Moments of the truncated multivariate-t distribution

by A. O'Hagan

(University of Warwick)*

*Work done while the author was at the University of Dundee

SUMMARY

The first and second moments of a multivariate-t distribution truncated in some or all of its variates are expressed in terms of the probability integral of the untruncated distribution. This considerably reduces the amount of computation required to calculate these moments.

1. INTRODUCTION

The multivariate-t distribution arises regularly in the Bayesian analysis of the linear model with a single unknown variance. Introducing inequality constraints on the parameters results in their prior and posterior distributions each being a truncated multivariate-t: a particular example is described in O'Hagan (1973). In problems concerning the treatment of outliers or the selection of a population with the highest mean we are often interested in the distribution of the maximum of a set of multivariate-t distributed variables: Afonja (1972) gives expressions relating the moments of this distribution to those of the truncated multivariate-t.

The amount of computation required to find moments of the truncated multivariate-t, useful in both the above applications, is considerably reduced by the formulae derived in this paper. The moments are expressed in terms of the distribution function of the untruncated distribution, for which a highly efficient algorithm already exists: Dutt (1975). In addition the dimensionality of integration is reduced. Without these formulæ the computation of the mean vector and covariance matrix of a p-dimensional distribution would require $1+p+\frac{1}{2}p(p+1)$ integrations in p dimensions, whereas with the use of equations (14) and (16) this is reduced to 2 integrations in p dimensions, p in (p-1) dimensions and $\frac{1}{2}p(p-1)$ in (p-2) dimensions.

We consider in this paper truncation on the right of each of the variables x_i in a p-variate t distribution, i.e. $-\infty<x_i\leq u_i$ $(i = 1,2,..,p)$. Since any linear transformation of the variables also has a multivariate-t distribution, results for truncation on the left or more general planar truncation are easily derived. In some cases it may be necessary to add and subtract integrals obtained from these results, e.g. when a variable

is truncated on both the left and the right. The linear transformation property also implies that we can eliminate location parameters, assuming zero means in the untruncated distribution. In Section 3 we derive a relationship between moments which is used in Section 4 as a reduction formula to obtain the mean and covariance matrix. These are then compared in Section 5 with known results in some special or limiting cases.

2. DEFINITIONS AND NOTATION

A $p\times 1$ vector random variable $\underset{\sim}{x}$ has a multivariate-t distribution with ν degrees of freedom, zero mean and scale matrix $\underset{\sim}{\Sigma}$ if its density function is

$$t_p(\underset{\sim}{x}|\nu,\underset{\sim}{\Sigma}) = \frac{\nu^{\frac{1}{2}\nu}}{\pi^{\frac{1}{2}p}|\underset{\sim}{\Sigma}|^{\frac{1}{2}}} \frac{\Gamma\{\frac{1}{2}(\nu+p)\}}{\Gamma(\frac{1}{2}\nu)} (\nu + \underset{\sim}{x}' \underset{\sim}{\Sigma}^{-1} \underset{\sim}{x})^{-\frac{1}{2}(\nu+p)} .$$

Its covariance matrix is in fact $\{\nu/(\nu-2)\}\underset{\sim}{\Sigma}$. We may equivalently regard $\underset{\sim}{x}$ as having conditionally a Normal distribution with zero mean and covariance matrix $h^{-1}\underset{\sim}{\Sigma}$, given h, where νh has a χ^2 distribution with ν degrees of freedom: although we use this representation below, we cannot simply integrate the moments of the truncated Normal distribution (see e.g. Tallis (1961)) with respect to h, for the truncation affects the distribution of h, see section 5(e). Truncating $\underset{\sim}{x}$ on the right by the $p\times 1$ vector $\underset{\sim}{u}$ results in $\underset{\sim}{x}$ having the density

$$f_p(\underset{\sim}{x}|\nu,\underset{\sim}{\Sigma}) = F_p(\underset{\sim}{u}|\nu,\underset{\sim}{\Sigma})^{-1}t_p(\underset{\sim}{x}|\nu,\underset{\sim}{\Sigma}), \tag{1}$$

for $-\infty < x_i \leq u_i$ $(i = 1,\ldots,p)$, and zero elsewhere. F_p is the distribution function of the multivariate-t distribution,

$$F_p(\underset{\sim}{u}|\nu,\underset{\sim}{\Sigma}) = \int_{-\underset{\sim}{\infty}}^{\underset{\sim}{u}} t_p(\underset{\sim}{x}|\nu,\underset{\sim}{\Sigma})d\underset{\sim}{x}.$$

