s and integrate for s from 0 to ∞, and the distribution of the $\bar{t}_i$ and T_{ij} takes the form

$$\frac{\left|\frac{1}{n}\mathbf{R}\right|^{-\frac{1}{2}}\Gamma\frac{1}{2}(\nu+pn)}{(\pi\nu)^{\frac{1}{2}p}\Gamma\frac{1}{2}\nu} \; \frac{|\,\mathbf{R}\,|^{-\frac{1}{2}(n-1)}\,|\,\mathbf{T}\,|^{\frac{1}{2}(n-p-2)}}{\pi^{\frac{1}{4}p(p-1)}\nu^{\frac{1}{2}p(n-1)}\prod\limits_{i=1}^{p}\Gamma\frac{1}{2}(n-i)}$$

$$\times\frac{\mathrm{d}\bar{\mathbf{t}}\prod\limits_{i\leqslant j}\mathrm{d}T_{ij}}{\left\{1+\dfrac{\bar{\mathbf{t}}'\left(\frac{1}{n}\mathbf{R}\right)^{-1}\bar{\mathbf{t}}+\mathrm{Tr}\,(\mathbf{R}^{-1}\mathbf{T})}{\nu}\right\}^{\frac{1}{2}(\nu+pn)}}, \qquad (8)$$

where $\mathbf{T}$ is the matrix $[T_{ij}]$. The limiting form of the distribution (8), as $\nu\to\infty$, is the Wishart distribution.

III. Marginal Distribution of the $\bar{t}_i$

To obtain the marginal distribution of the means, integration of (8) with respect to the T_{ij} will be made following a method similar to that used by Cramer (1946, Section 29.5). Since $\mathbf{R}^{-1}$ is positive definite, it may be reduced by an orthogonal matrix $\mathbf{H}$, so that

$$\mathbf{H}\mathbf{R}^{-1}\mathbf{H}'=\mathbf{\Lambda},$$

where $\mathbf{\Lambda}$ is diagonal, its diagonal elements λ_i being the latent roots of $\mathbf{R}^{-1}$. The same transformation applied to the matrix $\mathbf{T}$ yields

$$\mathbf{H}\mathbf{T}\mathbf{H}'=\mathbf{Y},\text{ say}, \qquad (9)$$

transforming the $\frac{1}{2}p(p+1)$ variates in $\mathbf{T}$ to $\frac{1}{2}p(p+1)$ new variates in $\mathbf{Y}$. The relation (9) represents a linear transformation of the variables whose determinant (the jacobian required) is a power of the determinant $|\,\mathbf{H}\,|$ (MacDuffee 1943; James 1954). As $\mathbf{H}$ is orthogonal, the jacobian is unity, and, omitting the constant, (8) becomes

$$\frac{|\,\mathbf{Y}\,|^{\frac{1}{2}(n-p-2)}\mathrm{d}\bar{\mathbf{t}}\prod\limits_{i\leqslant j}\mathrm{d}Y_{ij}}{\left\{1+\dfrac{\bar{\mathbf{t}}'\left(\frac{1}{n}\mathbf{R}\right)^{-1}\bar{\mathbf{t}}+\mathrm{Tr}\,(\mathbf{\Lambda}\mathbf{Y})}{\nu}\right\}^{\frac{1}{2}(\nu+pn)}}. \qquad (10)$$

Next, for $i\neq j$, let

$$Y_{ij}=z_{ij}\sqrt{Y_{ii}Y_{jj}}, \qquad (11)$$

and denote the diagonal matrix, whose ith diagonal element is $\sqrt{Y_{ii}}$, by $\mathbf{D}$.

The jacobian of the transformation (**11**) is $(Y_{11}\,Y_{22}\,.\,.\,.\,Y_{pp})^{\frac{1}{2}(p-1)}$ and

$$|\,\mathbf{Y}\,|=|\,\mathbf{DZD}\,|=Y_{11}\,Y_{22}\,.\,.\,.\,Y_{pp}\,|\,\mathbf{Z}\,|,$$

where

$$\mathbf{Z}=\begin{bmatrix} 1 & z_{12} & . & . & . & z_{1p} \\ z_{21} & 1 & . & . & . & z_{2p} \\ . & . & . & . & . & . \\ z_{p1} & z_{p2} & . & . & . & 1 \end{bmatrix}.$$

