RT-MAE 9410

ON SOME CHARACTERIZATIONS OF THE t-DISTRIBUTION

by

Reinaldo B. Arellano-Valle
and
Heleno Bolfarine

Palavras Chaves: Compound normal distributions, elliptical distributions,
(Key words) quadratic forms, spherical distributions.

Classificação AMS: 62E10, 62H05
(AMS Classification)

- abril de 1994 -

ON SOME CHARACTERIZATIONS OF THE t-DISTRIBUTION

Reinaldo B. Arellano-Valle and Heleno Bolfarine
Departamento de Estatística - IME
Universidade de São Paulo
Caixa Postal 20570 - CEP 01452-990
São Paulo - S.P. - Brasil

Summary

In this paper we discuss three diferent characterizations of the generalized t-distribution within the class of the eliptical distributions. We show that this distribution can be characterized in terms of its unconditional and conditional marginals and in terms of quadratic forms. Similar results have been proved for the normal distribution. An additional characterization of the t distribution within the subclass of the conpound normal distributions (or scale mixture of normal distributions) is also studied.

1. Introduction

The normal distribution has been characterized within the class of the elliptical distributions in several different ways by using some of its properties. Important references in this direction are Kelker (1970), Cambanis et al. (1981), Anderson and Fang (1987) and Khatri and Mukerjee (1987). We show in this paper that the t-distribution can be characterized within the class of the eliptical distributions by using some of its properties. Three results of this nature are presented. The first result charcterizes the generalized t-distribution in terms of its marginals. The second result presents a characterization of the t-distribution in terms of its marginal conditional distributions. The proof of this result, which considers the existence of a density, is similar to the proof considered in Kelker (1970) for the characterization of the normal distribution. Kelker's (1970) results are extendend by Cambanis et al. (1981), where characterizations of the normal distribution are provided within the class of the elliptical nondegenerate distributions. We conjecture that similar results also holds for the case of the generalized t-distribution with similar hypothesis as the ones considered in Cambanis et al. (1981). Further, it is shown that the generalized t-distribution can also be characterized in terms of some special quadratic forms. This result is proved by considering a preliminary fact stated in Anderson and Fang (1987) on the spherical distributions which put zero mass in the origin. A similar characterization for the normal distribution appears in Khatri and Mukerjee (1987) within the class of the spherical distributions which do not puts zero mass in the origin. Finally, we show that the t-distribution can be characterized within the subclass of the spherical distributions which are scale mixtures of the normal distribution. Linearity conditions on the variance of the univariate conditional distributions are used in this case. The main result is obtained as a consequence of a well known theorem due Diaconis ans Ylvisaker (1979), where it is shown that in the regular exponential family (with the natural parametrization), if the posterior expectation is linear then the prior distribution must be conjugated. We begin with some preliminary results of the elliptical and t-distributions.

2. Some preliminary results

In this section we present some known facts about the spherical, elliptical and t-distributions, which are used to prove the main results of this paper. We start with the spherical and elliptical distributions.

2.1. The spherical and elliptical distributions

Let $\mathcal{O}_n$ be the set (group) of the nxn orthogonal matrices. An n-dimensional random (column) vector $\mathbf{X}$ is said to have spherical (symmetric) distributions if for every $\mathbf{\Gamma} \in \mathcal{O}_n$ (see Kelker, 1990),

$$\mathbf{\Gamma X} \stackrel{d}{=} \mathbf{X},$$

where $Y \stackrel{d}{=} Z$ means that Y and Z have the same distribution. If ϕ_X denotes the characteristic function of $\mathbf{X}$, then

$$\phi_X(\mathbf{t}) = \phi(\mathbf{t}'\mathbf{t}), \quad \mathbf{t} \in \mathcal{R}^n,$$

for some function ϕ and we write $\mathbf{X} \sim S_n(\phi)$. Similarly, if $\mathbf{X}$ has a density function p_X, then

$$p_X(\mathbf{x}) = f(\mathbf{x}'\mathbf{x}), \quad \mathbf{x} \in \mathcal{R}^n,$$

for some function f. In this case we write $\mathbf{X} \sim S_n(f)$. For example, if

$$\phi(u) = e^{-\frac{u}{2}} \quad \text{or} \quad f(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u}{2}}, \quad u \geq 0,$$

then $S_n(\phi)$ or $S_n(f)$ is the spherical normal distribution, denoted by $N_n(0; \mathbf{I}_n)$, where $\mathbf{I}_n$ is the identity matrix of dimension n.

Now, let $\mathbf{U}^{(n)}$ denotes a random vector that is uniformly distributed on the unit sphere in $\mathcal{R}^{(n)}$. Cambanis et al. (1981) pointed out the fact that $\mathbf{X} \sim \mathbf{S}_n(\phi)$ if and only if

$$\mathbf{X} \stackrel{d}{=} R\mathbf{U}^{(n)}, \tag{2.1}$$

where $R \stackrel{d}{=} ||\mathbf{X}||$ (with $||\mathbf{x}|| = (\mathbf{x}'\mathbf{x})^{1/2}$) is independent of $\mathbf{U}^{(n)}$ and the distribution function F_R of R is related to ϕ through

$$\phi(u) = \int_{[0,\infty)} \Psi_n(r^2 u) dF_R(r),$$

where

$$\Psi_n(\mathbf{t}'\mathbf{t}) = E[e^{i\mathbf{t}'\mathbf{U}^{(n)}}].$$

Moreover, if $P[\mathbf{X} = 0] = 0$ (or $F_R(0) = 0$), then $||\mathbf{X}||$ and $\mathbf{X}/||\mathbf{X}||$ are independent and

$$\frac{\mathbf{X}}{||\mathbf{X}||} \stackrel{d}{=} \mathbf{U}^{(n)}.$$

It is easy to see that if $\mathbf{X} \sim S_n(f)$, then R has a density function p_R related to the density f through (Kelker, 1970; Cambanis et al., 1981)

$$(2.2) \qquad p_R(r) = \frac{2\pi^{n/2} r^{n-1}}{\Gamma[n/2]} f(r^2), \quad r \geq 0.$$