We will be concerned with moments of the truncated distribution (1), and we write the expectation of a general function $w(\underset{\sim}{x})$ as

$$E_p\{w(\underset{\sim}{x})|\nu,\underset{\sim}{\Sigma},\underset{\sim}{u}\} = \int_{-\underset{\sim}{\infty}}^{\underset{\sim}{u}} w(\underset{\sim}{x})f_p(\underset{\sim}{x}|\nu,\underset{\sim}{\Sigma})d\underset{\sim}{x}. \tag{2}$$

The representation of the t distribution in terms of the Normal and χ^2 distributions enables us to write (2) as

$$E_p\{w(\underset{\sim}{x})|\nu,\underset{\sim}{\Sigma},\underset{\sim}{u}\} = G_p(\underset{\sim}{u},\nu,\underset{\sim}{\Sigma})^{-1}\int_{-\underset{\sim}{\infty}}^{\underset{\sim}{u}} w(\underset{\sim}{x})\int_0^{\infty} h^{\frac{1}{2}(\nu+p)-1}\exp\{-\tfrac{1}{2}h(\nu+\underset{\sim}{x}'\underset{\sim}{\Sigma}^{-1}\underset{\sim}{x})\}d\underset{\sim}{x}dh, \tag{3}$$

where

$$G_p(\underset{\sim}{u},\nu,\underset{\sim}{\Sigma}) = \int_{-\underset{\sim}{\infty}}^{\underset{\sim}{u}}\int_0^{\infty} h^{\frac{1}{2}(\nu+p)-1}\exp\{-\tfrac{1}{2}h(\nu+\underset{\sim}{x}'\underset{\sim}{\Sigma}^{-1}\underset{\sim}{x})\}d\underset{\sim}{x}dh$$

$$= \pi^{\frac{1}{2}p}|\underset{\sim}{\Sigma}|^{\frac{1}{2}}2^{\frac{1}{2}(\nu+p)}\Gamma(\tfrac{1}{2}\nu)\nu^{-\frac{1}{2}\nu}F_p(\underset{\sim}{u}|\nu,\underset{\sim}{\Sigma}). \tag{4}$$

For manipulating quadratic forms it is necessary to introduce some rather complicated notation. We write $\underset{\sim}{S}$ for the inverse of $\underset{\sim}{\Sigma}$; x_i is the i-th element of a vector $\underset{\sim}{x}$, and x_{ij} is the (i,j)-th element of a matrix $\underset{\sim}{X}$, as usual. In addition we define:

(a) $\underset{\sim}{x}_{-i}$ to be the rest of $\underset{\sim}{x}$ after removing x_i,

(b) $\underset{\sim}{x}_{-ij}$ to be the rest of $\underset{\sim}{x}$ after removing x_i and x_j,

(c) $\underset{\sim}{X}_{-i}$ to be the rest of $\underset{\sim}{X}$ after removing the i-th row and column,

(d) $\underset{\sim}{X}_{-ij}$ to be the rest of $\underset{\sim}{X}$ after removing the i-th and j-th rows and columns,

(e) $\underset{\sim}{X}_{(i)}$ to be the rest of the i-th column of $\underset{\sim}{X}$ after removing x_{ii},

(f) $\underset{\sim}{X}_{(ij)}$ to be the rest of the i-th and j-th columns of $\underset{\sim}{X}$ after removing x_{ii}, x_{ij}, x_{ji} and x_{jj} (a $(n-2)\times 2$ matrix if $\underset{\sim}{X}$ is $n\times m$).

