SOME CHARACTERIZATIONS OF THE MULTIVARIATE $t$ DISTRIBUTION

by

Pi-Erh Lin

FSU Statistics Report M200

February, 1971
Department of Statistics
Florida State University
Tallahassee, Florida
ABSTRACT

SOME CHARACTERIZATIONS OF THE MULTIVARIATE t DISTRIBUTION

by

Pi-Erh Lin

The Florida State University

A multivariate t vector \( \mathbf{x} \) is represented in two different forms. The
distributions of a linear function, a quadratic form, and a subvector of \( \mathbf{x} \) are
derived under both representations. Some characterizations of the multivariate
t distribution are established through theories of the multivariate normal
distribution.
SOME CHARACTERIZATIONS OF THE MULTIVARIATE $t$ DISTRIBUTION

Pi-Erh Lin

The Florida State University

1. Introduction and Summary. The multivariate $t$ distribution was first derived independently by Cornish [4] and Dunnett and Sobel [8]. It arises in multiple decision problems concerned with the selection and ranking of population means of several normal populations having a common unknown variance (see Bechhofer, Dunnett and Sobel [3]). It can be used in setting up simultaneous confidence bounds for the means of correlated normal variables, simultaneous confidence bounds for parameters in a linear model and for future observations from a multivariate normal distribution (see John [12]). The multivariate $t$ distribution also appears in the Bayesian multivariate analysis of variance and regression, treated by Tiao and Zellner [17], Geisser and Cornfield [10], Raiffa and Schlaifer [14], and Ando and Kaufmann [2], where the normal-Wishart distribution is considered to be the conjugate prior distribution (in the sense of Raiffa and Schlaifer) of the mean vector and covariance matrix of a multivariate normal distribution. However, the representation of the multivariate $t$ vector used in the Bayesian multivariate analysis is somewhat different from that employed by Cornish [4], and Dunnett and Sobel [8].

* During the work on this paper support was received from training grant number 5T1 GM-913 from National Institute of General Medical Sciences.
In this paper we consider the distribution of a multivariate $t$ vector which may be represented in two different forms. The distributions of some statistics derived from a multivariate $t$ vector are obtained. Some characterization properties of a multivariate $t$ vector are developed through theories of the multivariate normal distribution. More specifically, we present two different representations for a multivariate $t$ vector in Section 2. The distributions of a linear function, a quadratic form, and a subvector of a multivariate $t$ vector are obtained in Section 3. Finally, in Section 4, some characterizations of a multivariate $t$ vector are established.

2. Multivariate $t$ distribution. A $p$-variate random vector $\mathbf{X} = (X_1, \cdots, X_p)^T$ is said to have a (non-singular) multivariate $t$ distribution with mean vector $\mu = (\mu_1, \cdots, \mu_p)^T$ and covariance matrix $\Sigma = \nu^{-1} \mathbf{I}$, where $\nu > 2$, denoted by $T_{\nu}(\mu, \Sigma, p)$, if it has the probability density function (pdf) given by

$$f(x) = \frac{\Gamma\left(\frac{\nu + p}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right) \nu^{\frac{p}{2}} \left|\Sigma\right|^{\frac{1}{2}}} \left(1 + \frac{(x - \mu)^T \Sigma^{-1} (x - \mu)}{\nu}\right)^{-\frac{\nu + p}{2}}, \quad \nu > 2.$$  

(1)

It should be noted that when $p = 1$, $T_{\nu}(0,1,1) = t_\nu$ is the univariate Student's $t$ distribution with $\nu$ degrees of freedom.

For the purpose of establishing the characterizations and the distribution theories it is convenient to define a multivariate $t$ vector in the following forms:

**Representation A.** Let $\mathbf{X} \sim T_{\nu}(\mu, \Sigma, p)$. Then $\mathbf{X}$ may be written as

$$\mathbf{Y} = S^{-1/2} \mathbf{X},$$

(2)
where $\mathbf{Y} \sim \mathcal{N}(\mu, \Sigma)$ and $vs^2 \sim \chi^2_v$, independent of $\mathbf{Y}$. This implies that

$$X|s^2 = s^2 \sim \mathcal{N}(\mu, s^2 \Sigma),$$

where $vs^2 \sim \chi^2_v$. The Student's t random variable is then represented as $X|s^2 = s^2 \sim \mathcal{N}(0, s^{-2})$, where $vs^2 \sim \chi^2_v$.

