The Korean Communications
in Statistics Vol. 4, No. 3, 1997
pp. 629 - 635

On an Approximation for Calculating Multivariate t Orthant Probabilities

Hea-Jung Kim[1)]

Abstract

An approximation for multivariate t probability for an orthant region(i.e., a rectangular region with lower limits of $-\infty$ for all margins) is proposed. It is based on conditional expectations, a regression with binary variables, and the exact formula for the evaluation of the bivariate t integrals by Dunnett and Sobel. It is noted that the proposed approximation method is especially useful for evaluating the multivariate t integrals where there is no simple method available until now.

1. Introduction

In multivariate analysis, many applications need to evaluate rectangle and orthant probabilities of the multivariate t-distribution. These include the multivariate t link model (e.g., Collett 1991), Bayesian analysis of the linear model (e.g., Press 1989), multivariate paired comparisons(e.g., Bökenholt 1992). In general the probabilities required to evaluate multidimensional integrals, which might be done by time consuming multidimensional quadrature method(cf. Genz 1992) or poor working Monte Carlo simulation. Thus one may prefer to use an approximation method if it's accuracy is good enough.

Earlier approximation methods for some special cases and a couple of exact methods are mentioned in Johnson and Kotz(1972). Dunnett and Sobel(1954) derived exact bivariate t orthant probability. For the general rectangular probabilities, John(1964) proposed an exact method for evaluating multivariate t probability integrals via inversion formula but it is too complicate to use. Therefore, several approximation methods are proposed for the specific types of the multivariate t variates(cf. Johnson and Kotz 1972), but they may not be applicable for calculating rectangle or orthant probabilities for general cases. In fact, Johnson and Kotz(1972) noted that there are no other tables of the probability integral or percentage points of the multivariate t-distribution.

In this paper we propose and develop an approximation method for the orthant probabilities which is similar to that used by Solow(1990) and Joe(1995) for evaluating multivariate normal

---

1) Professor, Department of Statistics, Dongguk University, Seoul, 100-715, Korea.

probability integrals. In Section 2 we extend the ideas of Solow(1990) and Joe(1995) to the case of multivariate t-distribution, and get approximations for the multivariate t integrals of orthant region (the rectangular region with lower limits of $-\infty$ for all margins). The extension of the ideas is quite simple. It consists of ascertaining conditional expectations of multivariate t-distribution to have the same form as those of multivariate normal distribution and constructing regression for binary variables designed to evaluate the integrals of orthant region. Numerical examples of the approximation for a specific case are given in Section 3, along with comparisons with some previously proposed approximations and the exact method.

2. Approximation Method

One generalization of the univariate Student t-distribution to the multivariate case was carried out by Cornish(1954), and Dunnett and Sobel(1954) (other generalizations are possible but this one appears to be the most useful in applications). Their definition of the multivariate t-distribution for $X=(X_1,\dots,X_p)'$ is as follows.

Definition 1. Let Y and u be independently distributed according to $N_p(0,\Sigma)$ and the χ^2_n -distribution, respectively, and let $(n/u)^{1/2}Y=X-\mu$. Then X has the multivariate t-distribution $t_p(\mu,\Sigma,n)$, that is,

$$f(x) = \frac{C_p|\Sigma|^{-1/2}}{\{n+(x-\mu)'\Sigma^{-1}(x-\mu)\}^{(n+p)/2}}, \quad -\infty < x_i < \infty, \quad n>0, \tag{1}$$

for $i=1,\dots,p$, where $C_p=(n^{n/2}\Gamma\{(n+p)/2\})/(\pi^{p/2}\Gamma\{n/2\})$.

The mean and covariance matrix of the t-distribution are easily shown to be $E(X)=\mu,\ n>1$, and $\mathrm{Var}(X)=n\Sigma/(n-2),\ n>2$.

If in the density of (1), X, μ, and Σ are partitioned as $X=(X_1',X_2')'$, $\mu=(\mu_1',\mu_2')'$, and $\Sigma=\{\Sigma_{ij}\}$, $i,j=1,2$, where $X_i:p_i\times 1$, and Σ_{ij} is a submatrix of order $p_i\times p_j$. Then the conditional distribution of X_2 is given by the following result.