Suppose now that

$$\mathbf{X} \stackrel{d}{=} V^{1/2}\mathbf{Z},$$

where $V \sim F_V$ on $[0, \infty)$ and $\mathbf{Z} \sim N_n(\mathbf{0}, \mathbf{I}_n)$ are independent. Then, $\mathbf{X} \sim S_n(\phi)$ with

$$(2.3) \qquad \phi(u) = \int_{[0,\infty)} e^{-\frac{1}{2}uv} dF_V(v), \quad u \geq 0,$$

and the distribution of $\mathbf{X}$ is a scale mixture of normal distributions. The subclass of such spherical distributions formed by varying F_V is called the class of the compound normal distributions. Further, if $F_V(0) = 0$, then $\mathbf{X} \sim S_n(f)$ (Cambanis et al., 1981) with

$$(2.4) \qquad f(u) = \int_0^\infty (2\pi v)^{-n/2} e^{-\frac{u}{2v}} dF_V(v), \quad u \geq 0.$$

In the special case where $V \sim IG(\nu/2, \lambda/2)$ or, equivalentely, $V \sim \lambda/\chi^2_\nu$, where $IG(\alpha, \beta)$ denotes the inverted gamma distribution with slope and scale parameters α and β, respectively, that is, V has density function given by

$$(2.5) \qquad p_V(v) = \frac{\beta^\alpha}{\Gamma[\alpha]}(1/v)^{\alpha+1} e^{-\frac{\beta}{v}}, \quad v > 0,$$

then, from (2.4), we have that

$$(2.6) \qquad f(u) = k(n, \nu)\lambda^{\frac{\nu}{2}}\{\lambda + u\}^{-\frac{(\nu+n)}{2}}, \quad u \geq 0,$$

where

$$(2.7) \qquad k(n, \nu) = \frac{\Gamma[\frac{(\nu+n)}{2}]}{\Gamma[\frac{\nu}{2}]\pi^{\frac{n}{2}}}.$$

In this case, $S_n(f)$ is a generalized version of the n-variate t-distribution (Dickey, 1967), which we denote by $t_n(\mathbf{0}, \mathbf{I}_n; \lambda, \nu)$. When $\lambda = \nu$, we have that $t_n(\mathbf{0}, \mathbf{I}_n; \nu, \nu) = t_n(\mathbf{0}, \mathbf{I}_n; \nu)$ is the usual n-variate t-distribution with ν degrees of freedom. Conversely, we can show that if $\mathbf{X} \sim t_n(\mathbf{0}, \mathbf{I}_n; \lambda, \nu)$ then in (2.4), $F_V \sim IG(\nu/2, \lambda/2)$ (Andrews and Mallows, 1974). Further, it is easy to see that $\mathbf{X} \sim t_n(\mathbf{0}, \mathbf{I}_n; \lambda, \nu)$ if and only if in (2.1), $R^2 \sim BeII(n/2, \lambda/2)$, or equivalently, $R^2 \sim (n\lambda/\nu)F_{n,\nu}$, where $BeII(\alpha, \beta)$ denotes the ordinary beta II distribution with parameters α and β and density given by

$$(2.8) \qquad \frac{\Gamma[\alpha + \beta]}{\Gamma[\alpha]\Gamma[\beta]} u^{\alpha-1}(1 + u)^{-(\alpha+\beta)}, \quad u > 0.$$

The class of elliptical distributions can be defined in a number of equivalent ways (see, for example, Kelker, 1970; Cambanis et al. 1981; Muirhead, 1982 and Fang et al. 1990). We adopt here the following definition. An nx1 random vector $\mathbf{X}$ is said to have an elliptical (symmetric) distribution with parameters $\mu \in \mathcal{R}^n$ and Σ of dimensions nx1 and nxn, respectively, with Σ being nonnegative definite ($\Sigma \geq 0$), that is, $\Sigma = \mathbf{A}\mathbf{A}'$ where $\mathbf{A}$ is a nxk matrix of rank k, if

$$\mathbf{X} \stackrel{d}{=} \mu + \mathbf{A}\mathbf{Y}, \quad \mathbf{Y} \sim S_k(\phi).$$

The notation used is $\mathbf{X} \sim E_n(\mu, \Sigma; \phi)$ and for $\mu = 0$ and $\Sigma = \sigma^2\mathbf{I}_n$, we have that $E_n(0, \sigma^2\mathbf{I}_n; \phi) = S_n(\phi)$. Moreober, if $\mathbf{X} \sim E_n(\mu, \Sigma; \phi)$, then

$$\phi_X(\mathbf{t}; \mu, \Sigma) = e^{i\mathbf{t}'\mu}\phi(\mathbf{t}'\Sigma\mathbf{t}), \quad \mathbf{t} \in \mathcal{R}^n.$$

In the case where Σ is positive definite ($\Sigma > 0$), and $\mathbf{X} \sim E_n(\mu, \Sigma; f)$, we have that

$$p_X(\mathbf{x}|\mu, \Sigma) = |\Sigma|^{-\frac{1}{2}} f((\mathbf{x} - \mu)'\Sigma^{-1}(\mathbf{x} - \mu)), \quad \mathbf{x} \in \mathcal{R}^n. \tag{2.9}$$

Note that $\mathbf{X} \sim E_n(\mu, \Sigma; f)$ is the nonsingular $t_n(\mu, \Sigma; \lambda, \nu)$ distribution when f is as given by (2.6).