**Representation B.** Let $X \sim T_v(\mu, I, p)$. Then $X$ may be written as

$$X = (V^T)^{-1} Y + \mu,$$

where $V^T$ is the symmetric square root of $\Sigma$, i.e., $V^T V = \Sigma \sim U(E^{-1}, v_{p+1})$ and $X \sim \mathcal{N}(\mu, V^{-1})$, independent of $\mu$. This implies that $X|V \sim \mathcal{N}(\mu, V^{-1})$, where $V \sim U(E^{-1}, v_{p+1})$. The Student's t random variable is then represented as $X|V \sim \mathcal{N}(0, V^{-1} q^2 / q^2)$ for any $p \times 1$ scalar vector $q \neq 0$, where $V \sim U(E^{-1}, v_{p+1})$.

Unless otherwise mentioned, it is assumed that $X$ and $\mu$ are $p \times 1$ vectors, $V$ and $\Sigma$ are $p \times p$ positive definite symmetric matrices.

3. **Distributions of statistics derived from multivariate t.** In this section we will derive the distributions of a linear function, a quadratic form, and a subvector of a multivariate $t$ vector defined by both Representations A and B.

**Theorem 1.** Let $X \sim T_v(\mu, I, p)$, then for any non-singular $p \times p$ scalar matrix $C$ and any $p \times 1$ scalar vector $d, C^T X \sim T_v(C \mu + d, C \Sigma C^T, p)$.

**Proof.** We will prove the theorem for a multivariate t vector $X$ defined by both representations.

(a) Under **Representation A**; Let $vs^2 \sim \chi^2_v$, then
(b) Under Representation B: Let \( V \sim \chi^2(\nu, v + p - 1) \), then

\[
\chi | V \sim \chi^2(y, v V^{-1}),
\]

\[
\Rightarrow \chi^2 + d | V \sim \chi^2(C \chi + d, v CV^{-1}C^*),
\]

where \( C^{-1}VC^{-1} \sim \nu([\text{CEC}]^{-1}, v + p - 1) \).

\[
\Rightarrow \chi^2 + d \sim \Gamma_y(C \chi + d, \text{CEC}^*; p).
\]

**Theorem 2.** Let \( \chi \sim \Gamma_y(y, v, p) \), defined by Representation A. Then

\[
p^{-1} \chi^2 \sim \chi^2_y,
\]

is noncentral \( F \) distributed with \( p \) and \( v \) degrees of freedom and noncentrality \( \tau^2 = p^{-1} \chi^2 \).

When \( y = 0 \), the distribution is the central \( F \).

**Proof.** Let \( v \chi^2 \sim \chi^2_y \). Then \( \chi | S^2 - \chi^2 \sim \chi^2(y, s^2 - \chi^2) \) and \( S^2(\chi^2 - \chi^2) | S^2 \sim F_p, v \),

i.e., noncentral chi-square distributed with \( p \) degrees of freedom and noncentrality \( \tau^2 \), and independent of \( S^2 \). This implies that the quadratic form

\[
p^{-1} \chi^2 \sim F_p, v(\tau^2).
\]

It is clear that \( F_p, v(0) \) is the central \( F_p, v \). Cornish [4] has obtained the result for the special case when \( y = 0 \).