The expression (**10**) thus changes to

$$\frac{(Y_{11}\,Y_{22}\,.\,.\,.\,Y_{pp})^{\frac{1}{2}(n-3)}\,|\,\mathbf{Z}\,|^{\frac{1}{2}(n-p-2)}\mathrm{d}\bar{\mathbf{t}}\prod_{i=1}^{p}\mathrm{d}Y_{ii}\prod_{i<j}\mathrm{d}z_{ij}}{\left\{1+\dfrac{\bar{\mathbf{t}}'\left(\dfrac{1}{n}\mathbf{R}\right)^{-1}\bar{\mathbf{t}}+\mathrm{Tr}\,(\boldsymbol{\Lambda}\mathbf{Y})}{\nu}\right\}^{\frac{1}{2}(\nu+pn)}}. \quad \ldots (\mathbf{12})$$

Integration of (**12**) over the domain where **Z** is positive definite yields the factor

$$\frac{\pi^{\frac{1}{4}p(p-1)}}{\{\Gamma\frac{1}{2}(n-1)\}^{p}}\prod_{i=1}^{p}\Gamma\tfrac{1}{2}(n-i),$$

and integration over the range 0 to ∞ for each of the Y_{ii} gives

$$\frac{\nu^{\frac{1}{2}p(n-1)}\{\Gamma\frac{1}{2}(n-1)\}^{p}\Gamma\frac{1}{2}(\nu+p)}{|\,\boldsymbol{\Lambda}\,|^{\frac{1}{2}(n-1)}\Gamma\frac{1}{2}(\nu+pn)},$$

leaving the marginal distribution of the $\bar{t}_i$ in the form

$$\frac{\left|\dfrac{1}{n}\mathbf{R}\right|^{-\frac{1}{2}}\Gamma\frac{1}{2}(\nu+p)}{(\pi\nu)^{\frac{1}{2}p}\Gamma\frac{1}{2}\nu}\cdot\frac{\mathrm{d}\bar{\mathbf{t}}}{\left\{1+\bar{\mathbf{t}}'\left(\dfrac{1}{n}\mathbf{R}\right)^{-1}\bar{\mathbf{t}}/\nu\right\}^{\frac{1}{2}(\nu+p)}}, \quad \ldots\ldots (\mathbf{13})$$

which is the multivariate t-distribution, characterized by the matrix $\{(1/n)\mathbf{R}\}^{-1}$. The variance-covariance matrix of the $\bar{t}_i$ is thus $(1/n)\{\nu/(\nu-2)\}\mathbf{R}$, and the marginal distribution of any variate $\bar{t}_i$ is

$$\frac{n^{\frac{1}{2}}\Gamma\frac{1}{2}(\nu+1)}{(\pi\nu)^{\frac{1}{2}}\Gamma\frac{1}{2}\nu}(1+n\bar{t}_i^2/\nu)^{-\frac{1}{2}(\nu+1)}\mathrm{d}\bar{t}_i$$

(Cornish loc. cit.). The variate $\bar{t}_i\sqrt{n}$ is then distributed in Student's distribution with ν degrees of freedom.

IV. Marginal Distribution of the T_{ij}

Since $\left(\frac{1}{n}\mathbf{R}\right)^{-1}$ is a positive definite symmetric matrix, the distribution (**8**) may be integrated easily over the range $-\infty$ to ∞ for each of the variables in $\bar{\mathbf{t}}$, yielding the numerical factor

$$\frac{\left|\frac{1}{n}\mathbf{R}\right|^{\frac{1}{2}}(\pi\nu)^{\frac{1}{2}p}\Gamma\frac{1}{2}[\nu+p(n-1)]}{\Gamma\frac{1}{2}(\nu+pn)},$$

and hence the marginal distribution of the T_{ij} is

$$\frac{|\mathbf{R}|^{-\frac{1}{2}(n-1)}\Gamma\frac{1}{2}[\nu+p(n-1)]}{\pi^{\frac{1}{4}p(p-1)}\nu^{\frac{1}{2}p(n-1)}\Gamma\frac{1}{2}\nu\prod_{i=1}^{p}\Gamma\frac{1}{2}(n-i)}\cdot\frac{|\mathbf{T}|^{\frac{1}{2}(n-p-2)}\prod_{i\leqslant j}dT_{ij}}{\{1+\mathrm{Tr}(\mathbf{R}^{-1}\mathbf{T})/\nu\}^{\frac{1}{2}[\nu+p(n-1)]}}. \quad ..(\mathbf{14})$$

The substitutions $(n-1)s_i^2=T_{ii}$ and $(n-1)w_{ij}=T_{ij}$ in (**14**) then give the simultaneous marginal distribution of the variances and covariances.