Remark 1. Suppose $X:p\times 1$ follows the distribution given in (1) and $X'=(X_1',X_2')$. the conditional distribution of $X_2:p_2\times 1$ given X_1 is also a multivariate t-distribution. Specifically, conditional on X_1, X_2 has distribution

$$t_{p_2}(\mu_{2.1},\alpha_{2.1}\Sigma_{22.1},n+p_1),$$

where

$$\mu_{2.1}=\mu_2+\Sigma_{21}\Sigma_{11}^{-1}(x_1-\mu_1),\quad \Sigma_{22.1}=\Sigma_{22}-\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$$

and

$$a_{2.1}=(n+p_1)^{-1}\{n+(x_1-\mu_1)'\Sigma_{11}^{-1}(x_1-\mu_1)\}.$$

Proof. Expand the quadratic for in (1);

$$(x-\mu)'\Sigma^{-1}(x-\mu) = (x_1-\mu_1)'\Sigma_{11}^{-1}(x_1-\mu_1)$$
$$+\{(x_2-\mu_2)-\Sigma_{21}\Sigma_{11}^{-1}(x_1-\mu_1)\}'\Sigma_{22.1}^{-1}\{(x_2-\mu_2)-\Sigma_{21}\Sigma_{11}^{-1}(x_1-\mu_1)\}.$$

Complete the square in x_2, and integrate with respect to x_2. This gives the marginal distribution of X_1 as $t_{p_1}(\mu_1, \Sigma_{11}, n)$ (cf. Press 1989). The ratio of the expanded quadratic form for (1) to the marginal density of X_1 yields the conditional distribution(cf. Box and Tiao 1973).

To work with standard form, we transform X by $t=D^{-1/2}(X-\mu)$ so that $t=(t_1, \cdots, t_p)'$ is distributed as $t_p(0, R, n)$, where $D=diag(\Sigma)$ and $R=\{\rho_{ij}\}$ is the correlation matrix of Y defined in Definition 1. Note that the marginal distribution of each $t_i(i=1, \cdots, p)$ is a univariate Student's t with n degrees of freedom.

Suppose we are interested in an orthant probability

$$\Pr(t_1 < b_1, \cdots, t_p < b_p). \tag{2}$$

The probability can be decomposed as the product of conditional probabilities,

$$\Pr(t_1 < b_1, t_2 < b_2) \times \prod_{k=3}^{p} \Pr(t_k < b_k | t_1 < b_1, \cdots, t_{k-1} < b_{k-1}). \tag{3}$$

Let $I_i = I(t_i < b_i)$, $i=1, \cdots, p$ where $I(A)$ denotes the indicator of the event A. Thus we can see that $E(I_i)=F(b_i)$, where F is cumulative distribution function of univariate Student's t with n degrees of freedom. Using the regression method with binary variables, we approximate the conditional probabilities

$$\Pr(t_k < b_k | t_1 < b_1, \cdots, t_{k-1} < b_{k-1}) = E(I_k | I_1=1, \cdots, I_{k-1}=1) \tag{4}$$

by

$$E(I_k) + \Omega_{21}\Omega_{11}^{-1}(1-E(I_1), \cdots, 1-E(I_{k-1}))' \tag{5}$$

where Ω_{21} is a row vector whose components are consist of $cov(I_k, I_i) = E(I_k I_i) - E(I_k)E(I_i)$, $i=1, \cdots, k-1$, and Ω_{11} is a $k-1 \times k-1$ matrix with (i, j) element being $\text{Cov}(I_i, I_j) = E(I_i I_j) - E(I_i)E(I_j)$, $1 \le i$, $j \le k-1$, where the exact formula for $E(I_i I_j) = \Pr(t_i < b_i, t_j < b_j)$ is obtained by Dunnett and Sobel(1954) ;

For odd n

$$
\begin{aligned}
&\Pr(t_i < b_i, t_j < b_j) \\
&= \frac{1}{2\pi}\tan^{-1}\left[-n^{1/2}\frac{(b_i+b_j)(b_ib_j+\rho_{ij}n)-(b_ib_j-n)g}{(b_ib_j-n)(b_ib_j+\rho_{ij}n)-n(b_i+b_j)g}\right] \\
&\quad + \frac{1}{4(\pi n)^{1/2}}\sum_{k=1}^{(n-1)/2}\frac{\Gamma(k)}{\Gamma(k+1/2)}\left\{\frac{b_i[1+\operatorname{sgn}(b_j-\rho_{ij}b_i)I_{f(b_j,b_i)}(1/2,k)]}{(1+b_i^2/n)^k}\right. \\
&\qquad\qquad \left. + \frac{b_j[1+\operatorname{sgn}(b_i-\rho_{ij}b_j)I_{f(b_i,b_j)}(1/2,k)]}{(1+b_j^2/n)^k}\right\},
\end{aligned}
$$

where

$$g=\{b_1^2-2\rho_{ij}b_ib_j+b_j^2+n(1-\rho_{ij}^2)\}^{1/2},$$

$$f(b_i,b_j)=\frac{(b_i-\rho_{ij}b_j)^2}{(b_i-\rho_{ij}b_j)^2+(1-\rho_{ij}^2)(n+b_j^2)},$$

$\operatorname{sgn}(\theta)=1$ for $\theta\geq 0$, otherwise 0 and $I_p(c,d)$ is the incomplete beta function ratio(cf. Johnson and Kotz, 1972). See Johnson and Kotz(1972), for the formula for even n. Noticing that $E(I_i^2)=E(I_i)$ we can easily verify that (4) and (5) are identical for $k=2$. If (5) happens to exceed 1 or less than zero, we use 1 and zero as their approximation values, respectively.