**Theorem 3.** Let \( \chi \sim \Gamma_y(y, v, p) \), defined by Representation B. Then

\[
p^{-1}(\chi - y)^{-1}(\chi - y) \text{ is distributed as central } F \text{ with } p \text{ and } v \text{ degrees of freedom.}
\]
Proof. Let $C$ be a $p \times p$ nonsingular matrix such that $C C' = I$ and let $Z = C(X-Y)$. Then by Theorem 1, $Z \sim T_p(q,1,p)$. Thus $Z$ may be defined by

$$
Z = (V^k)^{-1}Y,$n

where $V^k$ is the symmetric square root of $V$, i.e., $V^{k/2} = V \sim W(1, v+p-1)$ and $Y \sim N(q,v1)$, independent of $V$. Then we have

$$
\frac{v+p-1}{v} Z^2 = \frac{v+p-1}{v} Y V^{-1} Y,'n

which is the Hotelling's $T^2$-statistic. Therefore,

$$p^{-1}(X-Y)^T (X-Y) = p^{-1}Z^2 \sim F_{p, v}.$n

(See e.g., Anderson [1], p. 106).

Theorem 3 states that the quadratic form $p^{-1}X'X$ is central $F$ distributed with $p$ and $v$ degrees of freedom when $Y = 0$. For the case when $Y \neq 0$, we have not been able to obtain its distribution.

We will now obtain the marginal distribution of a subvector of a multivariate $t$. Let us partition $p \times 1$ vectors $X$ and $Y$, and $p \times p$ positive definite symmetric matrices $\Sigma$ and $V$ as follows:

$$
X = \begin{pmatrix} X^{(1)} \\ \Sigma^{(2)} \end{pmatrix}, \quad Y = \begin{pmatrix} Y^{(1)} \\ \Sigma^{(2)} \end{pmatrix}, \quad \Sigma = \begin{pmatrix} E_{11} & E_{12} \\ E_{21} & E_{22} \end{pmatrix}, \quad V = \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix},
$$

where $X^{(1)}$ and $Y^{(1)}$ are $q \times 1$, $(q < p)$, $X^{(2)}$ and $Y^{(2)}$ are $(p-q) \times 1$ subvectors, while $E_{11}$ and $V_{11}$ are $q \times q$, $E_{12}$ and $V_{12}$ are $q \times (p-q)$, $E_{22}$ and $V_{22}$
are \((p-q) \times (p-q)\) matrices. It is well known that if \(\Sigma\) is positive definite symmetric, so are \(\Sigma_{11}\) and \(\Sigma_{22}\).

**Theorem 4.** Let \(\chi \sim T_V(\xi, \Sigma, \nu)\), be defined by Representation A. Then

1. \(\chi^{(1)} \sim T_V(\xi^{(1)}, \Sigma_{11}, \nu)\) and \(\chi^{(2)} \sim T_V(\xi^{(2)}, \Sigma_{22}, \nu-p-q)\); 

2. If \(\Sigma_{12} = 0\), \(\frac{(\xi^{(1)} - \xi^{(2)})' \Sigma_{11}^{-1} (\xi^{(1)} - \xi^{(2)})}{(\xi^{(1)} - \xi^{(2)})' \Sigma_{22}^{-1} (\xi^{(1)} - \xi^{(2)})} \sim F_{p-q, \nu-p-q}\). 

**Proof.** (1) Let \(\nu \sim \chi_\nu^2\). Then

\[\nu|s^2 = s^2 \sim \chi(\nu, s^2),\]

\[\Rightarrow \chi^{(1)}|s^2 = s^2 \sim \chi(\nu^{(1)}, s^2, \nu_{11}),\]

\[\Rightarrow \chi^{(1)} \sim T_V(\xi^{(1)}, \Sigma_{11}, \nu).\]

Similarly, we can obtain that \(\chi^{(2)} \sim T_V(\xi^{(2)}, \Sigma_{22}, \nu-p-q)\).

(11) It follows from (1) and the proof of Theorem 2 that

\[s^2 (\xi^{(1)} - \xi^{(2)})' \Sigma_{11}^{-1} (\xi^{(1)} - \xi^{(2)}) | s^2 \sim \chi_q^2,\]

and

\[s^2 (\xi^{(2)} - \xi^{(2)})' \Sigma_{22}^{-1} (\xi^{(2)} - \xi^{(2)}) | s^2 \sim \chi_{p-q}^2.\]

If \(\Sigma_{12} = 0\), \(\chi^{(1)}\) and \(\chi^{(2)}\) are conditionally independent, given \(s^2\). Hence
\[
\begin{align*}
\frac{\binom{p-q}{q}}{inom{(1)}{1} \binom{(1)}{1} \binom{(1)}{1} \binom{(1)}{1}} & = \frac{\binom{(2)}{2} \binom{(2)}{2} \binom{(2)}{2}}{inom{(2)}{2} \binom{(2)}{2} \binom{(2)}{2}} \binom{2}{2} \cdot \psi_{1,p-q},
\end{align*}
\]

which is independent of \( S^2 \), completing the proof of the theorem.