From (**14**) the marginal distribution of T_{ii} is

$$\frac{\Gamma\frac{1}{2}(\nu+n-1)}{\nu^{\frac{1}{2}(n-1)}\Gamma\frac{1}{2}\nu\Gamma\frac{1}{2}(n-1)}\cdot\frac{T_{ii}^{\frac{1}{2}(n-3)}dT_{ii}}{(1+T_{ii}/\nu)^{\frac{1}{2}(\nu+n-1)}}, \quad ....(\mathbf{15})$$

which is equivalent to the distribution of Fisher's z with $(n-1)$ and ν degrees of freedom. The χ^2 distribution, as the distribution of the sum of the squares of independent normal variates in standard measure, is thus replaced by the distribution of z as the sum of the squares of uncorrelated but dependent t-variates.

Writing $T_{ii}=(n-1)s_i^2$ in (**15**) will give the marginal distribution of the variance, and from this the distribution of the average conditional variance of any admissible order (partial variance in Bartlett's (1933) terminology) can be obtained by appropriately adjusting the degrees of freedom involving the sample number n.

V. Marginal Distributions of Correlation and Regression Coefficients

Since regression and correlation coefficients, both total and partial, and the multiple correlation coefficient are essentially ratios of quantities involving the t-variates, in which the standardizing variate, the estimated standard deviation, cancels out, their distributions are identical with those of their counterparts relating to normally distributed quantities. These distributions may, however, be obtained directly, and as an illustration we give the derivation of the distribution of the regression coefficient, from the marginal distribution (**14**) when $p=2$.

With this value of p, (**14**) reduces to

$$\frac{\Gamma\frac{1}{2}[\nu+2(n-1)]}{\pi^{\frac{1}{2}\nu(n-1)}(1-\rho^2)^{\frac{1}{2}(n-1)}\Gamma\frac{1}{2}\nu\Gamma\frac{1}{2}(n-1)\Gamma\frac{1}{2}(n-2)}$$
$$\times\frac{\begin{vmatrix} T_{11} & T_{12} \\ T_{12} & T_{22}\end{vmatrix}^{\frac{1}{2}(n-4)} \mathrm{d}T_{11}\mathrm{d}T_{12}\mathrm{d}T_{22}}{\left\{1+\frac{1}{\nu(1-\rho^2)}(T_{11}-2\rho T_{12}+T_{22})\right\}^{\frac{1}{2}[\nu+2(n-1)]}}. \quad \ldots\ldots (16)$$

Substitute

$$T_{12}=b_{21}T_{11}$$

in (**16**), and take the integral with respect to T_{11} and T_{22}. Hence

$$\int\int\frac{T_{11}{}^{\frac{1}{2}(n-2)}(T_{22}-b_{21}^2T_{11})^{\frac{1}{2}(n-4)}\mathrm{d}T_{11}\mathrm{d}T_{22}\mathrm{d}b_{21}}{\left[1+\frac{1}{\nu(1-\rho^2)}\{T_{22}+T_{11}(1-2\rho b_{21})\}\right]^{\frac{1}{2}[\nu+2(n-1)]}}$$

has to be taken over the region defined by the inequality

$$T_{22}-b_{21}^2T_{11}\geqslant 0.$$

Change the variable to $x=T_{22}-b_{21}^2T_{11}$, and integrate with respect to both x and T_{11} from 0 to ∞, and the distribution of b_{21} takes the form

$$\frac{(1-\rho^2)^{\frac{1}{2}(n-1)}\Gamma\frac{1}{2}n}{\pi^{\frac{1}{2}}\Gamma\frac{1}{2}(n-1)}(1-2\rho b_{21}+b_{21}^2)^{-\frac{1}{2}n}\mathrm{d}b_{21},$$

as first found for normally distributed variates by Pearson (1926) and Romanovsky (1926).

VI. Marginal Distribution of the Covariance

As the starting point for the derivation of the marginal distribution of the covariance, we use the result for normally distributed quantities (Pearson, Jeffery, and Elderton 1929; Wishart and Bartlett 1932). With the notation of Section II, and taking $p=2$, the distribution of C_{12} is

$$\frac{(1-\rho^2)^{\frac{1}{2}(n-3)}\exp\{\rho C_{12}/\sigma^2(1-\rho)^2\}}{\pi^{\frac{1}{2}}\sigma^2 2^{\frac{1}{2}(n-2)}\Gamma\frac{1}{2}(n-1)}\left\{\frac{|C_{12}|}{\sigma^2(1-\rho^2)}\right\}^{\frac{1}{2}(n-2)}K_{\frac{1}{2}(n-2)}\left\{\frac{|C_{12}|}{\sigma^2(1-\rho^2)}\right\}\mathrm{d}C_{12}, \quad \ldots\ldots (17)$$

where the vertical bars now designate the modulus, and $K_m(x)$ is the Bessel function of the second kind with imaginary argument.

