THE EGYPTIAN STATISTICAL JOURNAL
ISSR, CAIRO UNIV., VOL., 41, NO.1, 1997

On the Distribution of Linear Combination of t-variables

Amina Ali Saleh *

Walker and Saw (1978) derived the distribution of an arbitrary linear combination of student's random variables with odd degrees of freedom. An extension to the case of any degrees of freedom is demonstrated in this paper. The proposed procedure uses approximation of the characteristic function of the student's t distribution given by Saleh (1994). For the special case of the sum or the difference of two independent t-variables, our approximation is compared with the exact probability given earlier by Ghosh (1975), and good results are obtained.

Key words: Distribution function ; Characteristic function ; t-variables ; Linear combination of " t-variables".

1. Introduction

The distribution of a linear combination of independent t- variables has been of considerable interest. For example, the distribution of the difference of two independent t-variables which arises in certain well-known procedures for problems concerning the difference between the means of two normal populations with unknown variances.This special case received extensive attention in the literature. Fisher and Yates (1957) tabulated values of the percentage points of the distribution of the difference of two independent t-variables for some selected value of degrees of freedom. Ghosh(1975)gave explicit formulas and values of the percentage points and the incomplete probability integral when the two t- variables have equal degrees of freedom, and gave approximation for the percentage points for the case of unequal degrees of freedom and compared them with the earlier one given by Fisher and Yates.

For any set of real constants $c_1, c_2, \cdots, c_n$, let

$$T^* = c_1 t_{f_1} + c_2 t_{f_2} + \cdots + c_n t_{f_n} , \qquad (1,1)$$

here the t_{f_i}'s are Student-t variables independently distributed with f_i degrees of freedom. If $f_i = 1$ for $i = 1, 2, \ldots, n$, then T^* will have the Cauchy distribution, and if $f_i < \infty$ for each i; T^* will be normally distributed. If $f_i < \infty$ and odd for all i, Walker and Saw (1978) gave an expression of T^* as a mixture of t- distributions which enabling the calulation of

*Institute of Statistical Studies and Research, Cairo University

percentage points using only tables of the t-distribution. For other values of f_i , the distribution of T^* no longer has a recognizable form.

In this articale we provide useful approximation for the distribution of T^* for any degrees of freedom. This distribution function may be expressed approximately as a weighted sum of student- t distribution functions with odd degrees of freedom.

Define the "pre-t" random variable as

$$X_f = t_f / \sqrt{f}$$

The distribution function of T^* is

$$\Pr\{ \sum_{i=1}^{n} c_i \, t_{f_i} \le u' \} = \Pr\{ \sum_{i=1}^{n} a_i \, X_{f_i} \le u \},$$

where $a_i = c_i \sqrt{f_i} / \omega$, $u = u'/\omega$ and $\omega = \sum_{i=1}^{n} c_i \sqrt{f_i}$.

The distribution of T^* is invariant over sign changes in a_i , therefore, we need only consider the distribution of

$$T = \sum_{i=1}^{n} a_i \, X_{f_i} = T^* / \omega \qquad (1.2)$$

where $a_i \ge 0$ for all i and $\sum_{i=1}^{n} a_i = 1$.

2. Approximate Distribution of T

The characteristic function of X_f is

$$\phi\,(\theta\,;f) = 2\int_0^{\infty} \left(\cos|\theta|\,x\right) g(x)\,dx \;.$$

where $g(x)$ is the density function of X_f .

When f is odd, say $f = 2m + 1$, it is known (see , e.g., Chapman 1950) that $\phi(\theta; f)$ reduces to

$$\phi(\theta; f) = e^{-|\theta|} q_m(|\theta|), \tag{2.1}$$

where

$$q_m(|\theta|) = \frac{m!}{(2m)!} \sum_{k=0}^{m} \frac{(2m-k)!}{k!(m-k)!} (2|\theta|)^k \ , \tag{2.2}$$

When f is even , say $f = 2m$, no such simple form for the characteristic function exists. However, Saleh (1994) , obtained an explicit and simple approximation form for it as a linear combination of the characteristic function when $f = 2m - 1$, and the charactristic function when $f = 2m + 1$, as follows :

(I) For $m \geq 2$,

$$\phi(\theta; 2m) = \hat{\alpha}_{2m} \phi(\theta; 2m-1) + (1 - \hat{\alpha}_{2m}) \phi(\theta; 2m+1) \ , \tag{2.3}$$

where ,

$$\hat{\alpha} = \frac{\sum_{i=1}^{n} [\mu_{2r}(2m-1) - \mu_{2r}(2m+1)][\mu_{2r}(2m) - \mu_{2r}(2m+1)]}{\sum_{i=1}^{n} [\mu_{2r}(2m-1) - \mu_{2r}(2m+1)^2} \tag{2.4}$$

and where, $\mu_{2r}(n)$ is the $2r$th moment of the t-distribution with n degrees of freedom and where s is a positive integer arbitrary chosen. The numerical examples of Saleh's study indicates that the best choice of s is 1.

(II) For $m = 1$,

$$\phi(\theta, 2) \cong \tfrac{1}{4} \phi(\theta; 1) + \tfrac{3}{4} \phi(\theta; 3) \tag{2.5}$$

Walker and Saw (1978), demonstrated that the charactristic function of T in (1.2), may be expressed as a linear compination of the " pre-t " charactristic functions as long as f_i's are odd. In the following we shall demonstrste that, if some or all f_i's are even, then ,by applying the approximation forms in (2.3)-(2.4), the charactristic function of T may be still has the same property. Consequently, the distribution function of T may always be expressed approximately as a weighted sum of Student-t distribution functions with odd degrees of freedom..