2.2. Some properties of the t-distribution

The $t_n(\mu, \Sigma; \lambda, \nu)$ distribution can be defined as the distribution of the nx1 random vector $\mathbf{X}$ such that

$$\mathbf{X} \stackrel{d}{=} \mu + V^{\frac{1}{2}}\mathbf{Z}, \tag{2.10}$$

where $V \sim GI(\nu/2, \lambda/2)$ and $\mathbf{Z} \sim N_n(0, \Sigma)$ are independent. For $\Sigma > 0$ we have that $\mathbf{X} \sim E_n(\mu, \Sigma; f)$, that is, $\mathbf{X}$ has density as given in (2.9) with f as given in (2.6). From (2.10) and (2.7) it is easy to see that

$$E[\mathbf{X}] = \mu, \quad Var[\mathbf{X}] = E[(\mathbf{X} - \mu)(\mathbf{X} - \mu)'] = \frac{\lambda}{\nu - 2}\Sigma, \tag{2.11}$$

for $\nu > 1$ and $\nu > 2$, respectively. Further, from (2.10) it follows that

$$\mathbf{Y} = \eta + \mathbf{B}\mathbf{X} \stackrel{d}{=} (\eta + \mathbf{B}\mu) + V^{1/2}\mathbf{B}\mathbf{Z} \sim t_m(\eta + \mathbf{B}\mu; \mathbf{B}\Sigma\mathbf{B}'; \lambda, \nu), \tag{2.12}$$

since $\mathbf{B}\mathbf{Z} \sim N_m(0, \mathbf{B}\Sigma\mathbf{B}')$, where $\eta \in \mathcal{R}^n$ is a mx1 vector and $\mathbf{B}$ a mxn matrix. Now, let

$$\mathbf{X} = \begin{pmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{pmatrix}, \quad \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$$

and

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix},$$

where $\mathbf{X}_1$ and μ_1 are mx1 vectors, Σ_{11} is a mxm matrix and so on. Thus, taking $\mathbf{B} = [\mathbf{I}_m \quad \mathbf{0}]$ in (2.12) we have that

$$\mathbf{X}_1 \sim t_m(\mu_1, \Sigma_{11}; \lambda, \nu).$$

The marginal distribution of $\mathbf{X}_2$ follows by symmetry. Assuming that $\Sigma > 0$, let

$$\mu_1(\mathbf{x}_2) = \mu_1(\mathbf{x}_2) + \Sigma_{12}\Sigma_{22}^{-1}(\mathbf{x}_2 - \mu_2),$$

$$\Sigma_{11.2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$$

and

$$q(\mathbf{x}_2) = (\mathbf{x}_2 - \mu_2)'\Sigma_{22}^{-1}(\mathbf{x}_2 - \mu_2).$$

Using the fact that (Muirhead, 1982)

$$|\Sigma| = |\Sigma_{11.2}||\Sigma_{22}|$$

and

$$(\mathbf{x} - \mu)'\Sigma^{-1}(\mathbf{x} - \mu) = (\mathbf{x}_1 - \mu_1)'\Sigma_{11.2}^{-1}(\mathbf{x}_2 - \mu_1(\mathbf{x}_2)) + q(\mathbf{x}_2),$$

we have that the conditional density of $\mathbf{X}_1$ given $\mathbf{X}_2 = \mathbf{x}_2$ is given by

$$p_{\mathbf{X}_1}(\mathbf{x}_1|\mu, \Sigma, \lambda, \nu) =$$

$$= |\Sigma|^{-\frac{1}{2}} f_{q(\mathbf{x}_2)}((\mathbf{x}_1 - \mu_1(\mathbf{x}_2))'\Sigma_{11.2}^{-1}(\mathbf{x}_2 - \mu_1(\mathbf{x}_2)), \quad \mathbf{x}_1 \in \mathcal{R}^m,$$

where

(2.13) $$f_a(u) = k(m, \nu_m)\lambda_a^{\frac{1}{2}\nu_m}\{\lambda_a + u\}^{-\frac{1}{2}(\nu_m + m)}, u \geq 0,$$

with

$$\nu_m = \nu + m - m \quad \text{and} \quad \lambda_a = \lambda + a,$$

and $k(.,.)$ is as given in (2.7). This means that

(2.14) $$(\mathbf{X}_1|\mathbf{X}_2 = \mathbf{x}_2) \sim t_m(\mu_1(\mathbf{x}_2), \Sigma_{11.2}; \lambda_{q(\mathbf{x}_2)}, \nu_m),$$

or, equivalentely,

$$(\mathbf{X}_1|\mathbf{X}_2 = \mathbf{x}_2) \sim E_m(\mu_1(\mathbf{x}_2), \Sigma_{11.2}; f_{q(\mathbf{x}_2)}),$$

with f_a given by (2.13). Note that for $\lambda = \nu$,

$$(\mathbf{X}_1|\mathbf{X}_2 = \mathbf{x}_2) \sim t_m(\mu_1(\mathbf{x}_2), \Sigma_{11.2}; \nu_{q(\mathbf{x}_2)}, \nu_m),$$

with $\nu_a = \nu + a \neq \nu_m = \nu + n - m$, which means that the usual t-distribution does not retain its conditional distributions. Finally, from (2.14) and (2.11) (with $\lambda = \lambda_a$ and $\nu = \nu_m$) it follows that

$$E[\mathbf{X}_1|\mathbf{X}_2] = \mu_1(\mathbf{X}_2) = \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(\mathbf{X}_2 - \mu_2)$$

and

$$(2.15)\ Cov[\mathbf{X}_1|\mathbf{X}_2] = \frac{\lambda_{q(\mathbf{X}_2)}}{\nu_m - 2}\Sigma_{11.2} = \frac{\lambda + (\mathbf{X}_2-\mu_2)'\Sigma_{22}^{-1}(\mathbf{X}_1-\mu_2)}{\nu+n-m-2}(\Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}).$$