For example

$$\underset{\sim}{X} = \left(\begin{array}{cc:c} x_{11} & x_{12} & \underset{\sim}{X}_{(12)}' \\ x_{21} & x_{22} & \\ \hdashline \multicolumn{2}{c:}{\underset{\sim}{X}_{(12)}} & \underset{\sim}{X}_{-12} \end{array}\right) .$$

We will require the standard identities

$$\underset{\sim}{x}'\underset{\sim}{S}\underset{\sim}{x} = s_{ii}x_i^2 + 2x_i\underset{\sim}{S}_{(i)}'\underset{\sim}{x}_{-i} + \underset{\sim}{x}_{-i}'\underset{\sim}{S}_{-i}\underset{\sim}{x}_{-i}$$

$$= \sigma_{ii}^{-1}x_i^2 + (\underset{\sim}{x}_{-i} - x_i\sigma_{ii}^{-1}\underset{\sim}{\Sigma}_{(i)})'\underset{\sim}{S}_{-i}(\underset{\sim}{x}_{-i} - x_i\sigma_{ii}^{-1}\underset{\sim}{\Sigma}_{(i)}) \quad (5)$$

and $$\underset{\sim}{S}_{-i}^{-1} = \underset{\sim}{\Sigma}_{-i} - \sigma_{ii}^{-1}\,\underset{\sim}{\Sigma}_{(i)}\underset{\sim}{\Sigma}_{(i)}' . \quad (6).$$

3. THE REDUCTION FORMULA

We now derive a reduction formula similar to that obtained by Birnbaum and Meyer (1953) for the truncated Normal distribution. Equation (3) may be written

$$s_{ii}G_p(\mathbf{u},\nu,\Sigma)E_p\{w(\mathbf{x})\,|\,\nu,\Sigma,\mathbf{u}\} = s_{ii}\int_{-\infty}^{\mathbf{u}}\int_0^{\infty} w(\mathbf{x})h^{\frac{1}{2}(\nu+p)-1}\exp\{-\tfrac{1}{2}h(\nu+\mathbf{x}'\Sigma^{-1}\mathbf{x})\}\,d\mathbf{x}\,dh$$

$$= \int_{-\infty}^{\mathbf{u}_{-i}}\int_0^{\infty} h^{\frac{1}{2}(\nu+p)-2}\exp\{-\tfrac{1}{2}h(\nu+\mathbf{x}'_{-i}\mathbf{S}_{-i}\mathbf{x}_{-i})\}$$

$$\int_{-\infty}^{u_i}\{s_{ii}hx_i\exp(-\tfrac{1}{2}hs_{ii}x_i^2)\}\{x_i^{-1}w(\mathbf{x})\exp(-hx_i\mathbf{S}'_{(i)}\mathbf{x}_{-i})\}\,d\mathbf{x}_{-i}\,dh\,dx_i\,,$$

and using

$$\int s_{ii}hx_i\exp(-\tfrac{1}{2}hs_{ii}x_i^2)\,dx_i = -\exp(-\tfrac{1}{2}hs_{ii}x_i^2)$$

we integrate by parts to obtain

$$s_{ii}G_p(\mathbf{u},\nu,\Sigma)E_p\{w(\mathbf{x})\,|\,\nu,\Sigma,\mathbf{u}\} = E_1-E_2, \tag{7}$$

where

$$E_1 = \int_{-\infty}^{\mathbf{u}_{-i}}\int_0^{\infty} h^{\frac{1}{2}(\nu+p)-2}\exp\{-\tfrac{1}{2}h(\nu+\mathbf{x}'_{-i}\mathbf{S}_{-i}\mathbf{x}_{-i})\}$$

$$\left[-\exp(-\tfrac{1}{2}hs_{ii}x_i^2)\{x_i^{-1}w(\mathbf{x})\exp(-hx_i\mathbf{S}'_{(i)}\mathbf{x}_{-i})\}\right]_{-\infty}^{u_i} d\mathbf{x}_{-i}\,dh\,,$$

$$E_2 = \int_{-\infty}^{\mathbf{u}_{-i}}\int_0^{\infty} h^{\frac{1}{2}(\nu+p)-2}\exp\{-\tfrac{1}{2}h(\nu+\mathbf{x}'_{-i}\mathbf{S}_{-i}\mathbf{x}_{-i})\}\int_{-\infty}^{u_i}\{-\exp(-\tfrac{1}{2}hs_{ii}x_i^2)\}\times$$

$$\times\left\{\frac{d}{dx_i}\,(x_i^{-1}w(\mathbf{x}))-h(\mathbf{S}'_{(i)}\mathbf{x}_{-i})x_i^{-1}w(\mathbf{x})\right\}\exp(-hx_i\mathbf{S}'_{(i)}\mathbf{x}_{-i})\,d\mathbf{x}_{-i}\,dh\,dx_i\,.$$