**Theorem 5.** Let \( Y \sim T_{v}(\bar{y}, \bar{z}, p) \), defined by Representation B. Then

\[(10) \quad \chi^{(1)} \sim T_{v}(\bar{y}^{(1)}, \bar{z}_{11}, q) \]

and

\[(11) \quad \chi^{(2)} \sim T_{v}(\bar{y}^{(2)}, \bar{z}_{22}, p-q). \]

**Proof.** Let \( V \sim N_{v}^{-1}(1, v+p-1) \). Then

\[
\begin{align*}
\chi | V & \sim N(\bar{y}, vv^{-1}), \\
\implies & \chi^{(1)} | V \sim N(\bar{y}^{(1)}, vv_{11}^{-1}), \\
where \quad V_{11+2} & = V_{11} - V_{22} vv_{22}^{-1}, \quad \sim N_{v_{11+2}}^{-1}(1, v_{11+2}^{11+2}-1), \\
\implies & \chi^{(1)} \sim T_{v}(\bar{y}^{(1)}, \bar{z}_{11}, q).
\end{align*}
\]

Similarly, we can obtain that \( \chi^{(2)} \sim T_{v}(\bar{y}^{(2)}, \bar{z}_{22}, p-q). \)

4. **Characterizations of the multivariate t distribution.** Cornish [6] gives a necessary and sufficient condition for a multivariate t vector defined by

Representation A. We present here the same necessary and sufficient condition under

Representation B and also give a simple proof under Representation A, utilizing

the results of a multivariate normal distribution. Two other characterizations

are also obtained under Representation A.
Theorem 6. $X \sim T_p(y, L, p)$ if, and only if, for any $p = 1$ scalar vector $g \neq 0$,

$$s^2(\frac{x - y}{a^2 z})^\frac{1}{2} \sim \tau_v.$$ 

Proof: (1) Under Representation $A$, let $\omega S^2 \sim X_Y^2$, then

$$\frac{x}{S^2} \sim N(y, s^{-2} I),$$

$$\Leftrightarrow g^2(\frac{x - y}{a^2 z})^\frac{1}{2} | S^2 \sim N(0, s^{-2}),$$

for any $g \neq 0$,

$$\Leftrightarrow g^2(\frac{x - y}{a^2 z})^\frac{1}{2} \sim \tau_v.$$ 

(ii) Under Representation $B$, let $V \sim N(\Sigma^{-1}, \nu + p - 1)$, then

$$\frac{v}{\nu} \sim N(y, \nu^{-1}),$$

$$\Leftrightarrow g^2(\frac{x - y}{a^2 z})^\frac{1}{2} | V \sim N(0, \nu(g^2 v^{-2})/(g^2 z)), $$

for any $g \neq 0$,

$$\Leftrightarrow g^2(\frac{x - y}{a^2 z})^\frac{1}{2} \sim \tau_v,$$

since $g^2 z / g^2 v^{-2} \sim s^2$. 

In order to establish some further characterizations, we need the following lemmas.

Lemma 1. Let $X_1, \ldots, X_p$ be independent and identically distributed (iid) random variables. Then $\sum_{k=1}^{p} X_k^2 \sim \chi_r^2$ if, and only if $\chi_{r_k}^2 \sim \chi_1^2, k = 1, \ldots, p.$
Proof. This is obvious, since

$$\mathbb{E} \exp\left(\frac{it}{\sqrt{\lambda}} X^2_k \right) = \left[ \mathbb{E} \exp(it X^2_k) \right]^\frac{1}{\lambda} = (1-2it)^{-\lambda/4}.$$ 

Lemma 2. Let \( Y \) be a continuous random variable defined on the real line symmetrically distributed about 0 with unit variance. Then \( Y^2 \sim X^2_1 \) if, and only if \( Y \sim N(0,1) \).