Now, let $\mathbf{X} \sim t_n(\mathbf{0}, \mathbf{I}_n; \lambda, \nu)$ and $\mathbf{A}$ a nxn symmetric matrix. Then we have that $\mathbf{X}'\mathbf{A}\mathbf{X} \sim \lambda BeII(m/2, \nu/2)$ (or, equivalentely, $(m\lambda/\nu)F_{m,\nu}$) if and only if $\mathbf{A}^2 = \mathbf{A}$ and $rank(\mathbf{A}) = m$, $1 \leq m < n$. In fact, since $\mathbf{X} \stackrel{d}{=} V^{1/2}\mathbf{Z}$, where $V \sim IG(\nu/2, \lambda/2)$ and $\mathbf{Z} \sim N_n(\mathbf{0}, \mathbf{I}_n)$ are independent, we have that $\mathbf{X}'\mathbf{A}\mathbf{X} \stackrel{d}{=} V\mathbf{Z}'\mathbf{A}\mathbf{Z}$ where V and $\mathbf{Z}'\mathbf{A}\mathbf{Z}$ are independent. Recalling that $\mathbf{Z}'\mathbf{A}\mathbf{Z} \sim \chi^2_m$ if and only if $\mathbf{A}^2 = \mathbf{A}$ and $rank(\mathbf{A}) = m$, the result follows.

Remark 1. Let $S \sim \chi^2_r/r$ and $T \sim s/\chi^2_s$ (or $T^{-1} \sim \chi^2_s/s$) independent. Then,

$$F = ST \sim F_{r,s} = (s/r)BeII(r/2, s/2).$$

Remark 2. Let $S \sim nF_{n,\nu}$ and $T \sim Be(m/2, (n-m)/2)$, namely, the ordinary beta distribution, independent, where $1 \leq m \leq n$. Then, $F = ST \sim (m/\nu)F_{m,\nu} = BeII(m/2, \nu/2)$.

3. A characterization in terms of its marginal distributions

The theorem that we present next provides a characterization of the generalized t-distribution in terms of its marginals. We use the notation and the results presented in Section 2.

Theorem 1. *Let $\mathbf{X} \sim E_n(\mu, \Sigma, \phi)$. Then, any marginal distribution is a generalized t-distribution if and only if $\mathbf{X}$ has a generalized t-distribution.*

Proof. (For the "only if" part see Section 2.2). To prove the converse, we take (without loss of generality) the spherical case, with $\mu = 0$ and $\Sigma = \mathbf{I}_n$. Thus, $\mathbf{X} = (\mathbf{X}_1', \mathbf{X}_2')' \sim E_n(\mathbf{0}, \mathbf{I}_n; \phi)$, where $\mathbf{X}_1$ is mx1, $1 \leq m < n$. Suppose that $\mathbf{X}_1 \sim t_m(\mathbf{0}, \mathbf{I}_m; \lambda, \nu)$. Then, $\mathbf{X}_1 \sim E_n(\mathbf{0}, \mathbf{I}_n; \phi)$, where ϕ is given by (2.3) with $F_\nu = GI(\nu/2, \lambda/2)$. Since $\mathbf{X}$ has characteristic function $\phi(\mathbf{t}'\mathbf{t})$, $t \in \mathcal{R}^n$, then $\mathbf{X} \sim t_n(\mathbf{0}, \mathbf{I}_n; \lambda, \nu)$, as was to be proved. ∎

4. A characterization in terms of its conditional distributions

We present now a characterization of the t-distribution in terms of its conditional distributions. Use is made of some results presented in Section 2.2.

Theorem 2. *Let $\mathbf{X} = (\mathbf{X}_1', \mathbf{X}_2')' \sim E_n(\mu, \Sigma; f)$ where $\mathbf{X}$ is $mx1$, $1 \leq m < n$. Then, the conditional distribution of $\mathbf{X}_1$ given $\mathbf{X}_2$ is the generalized m-variate t-distribution if and only if the distribution of $\mathbf{X}$ is the generalized n-variate t-distribution.*

Proof. For the "only if" part see Section 2.2. To prove the reverse statement, it sufficies to consider the spherical case only, that is, $\mu = \mathbf{0}$ and $\Sigma = \mathbf{I}$, since the general form of the conditional density of $\mathbf{X}_1$ given $\mathbf{X}_2$ is independent of μ and Σ. Thus, in the spherical case, the conditional density of $\mathbf{X}_1$ given $\mathbf{X}_2$ is given by (Kelker, 1980; Cambanis et al., 1981)

$$p_{\mathbf{X}_1}(\mathbf{x}_1|\mathbf{x}_2) = f_{||\mathbf{x}||^2}(\mathbf{x}_1'\mathbf{x}_1), \quad \mathbf{x}_1 \in \mathcal{R}^m,$$

where

$$(4.1) \qquad f_a(u) = \frac{f(u+a)}{f_2(a)}, \quad u \geq 0, a \geq 0,$$

with

$$f_2(a) = \frac{2\pi^{\frac{m}{2}}}{\Gamma[\frac{m}{2}]} \int_0^\infty f(r^2+a) r^{m-1} dr,$$

that is, $f(\mathbf{x}_2'\mathbf{x}_2)$, $\mathbf{x}_2 \in \mathcal{R}^{n-m}$, is the marginal density of $\mathbf{X}_2$.