The use of (5) as an approximation to (4) is analogous to the formula of Remark 1:

$$\mu_{2.1}=E(X_2|X_1=x_1)=\mu_2+\Sigma_{21}\Sigma_{11}^{-1}(x_1-\mu_1) \tag{6}$$

for a multivariate t random vector $(X_1',X_2')'$ with mean vector $(\mu_1',\mu_2')'$ and partitioned scale matrix $\Sigma=\{\Sigma_{ij}\}$, $i,j=1,2$.

The approximation formula (5) can be substituted into the decomposition (3) to get one approximation to the rectangular probability in (2). However, the decomposition into conditional probabilities is not unique, and hence we can expect that different approximations lead to different approximations. That is, (3) is equivalent to

$$\Pr(t_{i_1}<b_{i_1},t_{i_2}<b_{i_2})\prod_{k=3}^{m}\Pr(t_{i_k}<b_{i_k}|t_{i_1}<b_{i_1},\cdots,t_{i_{k-1}}<b_{i_{k-1}}), \tag{7}$$

where $(i_1,\cdots,i_m)$ is a permutation of $(1,\cdots,m)$ with $i_1<i_2$. Thus the number of permutations is $m!/2$. For example, for $m=3$, there are 3 permutations that could be considered:

$$
\begin{aligned}
\Pr(t_1<b_1,t_2<b_2,t_3<b_3) &= \Pr(t_1<b_1,t_2<b_2)\Pr(t_3<b_3|t_1<b_1,t_2<b_2) \\
&= \Pr(t_1<b_1,t_3<b_3)\Pr(t_2<b_2|t_1<b_1,t_3<b_3) \\
&= \Pr(t_2<b_2,t_3<b_3)\Pr(t_1<b_1|t_2<b_2,t_3<b_3).
\end{aligned}
$$

For each permutation, the approximation of the form (5) can be used for each conditional probability, so that we may get overall approximation of the orthant probability (2) by averaging the $m!/2$ approximations:

$$\Pr(t_1 < b_1, \cdots, t_p < b_p) \approx \frac{1}{m!/2} \sum \Big[\Pr(t_{i_1} < b_{i_1}, t_{i_2} < b_{i_2}) \times \prod_{k=3}^{m} \{ F(b_{i_k}) - \Omega_{21}^{*} \Omega_{11}^{*-1} (1 - F(b_{i_1}), \cdots, 1 - F(b_{i_{k-1}}))' \} \Big], \tag{8}$$

where F is cumulative distribution function of univariate Student's t with n degrees of freedom, Ω_{21}^{*} is a row vector whose components are consist of $\Pr(t_{i_k} < b_{i_k}, t_{i_\ell} < b_{i_\ell}) - F(b_{i_k}) F(b_{i_\ell})$, $\ell = 1, \cdots, k-1$, and Ω_{11}^{*} is a $k-1 \times k-1$ matrix with $(i_\ell, i_{\ell'})$ element being $\Pr(t_{i_\ell} < b_{i_\ell}, t_{i_{\ell'}} < b_{i_{\ell'}}) - F(b_{i_\ell}) F(b_{i_{\ell'}})$, $1 \leq \ell$, $\ell' \leq k-1$. The standard deviation of the $m!/2$ approximations gives rough measure of precision of the overall approximation. In practice, when m is large 10^2 to 10^4 randomly selected permutations can control the standard error associated with the average of permutation(cf. Joe 1995). The orthant probabilities of the general multivariate t-distribution (1) can be obtained in a similar way, with t_i replaced by $(X_i - \mu_i)/\sigma_{ii}$, $i = 1, \cdots, p$ in (3).