For a fixed value of s^2, make the substitution $C_{12}=s^2T_{12}$ in (**17**), multiply by the distribution of s^2, and integrate with respect to s^2 from 0 to ∞, obtaining

$$\frac{\nu^{\frac{1}{2}\nu}(1-\rho^2)^{\frac{1}{2}(n-3)}\mathrm{d}T_{12}}{\pi^{\frac{1}{2}}2^{\frac{1}{2}(\nu+n-2)}\sigma^{\nu+2}\Gamma\frac{1}{2}\nu\Gamma\frac{1}{2}(n-1)}\int_0^\infty(s^2)^{\frac{1}{2}\nu}\exp\left\{-\frac{s^2}{2\sigma^2}\left(\nu-\frac{2\rho T_{12}}{1-\rho^2}\right)\right\}$$
$$\times\left\{\frac{s^2|T_{12}|}{\sigma^2(1-\rho^2)}\right\}^{\frac{1}{2}(n-2)}K_{\frac{1}{2}(n-2)}\left\{\frac{s^2|T_{12}|}{\sigma^2(1-\rho^2)}\right\}\mathrm{d}(s^2). \quad \ldots\ldots (18)$$

A second substitution

$$y=\frac{s^2\mid T_{12}\mid}{\sigma^2(1-\rho^2)}$$

reduces the integral in this expression to

$$\int_0^\infty y^{\frac{1}{2}(\nu+n-2)}\exp\left\{-\frac{1-\rho^2}{2\mid T_{12}\mid}\left(\nu-\frac{2\rho T_{12}}{1-\rho^2}\right)y\right\}K_{\frac{1}{2}(n-2)}(y)\mathrm{d}y,$$

and this may be evaluated using the result given by Watson (1922, p. 388 (**7**)), yielding

$$\frac{(\frac{1}{2}\pi)^{\frac{1}{2}}\Gamma\frac{1}{2}(\nu+2)\Gamma\frac{1}{2}[\nu+2(n-1)]P_{\frac{1}{2}(n-3)}^{-\frac{1}{2}(\nu+n-1)}(\cosh\alpha)}{(\sinh\alpha)^{\frac{1}{2}(\nu+n-1)}}, \quad \ldots . \quad (\mathbf{19})$$

where $\cosh\alpha=\nu(1-\rho^2)-2\rho T_{12}/(2\mid T_{12}\mid)$ and $P_\mu^\lambda(x)$ is the associated Legendre function of the first kind.

Multiplying (**19**) by the constant from (**18**) and substituting for the hyperbolic functions reduces the distribution of T_{12} to the form

$$\frac{\nu^{\frac{1}{2}(\nu+2)}\Gamma\frac{1}{2}[\nu+2(n-1)]}{2\Gamma\frac{1}{2}(n-1)}\cdot\frac{(1-\rho^2)^{\frac{1}{2}(\nu+n-1)}\mid T_{12}\mid^{\frac{1}{2}(n-3)}}{[\{\nu(1-\rho^2)-2\rho T_{12}\}^2-4\mid T_{12}\mid^2]^{\frac{1}{4}(\nu+n-1)}}$$
$$\times P_{\frac{1}{2}(n-3)}^{-\frac{1}{2}(\nu+n-1)}\left\{\frac{\nu(1-\rho^2)-2\rho T_{12}}{2\mid T_{12}\mid}\right\}\mathrm{d}T_{12}. \quad \ldots\ldots\ldots \quad (\mathbf{20})$$

The change of variable $T_{12}=(n-1)w_{12}$ will give the marginal distribution of the covariance, and from this the distribution of the partial covariance of any admissible order can be obtained by appropriately adjusting the degrees of freedom involving the sample number n.

VII. References

Bartlett, M. S. (1933).—*Proc. Roy. Soc. Edinb.* **53** (3): 260.

Bechhofer, R. E., Dunnett, C. W., and Sobel, M. (1954).—*Biometrika* **41**: 170.

Cornish, E. A. (1954).—*Aust. J. Phys.* **7**: 531.

Cramer, H. (1946).—"Mathematical Methods of Statistics." (Princeton Univ. Press.)

Dunnett, C. W., and Sobel, M. (1954).—*Biometrika* **41**: 153.

James, A. T. (1954).—*Ann. Math. Statist.* **25**: 40.

MacDuffee, C. C. (1943).—"Vectors and Matrices." (Math. Assoc. Amer.)

Pearson, K. (1926).—*Proc. Roy. Soc.* A **122**: 1.

Pearson, K., Jeffery, G. B., and Elderton, E. M. (1929).—*Biometrika* **21**: 164.

Romanovsky, V. (1926).—*Bull. Acad. Sci. U.R.S.S.* [6] **10**: 643.

Watson, G. N. (1922).—"Theory of Bessel Functions." (Cambridge Univ. Press.)

Wishart, J. (1928).—*Biometrika* **20A**: 32.

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