Suppose then that

$$(\mathbf{X}_1 | \mathbf{X}_2 = \mathbf{x}_2) \sim t_m(0; \mathbf{I}_m \lambda_{||\mathbf{x}||^2}, \nu_m),$$

where $\lambda_a = \lambda + a$ and $\nu_m = \nu + n - m$. This means that

$$(\mathbf{X}_1 | \mathbf{X}_2 = \mathbf{x}_2) \sim El_m(0; \mathbf{I}_m; f_{||\mathbf{x}||^2}),$$

where $f_a(.)$ is given by (2.13). Thus, it follows from (4.1) and (2.13) that

$$(4.2) \qquad f(u+a) = f_2(a) k(m, \nu+n-m) \{\lambda + a\}^{\frac{1}{2}(\nu+n-m)} \{\lambda + a + u\}^{-\frac{1}{2}(\nu+n)},$$

where the constant $k(.,.)$ is as defined in (2.7). From (4.2), we may write

$$(4.3) \qquad f(v) = f_2(a) k(m, \nu+n-m) \{\lambda + a\}^{\frac{1}{2}(\nu+n-m)} \{\lambda + v\}^{-\frac{1}{2}(\nu+n)}, \quad v \geq 0.$$

Morever, by taking $v = a$ in (4.3), it follows that

$$f_2(a) = \frac{f(a)\{\lambda + a\}^{\frac{1}{2}m}}{k(m, \nu+n-m)},$$

which may be replaced in (4.3) yielding

$$(4.4) \qquad f(v) = f(a) \{\lambda + a\}^{\frac{1}{2}(\nu+n)} \{\lambda + v\}^{-\frac{1}{2}(\nu+n)}, \quad v \geq 0.$$

Now, since $a = ||\mathbf{x}_2||^2$ is fixed and $f(\mathbf{x}'\mathbf{x})$, $\mathbf{x} \in \mathcal{R}^n$ is the density of $\mathbf{X}$, it follows from (4.4) that

$$1 = \int_{\mathcal{R}^n} f(\mathbf{x}'\mathbf{x}) d\mathbf{x} = f(a)\{\lambda + a\}^{\frac{1}{2}(\nu+n)} \int_{\mathcal{R}^n} \{\lambda + \mathbf{x}'\mathbf{x}\}^{-\frac{1}{2}(\nu+n)} d\mathbf{x},$$

from where we can write

$$f(a)\{\lambda + a\}^{\frac{1}{2}(\nu+n)} = \{\int_{\mathcal{R}^n} \{\lambda + \mathbf{x}'\mathbf{x}\}^{-\frac{1}{2}(\nu+n)} d\mathbf{x}\}^{-1}$$

$$= \frac{\Gamma[\frac{1}{2}(\nu+n)]}{\Gamma[\frac{1}{2}\nu]\pi^{\frac{1}{2}n}} \lambda^{\frac{1}{2}\nu} = k(n, \nu)\lambda^{\frac{1}{2}\nu}.$$

Thus,

$$f(v) = k(n,\nu)\lambda^{\frac{1}{2}\nu}\{\lambda + v\}^{-\frac{1}{2}(\nu+n)}, \quad v \geq 0,$$

showing that $\mathbf{X} \sim t_n(\mathbf{0}, \mathbf{I}_n; \lambda, \nu)$, which concludes the proof.

5. A characterization in terms of quadratic forms

We now procede to a characterization of the generalized t-distribution within the class of the spherical distributions in terms of quadratic forms. To prove this result, the following preliminary lemmas will be required.

Lemma 1. *Let* $\mathbf{Y} \sim E_p(\mathbf{0}, \mathbf{I}_p; \phi)$ *and* $\mathbf{\Delta} = diag(\lambda_1, \ldots, \lambda_p)$, *a diagonal pxp matrix with diagonal elements* $\lambda_1 \geq \ldots \geq \lambda_p > 0$. *Suppose that* $\mathbf{Y}'\mathbf{\Delta}\mathbf{Y} \sim \lambda BeII(q/2, \nu/2)$*(or, equivalente–*

ly, $(q\lambda/\nu)F_{q,\nu})$. *Then,* $p = q$ *and* $\mathbf{Y}$ *has density* $f(\mathbf{y}'\mathbf{y})$, $\mathbf{y} \in \mathcal{R}^p$, *that is,* $\mathbf{Y} \sim E_n(\mathbf{0}, \mathbf{I}_p; f)$, *with* $f(u)$, $u \geq 0$ *(*$f(.)$ *being infinetely differentiable).*

Proof. Initially we point out the fact that $P[\mathbf{Y} = \mathbf{0}] = 0$, since $\mathbf{\Delta} > 0$ and $\mathbf{Y}'\mathbf{\Delta}\mathbf{Y} \sim \lambda BeII(q/2, \nu/2)$. Now, as $\mathbf{Y} \sim E_p(\mathbf{0}, \mathbf{I}_p; \phi)$ and $P[\mathbf{Y} = \mathbf{0}] = 0$, we have that $S = ||\mathbf{Y}||^2$ and $\mathbf{U} = \mathbf{Y}/S^{1/2}$ are independent (see Section 2.2). Let F_S and F_T be the distribution functions of S and $T = \mathbf{U}'\mathbf{\Delta}\mathbf{U}$, respectively. Note that $\lambda_1 \geq T \geq \lambda_p > 0$,with probability one. Since the distribution of $ST = \mathbf{Y}'\mathbf{\Delta}\mathbf{Y}$ is $(q\lambda/\nu)F_{q,\nu}$ (or equivalentely, $\lambda BeII(q/2, \nu/2)$), we have that

$$\int_{\lambda_1}^{\lambda_p} F_S(u/t)dF_T(t) = \int_0^u \frac{1}{\lambda B(\frac{1}{2}q, \frac{1}{2}\nu)}\{\frac{v}{\lambda}\}^{\frac{1}{2}q-1}\{1 + \frac{v}{\lambda}\}^{-\frac{1}{2}(q+\nu)}dv, \tag{5.1}$$

where

$$B(\frac{1}{2}q, \frac{1}{2}\nu) = \frac{\Gamma[\frac{1}{2}(q+\nu)]}{\Gamma[\frac{1}{2}q]\Gamma[\frac{1}{2}\nu]}.$$