3. Numerical Comparison and Conclusion

In this section we make a comparison of the suggested approximation(NEW) and two other previously proposed approximations with exact orthant probability. The two other approximation methods consisting of the method by Dunnett and Sobel (1954)(DS) and that based directly on Bonferroni's inequality(BON). They are well explained in Johnson and Kotz (1972, p.140-p.141). The orthant probability to be compared in this section is

$$\Pr\left(\bigcap_{i=1}^{p} t_p \leq t_{n(p),\alpha} \right) = \Pr\left(\max(t_1, t_2, \cdots, t_p) \leq t_{n(p),\alpha} \right) = \alpha.$$

Comparison of exact values of $t_{n(p),\alpha}$ points with the approximations for selected values of n, p and $\rho_{ij} = 1/2$ for all i and j is noted in Table 1. For constructing the table we used the exact values of $t_{n(p),\alpha}$ obtained by Cornish(1962). For $p = 6$, we tabulated the average value of randomly selected 10^2 permutations for NEW with the standard deviation in the parenthesis.

TABLE 1. Comparison of the percentage points $t_{n(p),\alpha}$ for $\rho_{ij}=1/2$ for all i and j

		$p=3$				$p=6$			
α	n	BON	DS	NEW	EXACT	BON	DS	NEW	EXACT
.99	5	4.46	4.46	4.20	4.21	5.75	5.75	5.02(.018)	5.03
	∞	2.71	2.71	2.67	2.68	3.06	3.06	3.01(.015)	3.00
.95	5	2.91	2.90	2.68	2.68	3.93	3.90	3.29(.012)	3.30
	∞	2.13	2.12	2.04	2.05	2.54	2.53	2.07(.021)	2.06
.75	5	1.62	1.55	1.33	1.32	2.48	2.38	1.83(.025)	1.81
	∞	1.38	1.33	1.19	1.19	1.19	1.86	1.60(.013)	1.60
.50	5	1.07	0.89	0.63	0.62	1.93	1.71	1.11(.024)	1.10
	∞	0.97	0.82	0.60	0.59	1.59	1.45	1.04(.013)	1.04

We have suggested an approximation method for an orthant probability of the multivariate t-distribution. It is quite simple and depend on conditional expectations and regression for binary variables. The approximation is based upon standardized form of the multivariate t-distribution, so that the approximation is also usable for other multivariate t-distributions.

Since, except for some special distributions(cf. Johnson and Kotz 1972), an exact method and an approximation method for the probability of the orthant region were not available, we conducted a limited but informative numerical study that confined to the cases in Table 1. As noted by the table the suggested approximation is good enough to be taken as an approximation method for the orthant probability. Moreover, we may highlight the merit of the suggested approximation that it also provides the approximate probability of the orthant region in the case where the ρ_{ij}'s are not all the same and where the t_i's and b_i's in (2) are not all the same.

In this paper, we considered an approximation for the lower orthant probabilities. The main ideas of the approximation can be easily extended to an approximation for the upper orthant probabilities of the form $\Pr(t_i > a_i, i=1,\cdots,p)$ and that for the probabilities of rectangular regions. This is another interesting research topic and it is left for a future study.

References

[1] Bökenholt, U. (1992). Thurstonian representation for partial ranking data. *British Journal of Mathematical and Statistical Psychology*, Vol. 45, 31-49.

[2] Box, G. E. P. and Tiao, G. C. (1973). *Bayesian Inference in Statistical Analysis*. Addison-Wesley, Massachusetts.

[3] Collett, D. (1991). *Modelling Binary Data*. Capman Hall, London.

[4] Cornish, E. A. (1962). *The Multivariate t-distribution Associated with the General Multivariate Normal Distribution.* Technical Report No. 13, Division of Mathematical Statistics, Commonwealth Scientific and Industrial Research Organization, Australia.

[5] Dunnett, C. W. and Sobel, M. (1954). A bivariate generalization of Student's t-distribution, with tables for certain special cases. *Biometrika*, Vol. 41, 153-169.

[6] Genz, A. (1992). Numerical computation of multivariate normal probabilities. *Journal of Computational 'and Graphical Statistics*, Vol. 1, 141-149.

[7] Joe, H. (1995). Approximations to multivariate normal rectangle probabilities based on conditional expectations. *Journal of the American Statistical Association*, Vol. 90, 957-964.

[8] John, S. (1962). On the evaluation of the probability integral of multivariate t-distribution. *Biometrika*, Vol. 48, 409-417.

[9] Johnson, M. L. and Kotz, S. (1972). *Continuous Multivariate Distributions in Statistics*, John Wiley Sons, Inc..

[10] Press, S. J. (1989). *Bayesian Statistics*, John Wiley Sons, Inc..

[11] Solow, A. R. (1990). A method for approximating multivariate normal orthant probabilities. *Journal of Statistical Computation and Simulation*, Vol. 37, 225-229.