Proof: Since \( Y^2 \sim X^2_1 \) if, and only if its pdf can be written as

$$f(y) = \beta(y) e^{-\frac{y^2}{2}}, \quad -\infty < y < \infty,$$

where \( \beta(y) + \beta(-y) = c \), a constant for all \( y \). (See Roberts and Geisser [15]). The symmetry of the density implies that \( \beta(y) = (2\pi)^{-1/2} \) for all \( y \) and hence \( Y \sim N(0,1) \).

Lemma 3. Let \( Y_1, \ldots, Y_p \) be independent random variables with finite variances. Assume that \( Y_k \), \( k = 1, \ldots, p \), has the pdf \( h_k(y) \) such that \( h_k(y) > 0 \) and differentiable for all \( y \), \( -\infty < y < \infty \). Then the joint pdf of \( Y_1, \ldots, Y_p \) is a function only of \( Y_1^2 + \cdots + Y_p^2 \) if, and only if \( (Y_1, \ldots, Y_p)^T \sim N(0, \sigma^2 I) \), for some \( \sigma^2 < \infty \).

Proof. The sufficient condition is obvious. We shall prove the necessary condition. The joint pdf of \( Y_1, \ldots, Y_p \) can be written as

$$\prod_{k=1}^p h_k(y_k) = \rho_0\left( \frac{1}{\lambda} Y_k^2 \right),$$

where \( \rho_0(t) \) is a differentiable function of \( t \). This implies that the marginal pdf of \( Y_k \), \( k = 1, \ldots, p \), is
(14) \[ h_k(y_k) = s_k(y_k^2), \]

where \( s_k(t) \) is a differentiable function of \( t \). Hence we have

(15) \[ s_0(\sum_{k=1}^{p} y_k^2) = \prod_{k=1}^{p} s_k(y_k^2). \]

Differentiating both sides of (15) with respect to \( y_i^2, i = 1, \ldots, p \), we have

(16) \[ \frac{d}{dy_i} s_0(\sum_{k=1}^{p} y_k^2) = \prod_{k=1}^{p} s_k(y_k^2), \]

where \( s_k(t) \) denotes the derivative of \( s_k(t) \) with respect to \( t, k = 0, \ldots, p \).

Therefore, for all \( y_i \), \( i = 1, \ldots, p \), we have

(17) \[ s_i^2(y_i^2)/s_1(y_1^2) = c, \]

where \( c \) is a constant independent of \( i \). Solving the differential equation (17) gives

(18) \[ s_i(y_i^2) = d \cdot e^{c_1 y_i^2}, \quad i = 1, \ldots, p, \]

where \( d \) is also a constant. This shows that \( Y_i \sim N(0, \sigma_i^2) \), where \( \sigma_i^2 = -(2c)^{-1} \).

But since \( c \) does not depend on \( i \), the variances of \( Y_1, \ldots, Y_p \) must be equal, say \( \sigma^2 \). Hence, \( (Y_1, \ldots, Y_p)' \sim N(0, \sigma^2 I) \).

For the remainder of this section, we will consider a multivariate t vector defined by Representation A.

Theorem 7. Let \( \nu s^2 \sim \chi^2_\nu \) and let \( X_1, \ldots, X_p \) be continuous random variables defined on the real line. Assume that \( X_1, \ldots, X_p \) are conditionally iid, and that \( X_1 \) is symmetrically distributed about its mean \( \mu \) with finite variance, given \( s^2 = \sigma^2 \). Then
\[ (p-2\gamma)^{-1} \sum_{k=1}^{p} (X_k - \mu)^2 \sim \chi^2_p, \nu \]

if, and only if \((X_1, \ldots, X_p) \sim T_v(y_0, \sigma^2 I, p),\) where \(y = (\nu_1, \ldots, \nu_p)\), and \(I\) is the identity matrix of order \(p\).