Note that the right hand side of (5.1) is infinitely diferentiable with respect to u so that the left hand side must also be infinitely differentiable with respect to u. Moreover, since the domain of integration of the left hand side of (5.1) is finite, it is obvious that $F_S(u|t)$ has to be infinitely differentiable with respect to u. This shows that the density function of $\mathbf{Y}$ exists and must be infinitely differentiable. If the density of $\mathbf{Y}$ is $f(\mathbf{y}'\mathbf{y})$, $\mathbf{y} \in \mathcal{R}^p$, then the density of S is given by (see (2.2))

$$p_S(s) = \frac{\pi^{\frac{1}{2}p}}{\Gamma[\frac{1}{2}p]}s^{\frac{1}{2}p-1}f(s).$$

Thus, from (5.1), it follows that

$$\int_{\lambda_1}^{\lambda_p} \frac{2\pi^{\frac{1}{2}p}}{\Gamma[\frac{1}{2}p]}u^{\frac{1}{2}p-1}f(u|t)t^{-\frac{1}{2}p}dF_T(t) =$$

$$= \frac{1}{\lambda B(\frac{1}{2}q, \frac{1}{2}\nu)}\{\frac{u}{\lambda}\}^{\frac{1}{2}q-1}\{1 + \frac{u}{\lambda}\}^{-\frac{1}{2}(q+\nu)},$$

for all $u > 0$. Thus, if $q > p$, then

$$(5.2) \qquad \int_{\lambda_p}^{\lambda_1} f(u/t)t^{-\frac{1}{2}p}dF_T(t) = \frac{\Gamma[\frac{1}{2}(q+\nu)]}{\Gamma[\frac{1}{2}\nu]\pi^{\frac{1}{2}p}}\lambda^{-\frac{1}{2}p}\{\frac{u}{\lambda}\}^{\frac{1}{2}(q-p)}\{1+\frac{u}{\lambda}\}^{-\frac{1}{2}(q+\nu)},$$

for all $u > 0$. Note that by taking $u = 0$ in (5.2) we have that

$$f(0)E[T^{-\frac{1}{2}p}] = 0,$$

what is impossible, since it contradicts $f(0)E[T^{-\frac{1}{2}}] > 0$. Thus (5.2) does not hold for $u \to 0$ so we have that $q \leq p$. In a similar fashion, it can be shown that $q < p$ is also impossible. Thus, we must have $q = p$, which proves the lemma.

A similar lemma can be found in Khatri and Mukerjee (1987) and is used to prove a similar characterization for the normal distribution, under the hypothesis that $\mathbf{Y}'\mathbf{\Delta}\mathbf{Y} \sim \chi^2_q$.

Suppose now that $\mathbf{X} \stackrel{d}{=} R\mathbf{U}^{(n)} \sim E_n(\mathbf{0}, \mathbf{I}_n; \phi)$, where R and $\mathbf{U}^{(n)}$ are as defined in (2.1). As before, consider the partition $\mathbf{X} = (\mathbf{X}_1', \mathbf{X}_2')'$, where $\mathbf{X}_1$ is mx1. Cambanis et al. (1981) pointed out the fact that $(\mathbf{X}_1', \mathbf{X}_2')' \stackrel{d}{=} (R\sqrt{V_1}\mathbf{U}^{(m)\prime}, R\sqrt{V_2}\mathbf{U}^{(m-n)\prime})'$, where $V_1 \geq 0$, $V_1 + V_2 = 1$; $(V_1, V_2)'$, R, $\mathbf{U}^{(m)}$ and $\mathbf{U}^{(n-m)}$ are all independent and $V_1 \sim Be(m/2, (n-m)/2)$ (see also Anderson and Fang, 1987). Thus, $||\mathbf{X}_1||^2 \stackrel{d}{=} R^2V_1$, with a distribution denoted by $G(m/2, (n-m)/2; \phi)$ by Anderson and Fang (1987). With this notation, Anderson and Fang (1987) prove the following result.

Lemma 2. *Suppose that* $\mathbf{X} \stackrel{d}{=} R\mathbf{U}^{(n)} \sim E_n(\mathbf{0}, \mathbf{I}_n; \phi)$, $P[\mathbf{X} = \mathbf{0}] = 0$ *and* $\mathbf{A}$ *is an nxn symmetric matrix. Then,* $\mathbf{X}'\mathbf{A}\mathbf{X} \sim G(m/2, (n-m)/2; \phi)$ *if and only if* $\mathbf{A}^2 = \mathbf{A}$ *and* $rank(\mathbf{A}) = m$. The main result of this section is presented and proved next.

Theorem 3. *Suppose that* $\mathbf{X} \sim E_n(\mathbf{0}, \mathbf{I}_n; \phi)$ *and let* $\mathbf{A}$ *be a symmetric nxn matrix. Then* $\mathbf{X}'\mathbf{A}\mathbf{X} \sim \lambda BeII(m/2, (n-m)/2)$ *(or, equivalentely,* $(m\lambda/\nu)F_{m,\nu}$*) if and only if* $\mathbf{X} \sim t_n(\mathbf{0}, \mathbf{I}_n; \lambda, \nu)$, $\mathbf{A}^2 = \mathbf{A}$ *and* $rank(\mathbf{A}) = m$.