To evaluate E_1 we assume that the value of the expression in square brackets is zero at $x_i = -\infty$. This will certainly be true if $w(\mathbf{x})$ is a polynomial in x_i, as will be required in Section 4, and for a wide range of other functions. Writing

$$w(\mathbf{x}_{-i},u_i) \equiv w(\mathbf{x})\big|_{x_i = u_i},$$

then

$$E_1 = -\int_{-\boldsymbol{\infty}}^{\mathbf{u}_{-i}}\int_0^{\infty} u_i^{-1} w(\mathbf{x}_{-i}, u_i) h^{\frac{1}{2}(\nu+p)-2} \exp\{-\tfrac{1}{2}h(\nu + s_{ii}u_i^2 + 2u_i \mathbf{S}'_{(i)}\mathbf{x}_{-i} + \mathbf{x}'_{-i}\mathbf{S}_{-i}\mathbf{x}_{-i})\}\, d\mathbf{x}_{-i}\, dh$$

$$= -\int_{-\boldsymbol{\infty}}^{\mathbf{u}_{-i}}\int_0^{\infty} u_i^{-1} w(\mathbf{x}_{-i}, u_i) h^{\frac{1}{2}(\nu+p)-2}$$

$$\exp[-\tfrac{1}{2}h\{\nu + \sigma_{ii}^{-1}u_i^2 + (\mathbf{x}_{-i} - u_i\sigma_{ii}^{-1}\boldsymbol{\Sigma}_{(i)})'\mathbf{S}_{-i}(\mathbf{x}_{-i} - u_i\sigma_{ii}^{-1}\boldsymbol{\Sigma}_{(i)})\}]\, d\mathbf{x}_{-i}\, dh$$

using (5). By a simple transformation of variables we may write this in the form (3) so that

$$E_1 = -u_i^{-1}\left(\frac{\nu-1}{\nu+\sigma_{ii}^{-1}u_i^2}\right)^{\frac{1}{2}(\nu+p)-1} G_{p-1}\{\mathbf{u}(i), \nu-1, \boldsymbol{\Sigma}_{-i}^*\}$$

$$E_{p-1}\{w(\mathbf{x}_{-i} + u_i\sigma_{ii}^{-1}\boldsymbol{\Sigma}_{(i)}, u_i) \mid \nu-1, \boldsymbol{\Sigma}_{-i}^*, \mathbf{u}(i)\}, \quad (8)$$

where

$$\mathbf{u}(i) = \mathbf{u}_{-i} - u_i\sigma_{ii}^{-1}\boldsymbol{\Sigma}_{(i)},\quad \boldsymbol{\Sigma}_{-i}^* = \frac{\nu+\sigma_{ii}^{-1}u_i^2}{\nu-1}(\boldsymbol{\Sigma}_{-i} - \sigma_{ii}^{-1}\boldsymbol{\Sigma}_{(i)}\boldsymbol{\Sigma}'_{(i)})$$

using (6). Turning now to E_2 we find that it is composed of two terms:

$$E_2 = E_3 + E_4, \quad (9)$$

where

$$E_3 = -\int_{-\boldsymbol{\infty}}^{\mathbf{u}}\int_0^{\infty} \frac{d}{dx_i}\{x_i^{-1}w(\mathbf{x})\} h^{\frac{1}{2}(\nu+p)-2} \exp\{-\tfrac{1}{2}h(\nu + \mathbf{x}'\boldsymbol{\Sigma}^{-1}\mathbf{x})\}\, d\mathbf{x}\, dh$$

$$= -\{(\nu-2)/\nu\}^{\frac{1}{2}(\nu+p)-1} G_p(\mathbf{u}, \nu-2, \tfrac{\nu}{\nu-2}\boldsymbol{\Sigma}) E_p\{\tfrac{d}{dx_i}(x_i^{-1}w(\mathbf{x})) \mid \nu-2, \tfrac{\nu}{\nu-2}\boldsymbol{\Sigma}, \mathbf{u}\}, \quad (10)$$

$$E_4 = \int_{-\boldsymbol{\infty}}^{\mathbf{u}}\int_0^{\infty} \Big(\sum_{j\neq i} s_{ij}x_j\Big) x_i^{-1} w(\mathbf{x}) h^{\frac{1}{2}(\nu+p)-1} \exp\{-\tfrac{1}{2}h(\nu + \mathbf{x}'\boldsymbol{\Sigma}^{-1}\mathbf{x})\}\, d\mathbf{x}\, dh$$

$$= \sum_{j\neq i} s_{ij} G_p(\mathbf{u}, \nu, \boldsymbol{\Sigma}) E_p\{x_i^{-1}x_j w(\mathbf{x}) \mid \nu, \boldsymbol{\Sigma}, \mathbf{u}\}. \quad (11)$$