**Proof.** The sufficient condition is a special case of Theorem 2. We shall show the necessary condition. Let \(Y_k = (X_k - \mu)/\sigma, k = 1, \ldots, p.\) Then the \(Y_k\)'s also satisfy the conditions of the theorem with mean 0 and finite variance, given \(S^2 = \sigma^2\).

It is well known that an \(F_{p, \nu}\)-distributed random variable may be represented as the ratio of two independent chi-square random variables with \(p\) and \(\nu\) degrees of freedom, respectively, multiplied by a suitable constant. Thus, let \(S^2_1 \sim \chi^2_p\) independent of \(S^2\), then we have

\[ \sum_{k=1}^{p} Y_k^2 \sim \frac{S^2_1}{S^2}, \]

or

\[ \sum_{k=1}^{p} (SY_k)^2 \sim \frac{S^2}{\frac{S^2_1}{S^2}}, \sim \chi^2_p, \]

which is independent of \(S^2\). By Lemma 1, this implies that

\[ (SY_k)^2 \sim \chi^2_1, \quad k = 1, \ldots, p, \]

and by Lemma 2, we have \(SY_j^2 \sim \chi^2 \sim \chi^2 \sim \chi^2_1\), where \(Y = (Y_1, \ldots, Y_p)\), and

\(S^2 \sim \chi^2_1.\) Thus, by Representation \(A, Y \sim T_v(\nu, 0, I, p)\) and by Theorem 1, \(X \sim T_v(y_0, \sigma^2 I, p).\)

**Theorem 3.** Let \(n \chi^2 \sim \chi^2_0\) and let \(X_k, k = 1, \ldots, p,\) be conditionally independent random variables, given \(S^2,\) with \(\text{Var}(X_k | S^2) = \sigma_k^2 / \sigma^2 < \infty.\) Assume that \(X_k\)

\((k = 1, \ldots, p),\) given \(S^2 = \sigma^2,\) has the conditional pdf \(f \chi^2_k(x | \sigma^2)\) which is positive
and differentiable for all $x_1, \ldots, x_p$. Then the joint pdf of $X_1, \ldots, X_p$ is a function only of $x_1^2 + \cdots + x_p^2$ if, and only if $(X_1, \ldots, X_p)' \sim T_{p}(\Omega, \sigma^2 I, p)$, for some $\sigma^2 < \infty$.

**Proof.** The sufficient condition is easily seen from the joint pdf of $X_1, \ldots, X_p$. We shall prove the necessary condition. Since the joint pdf of $X_1, \ldots, X_p$ is a function of $x_1^2 + \cdots + x_p^2$, it follows that the conditional joint pdf of $X_1, \ldots, X_p$, given $S^2 = s^2$, is a function of $x_1^2 + \cdots + x_p^2$ and $s^2$ almost surely with respect to a $\sigma$-finite measure $\mu$, i.e.,

$$
\frac{p}{k=1} f_k(x_k | s^2) = \frac{p}{k=1} \frac{f_k(x_k | s^2)}{f_k(x_k | s^2)}, \quad a.s. (\mu).
$$

Similarly, the conditional marginal pdf of $X_k \ (k = 1, \ldots, p)$, given $S^2 = s^2$, is

$$
f_k(x_k | s^2) = \frac{f_k(x_k | s^2)}{f_k(x_k | s^2)}, \quad a.s. (\mu),
$$

where $f_k(t, s^2) \ (k = 0, \ldots, p)$ is a positive and differentiable function of $t \geq 0$. We will drop the argument $s^2$ in the functions $f_k(t, s^2)$ since it is treated as a constant when $S^2 = s^2$ is given. Now, the result of Lemma 3 and the fact that $\text{Var}(X_k | s^2) = \sigma_k^2/s^2$ show that $(X_1, \ldots, X_p)' | S^2 = s^2 \sim T_{p}(\Omega, s^2 I, p)$, where $\sigma^2 = \sigma_k^2$ for all $k = 1, \ldots, p$. Therefore, $(X_1, \ldots, X_p)' \sim T_{p}(\Omega, \sigma^2 I, p)$.