Proof. For the part "only if" see Section 2.2. We prove now the converse. Note that $\mathbf{x}'\mathbf{A}\mathbf{x} \geq 0$ for all $\mathbf{x} \in \mathcal{R}^n$, since $\mathbf{X}'\mathbf{A}\mathbf{X} \sim \lambda BeII(m\lambda/\nu)$ (or equivalentely, $(m\lambda/\nu)F_{m,\nu}$). Let $k = rank(\mathbf{A})$ and let $\mathbf{\Gamma}$ be an nxk matrix such that

$$\mathbf{A} = \mathbf{\Gamma}\mathbf{\Delta}\mathbf{\Gamma}', \quad \text{and} \quad \mathbf{\Gamma}'\mathbf{\Gamma} = \mathbf{I}_k,$$

where $\mathbf{\Delta} = diag(\lambda_1, \ldots, \lambda_k)$, with $\lambda_1 \geq \ldots \lambda_k > 0$. Let $\mathbf{Y} = \mathbf{\Gamma}'\mathbf{X}$. Thus, $\mathbf{Y} \sim E_k(\mathbf{0}, \mathbf{I}_k; \phi)$ and $\mathbf{Y}'\mathbf{\Delta}\mathbf{Y} \stackrel{d}{=} \mathbf{X}'\mathbf{A}\mathbf{X} \sim (m\lambda/\nu)F_{m,\nu}$. Using Lemma 1, we have that $\mathbf{Y} \sim E_k(\mathbf{0}, \mathbf{I}_k; f_1)$, that is, $\mathbf{Y}$ has density $f_1(\mathbf{y}'\mathbf{y})$, $\mathbf{y} \in \mathcal{R}^k$. Since $P(\mathbf{Y} = \mathbf{0}) = 0$ and $\mathbf{Y}'\mathbf{\Delta}\mathbf{Y} \sim (m\lambda/\nu)F_{m,\nu}$ we have that (see Remark 2)

$$\mathbf{Y}'\mathbf{\Delta}\mathbf{Y} \stackrel{d}{=} R_1^2V_1,$$

where

$$R_1 \stackrel{d}{=} ||\mathbf{Y}||$$

and $V_1 \sim Beta(m/2,(k-m)/2)$ are independent. From Lemma 2, we have that $\mathbf{\Delta}^2 = \mathbf{\Delta}$ and $rank(\mathbf{\Delta}) = m$, that is, $k = m$, $\mathbf{\Delta} = \mathbf{I}_m$ and $V_1 = 1$. Hence,

$$\mathbf{Y}'\mathbf{\Delta}\mathbf{Y} = ||\mathbf{Y}||^2 \stackrel{d}{=} R_1^2 \sim (m\lambda/\nu)F_{m,\nu}.$$

This shows that

$$\mathbf{Y} = ||\mathbf{Y}||\frac{\mathbf{Y}}{||\mathbf{Y}||} \stackrel{d}{=} R_1^2\mathbf{U}^{(m)} \sim t_m(\mathbf{0},\mathbf{I}_m;\lambda,\nu),$$

is a marginal density of $\mathbf{X} \sim E_n(\mathbf{0},\mathbf{I}_n;\phi)$. Since m is arbitrary, Theorem 1 above shows that $\mathbf{X} \sim t_n(\mathbf{0},\mathbf{I}_n;\lambda,\nu)$, which concludes the proof.

An alternative proof can be obtained by following the proof given in Khatri and Mukherjee (1987) to the characterization of the normal distribution using quaratic forms. Their approach is based on a Taylor series expansion of the density function.

We believe that similar results hold for any distribution that is scale mixture of the normal distribution. In fact, in this class the corresponding conditional distribution given the scale parameter is normal. Accordingly, the characterizations derived for the normal distribution also applies in this case.

6. A characterization of the t-distribution within the class of the compound normal distributions

As we have seen in Section 2, an important subclass of the elliptical distributions is given by the compound normal distributions, which are mixture in a scale parameter of the normal distribution. In Section 2 it is considered that the scale parameter follows the $IG(\nu/2,\lambda/2)$ so that the t-distribution follows. Thus, any distribution in this class is determined once the measure of the mixture is obtained. The question we try to answer in this section is *under what conditions on the observations we can identify the measure of the mixture.* In this sence, Diaconis and Ylvisaker (1979,1985) provide some results in the context of the regular exponential family of distributions with the natural parametrization. Their main result is stated next in the uniparameter case.

Considering the natural parametrization, the probability distribution of a regular exponential family is defined by

$$P_\theta(dy) = e^{\theta y - K(\theta)}m(dy), \quad \theta \in \Theta, \tag{6.1}$$

where the parameter space Θ is convex and open in $\mathcal{R}$ and $m(dy)$ is and absolutely continuous measure with respect to the Lebesgue measure. Let $\mathbf{\Theta}$ denote the random variable taking values θ in Θ. Given $\mathbf{\Theta} = \theta$, the mean value of Y is the derivative of $K(\theta)$, that is, $K'(\theta)$. The mean value of Y is given by

$$E[Y] = E[K'(\mathbf{\Theta})],$$

and the posterior mean is $E[K'(\mathbf{\Theta}|\mathbf{Y})]$. To obtain the results that we present next, we restrict ourselves to the class of the nondegenerate prior distributions for which $E[Y]$ and

$E[K'(\Theta)|\mathbf{Y}]$ are well defined. Diaconis and Ylvisaker (1979) prove the following result concerning the family (6.1), where it is established that the linearity condition in the posterior mean implies that the prior distribuiton must be conjugated.

Lemma 3. (Diaconis and Ylvisaker, 1979) *Suppose that Θ is open in $\mathcal{R}$. Let Y_1 be a sample of size $n = 1$ from P_θ of (3.1) and suppose that the support ofthe measure m contains an open interval in $\mathcal{R}$. If Θ has a priori distribution $\pi(.)$ which does not concentrate at a single point and if*

(6.2)
$$E[K'(\Theta)|Y_1] = b + aY_1,$$

for some constants a and b, then $a \neq 0$, π is absolutely continuous with respect to the Lebesgue measure and has density

(6.3)
$$\pi(d\theta) = ce^{(b/a)\theta - [(1-a)/a]M(\theta)}d\theta.$$

The theorem that we consider next presents another characterization of the t-distribution.