Collecting together (7), (8), (9), (10) and (11), and using (4) for each of the G functions, we eventually find that

$$\sum_{j=1}^{p} s_{ij} E_p \{x_i^{-1} x_j w(\underset{\sim}{x}) \mid \nu, \underset{\sim}{\Sigma}, \underset{\sim}{u}\}$$

$$= \frac{\nu}{\nu-2} \frac{F_p(\underset{\sim}{u}\mid\nu-2, \frac{\nu}{\nu-2}\underset{\sim}{\Sigma})}{F_p(\underset{\sim}{u}\mid\nu,\underset{\sim}{\Sigma})} E_p\{\frac{d}{dx_i}(x_i^{-1}w(\underset{\sim}{x}))\mid\nu-2, \frac{\nu}{\nu-2}\underset{\sim}{\Sigma}, \underset{\sim}{u}\}$$

$$- u_i^{-1}\xi_i \; E_{p-1}\{w(\underset{\sim}{x}_{-i} + u_i\sigma_{ii}^{-1}\underset{\sim}{\Sigma}_{(i)}, u_i)\mid\nu-1, \underset{\sim}{\Sigma}_{-i}^{*}, \underset{\sim}{u}_{(i)}\}, \tag{12}$$

where

$$\xi_i = \frac{1}{\sqrt{(2\pi\sigma_{ii})}} \cdot \left(\frac{\nu}{\nu+\sigma_{ii}^{-1}u_i^2}\right)^{\frac{1}{2}(\nu-1)} \cdot \frac{\Gamma\{\frac{1}{2}(\nu-1)\}\cdot(\frac{1}{2}\nu)}{\Gamma(\frac{1}{2}\nu)} \cdot \frac{F_{p-1}\{\underset{\sim}{u}(i)\mid\nu-1, \underset{\sim}{\Sigma}_{-i}^{*}\}}{F_p(\underset{\sim}{u}\mid\nu,\underset{\sim}{\Sigma})}. \tag{13}$$

Notice that the last term in (12) is an expectation with respect to the conditional multivariate-t distribution of $\underset{\sim}{x}_{-i}$ given $x_i = u_i$.

4. FIRST AND SECOND MOMENTS

We will now use (12) to obtain the mean vector and covariance matrix of $\underset{\sim}{x}$. We first suppose that $w(\underset{\sim}{x}) = x_i$, then

$$x_i^{-1}x_j w(\underset{\sim}{x}) = x_j, \quad \frac{d}{dx_i}(x_i^{-1}w(\underset{\sim}{x})) = 0, \quad w(\underset{\sim}{x}_{-i} + u_i\sigma_{ii}^{-1}\underset{\sim}{\Sigma}_{(i)}, u_i) = u_i.$$

Therefore equation (12) becomes, in this case,

$$\sum_{j=1}^{p} s_{ij} E_p(x_j \mid \nu, \underset{\sim}{\Sigma}, \underset{\sim}{u}) = -\xi_i.$$

Taking $w(\underset{\sim}{x})$ to be $x_1, x_2, \ldots, x_p$ successively the right-hand sides of the resulting equations form the vector $(-\underset{\sim}{\xi})$. The left hand sides form $\underset{\sim}{S}\underset{\sim}{E}_p(\underset{\sim}{x}\mid\nu,\underset{\sim}{\Sigma},\underset{\sim}{u})$, where $\underset{\sim}{E}_p(\underset{\sim}{x}\mid\nu,\underset{\sim}{\Sigma},\underset{\sim}{u})$ is the vector whose i-th element is $E_p(x_i\mid\nu,\underset{\sim}{\Sigma},\underset{\sim}{u})$, and is therefore the mean vector we require. Therefore

$$\underset{\sim}{E}_p(\underset{\sim}{x}\mid\nu,\underset{\sim}{\Sigma},\underset{\sim}{u}) = -\underset{\sim}{\Sigma}\underset{\sim}{\xi}, \tag{14}$$