Theorem 4. *Let $X_1, \ldots, X_n$ and V be random variables such that given $V = v$, $X_1, \ldots, X_n$ are independent and identically distributed (iid) $N(0, v)$, where V has distribution F_V over $[0, \infty)$, which means that the joint distribution of $\mathbf{X} = (X_1, \ldots, X_n)'$ is compound normal (see (2.3) and (2.4)). If $P[X_1 = 0] = 0$ and*

(6.4)
$$E[X_2^2|X_1] = b + aX_1^2, \quad 0 < a < 1, \quad b > 0,$$

then $\mathbf{X} = (X_1, \ldots, X_n)' \sim t_n(\mathbf{0}, \mathbf{I}_n; b/a, (a+1)/a)$. The converse also holds.

Proof. Note that $P[X_1 = 0] = 0$ implies that $V > 0$ with probability one. let $Y_i = X_i^2$, $i = 1, \ldots, n$. Then, given $V = v$, $Y_1, \ldots, Y_n$ are iid $Ga(1/2, 1/(2v))$. Thus, the conditional distribution of Y_1 given $V = v$ is of the form (6.1) with natural parameter $\theta = -1/(2v)$,

(6.5)
$$K(\theta) = -\frac{1}{2}\log(-\theta), \quad -\infty < \theta < 0,$$

and

$$m(dy_1) = (\pi y_1)^{-\frac{1}{2}} I_{[0,\infty)}(y_1)dy_1,$$

where $I_{\mathbf{A}}(x)$ is 1 if $x \in \mathbf{A}$ and 0 otherwise. Then, let $\Theta = -1/(2V)$. Using (6.4) and some well known properties of conditional expectations, we have that

$$E[E(Y_1|\Theta)|Y_1] = E[E(Y_1|V)|Y_1]$$
$$= E[E(Y_2|V)|Y_1]$$
$$= E[E(Y_2|V, Y_1)|Y_1]$$
$$= E[Y_2|Y_1] = E[E(Y_2|X_1)|Y_1] = b + aY_1.$$

From Lemma 6.1 it follows that the distribution of Θ is of the form (6.3) with $K(\theta)$ given by (6.5), that is, $\Theta \sim Ga((a+1)/(2a), b/(2a))$. Accordingly, $V = -1/(2\Theta) \sim IG((a+1)/(2a), b/a)$, form where the result follows. The reverse statement follows from (2.15) with $\mu = 0$, $\Sigma = \mathbf{I}_n$, $\lambda = b/a$, $\nu = (a+1)/a$ and by taking $n - m = m = 1$.

Note that (6.4) can be expressed as

$$E[V|X_1] = b + aX_1^2.$$

References

Anderson and Fang, (1987). Cochran's theorem for elliptically contourned distribution. *Shankya, A*, 49, 305-315.

Andrews, D.F. and Mallows, C.L. (1974). Scale mixtures of normal distributions. *Journal of the Royal Statistical Society*, B, 36, 99-102.

Cambanis, Huang, and Simons, G. (1981). On the theory of elliptically contourned distributions. *Journal of Multivariate Analysis*, 11, 368-385.

Diaconis, P. and Ylvisaker, D. (1979). Conjugate priors for exponential families. *The Ann. of Statistic.*, 7, 269-281.

Diaconis, P. and Ylvisaker, D. (1985). Quantifying prior opinion. In Bayesian Statistics 2. Elsevier Science Publisher B.V. and Valencia University Press.

Dickey, J.M. (1967). Matric-variate generalization of the multivariate t-distribution and the inverted multivariate t-distribution. *Ann. Math. Statist.*, 38, 511-518.

Fang,K.T., Kotz, S. and Ng, K.W. (1990). *Symmetric multivariate and related distributions.* Chapman and Hall. London.

Khatri, C.G. and Mukerjee, R. (1987). Characterizations of normality within the class of elliptical contourned distributions. *Statistics and Probability letters*, 5, 187-190.

Kelker, D. (1970). Distribution theory of spherical distributions and location scale parameters. *Shankya*, A, 32, 419-430.

Muirhead, R.J. (1982). *Aspects of multivariate statistical theory.* Wiley, N.Y.

ULTIMOS RELATORIOS TECNICOS PUBLICADOS

9401 - PAULA, G. A., SEN P.K., Tests of ordered Hypotheses in Linkage in Heredity, 12p.

9402 - PAULA, G. A., SEN P. K. One-Sided Tests in Generalized Linear Models with Parallel Regression Lines, 18p.

9403 - PAULA, G. A., .Influence Diagnostics in Proper Dispersion Models,22 p.

9404 - SALINAS-TORRES, V.H. & PEREIRA, C.A.B. and TIWARI, R.C. - Convergence of Dirichlet Measures Arising in Context of Bayesian Analysis of Competing Risk Models, 17p.

9405 - CRIBARI-NETO, F. & FERRARI, S.L.P. - Second Order Asymptotics for Score Tests in Generalized Linear Models, 25p.

9406 - ZUAZOLA, P.L.I., PEREIRA, C.A.B. & WECHSLER, S. - Predictivistic Statistical Inference in Finite Populations, 10p.

9407 - BOLFARINE, H. & ROJAS, M.G. - Maximum Likelihood Estimation of Simultaneous Pairwise Linear Structural Relationships, 19p.

9408 - FERRARI, P.A. & KIPNIS, C. - Second Class Particles in the Rarefaction Fan, 9p.

9409 - ANDRÉ, C.D.S., NARULA, S.C. & PERES, C.A. - Asymptotic Properties of MSAE Estimates in Multistage Model. Preliminary Results, 19p.

The complete list of Relatórios do Departamento de Estatística, IME-USP, will be sent upon request.

- Departamento de Estatística
IME-USP
Caixa Postal 20.570
01452-990 - São Paulo, Brasil