Following Walker and Saw, let Q be the matrix whose element in row i and column j is the coefficient of $e^{-|\theta|}|\theta|^j$ in $\phi(\theta;2i+1)$.

$$(Q)_{ij} = i!(2i-j)!2^j/(2i)!j!(i-j)! \,, \qquad \text{if } j = 0,1,2,\dots,I \quad i = 0,1,2,3,\dots$$

$$= 0, \qquad \text{if } j > I$$

The matrix Q is lower-triangular with nonzero diagonal elements, hence it is nonsingular. Let Q_i represents the ith row of Q, $i = 0, 1, 2, \dots$. Since Q is lower-triangular Q_i has zeros after the first $i+1$ elements (the first element, corresponding to column 0, is unity). For m, and j in $\{0,1,2,\dots\}$, Q^{-1} is the matrix whose element in row m and column j is

$$(Q^{-1})_{m,j} = \frac{(-1)^{m-j}(2j)!(m+1)!}{2^m\,(m-j)!\,j!(2j-m+1)!}\,, \qquad \text{if } j \le m \le 2j+1$$

$$= 0 \qquad \text{otherwise}$$

For the proof, see Walker and Saw (1976), they also gave the first eleven rows and columns of Q^{-1} For completeness, Table 1 gives the first eleven rows and nine columns of Q^{-1}.

ISSR, Cairo Univ., Vol.,41, No.1, 1997, PP. 51-59

Table 1: The lower-triangular Elements of the first. Rows and columns of Q^{-1}

	Column								
Row	0	1	2	3	4	5	6	7	8
0	1								
1	-1	1							
2	0	-3	3						
3	0	3	-18	15					
4	0	0	45	-150	150				
5	0	0	-45	675	-1575	945			
6	0	0	0	-1575	11025	-19845	10395		
7	0	0	0	1575	-44100	198450	-291060	135135	
8	0	0	0	0	49225	1190700	3929310	-4864860	2027025
9	0	0	0	0	-99225	4465125	-32744250	85135050	-91216125
10	0	0	0	0	0	-9823275	180093375	-936485550	2006754750

Let Φ and $\Theta(c)$ denote the arrays

$$\Phi = \begin{pmatrix} \phi(\theta;1) \\ \phi(\theta;3) \\ \phi(\theta;5) \\ \cdot \\ \cdot \\ \cdot \end{pmatrix} \quad \text{and} \quad \Theta(c) = \begin{pmatrix} 1 \\ |\theta c| \\ |\theta c|^2 \\ |\theta c|^3 \\ \cdot \\ \cdot \\ \cdot \end{pmatrix}$$

Recall that,

$$\phi(\theta; 2m+1) = e^{-|\theta|} q_m(|\theta|)$$

we may write

$$\Phi = e^{-|\theta|} Q \Theta(1) \tag{2.5}$$

since Q is nonsingular, we may rewrite (2.5) as

$$Q^{-1} \Phi = e^{-|\theta|} Q \Theta(1) \tag{2.6}$$

The charactristic function of $T = \sum_{i=1}^{n} a_i X_{fi}$ is

$$\phi_T(\theta) = \pi_{j=1}^{n} \phi(\theta a_j ; f_j)$$

$$\approx e^{-|\theta|} \prod_{j=1}^{n} Q^{*}_{f_j} \Theta(a_j) \qquad (2.7)$$

$Q^{*}_{f_j}$ represents the $\left(\frac{f_j - 1}{2}\right)$th row of Q if f_j is odd, and the linear combination

$\{ \hat{\alpha}_{f_j} Q_{\frac{f_j}{2}-1} + (1 - \hat{\alpha}_{f_j}) Q_{\frac{f_j}{2}+1} \}$ of $(\frac{f_j}{2} - 1)$, $(\frac{f_j}{2} + 1)$ rows of the Q

if f_j is even and the wghits $\hat{\alpha}_{f_j}$ are as given in (2.3).

Note that $Q^{*}_{f_j} \Theta(a_j)$ is a polynomial in $|\theta|$ of degree at most v_j ,where

$$v_j = \frac{f_j - 1}{2} \qquad \text{if } f_j \text{ is odd}$$

$$= \frac{f_j}{2} + 1 \qquad \text{if } f_j \text{ is even}$$

Therefore $\phi_T(\theta)$ is a polynomial in $|\theta|$ of degree at most $v = \sum_{j=1}^{n} v_j$. We may find a vector λ such that

$$\phi_T(\theta) \approx e^{-|\theta|} \lambda' Q^{-1} \Phi$$

$$= \eta' \Phi \qquad (2.9)$$

where $\eta' Q^{-1}$

ISSR, Cairo Univ., Vol.,41, No.1, 1997, PP. 51-59

Since $\Phi_T(\theta)$ is a polynomial in $|\theta|$ of degree at most v, η has zeros after its first v+1 elements. Setting $\theta = 0$ in (2.9), we have $\sum \eta_i = 1$, η_i being the ith element of η (i = 1, 2, ...). Numerical investigation indicates also that η_i may be negative for some i , i = 1, 2, ...

From (2.9), the distribution function of T is given by

$$P(T \leq c) = \eta_0 G_0(c) + \eta_1 G_1(c) + \ldots + \eta_v G_v(c) \qquad (2.10)$$

where $G_i(c) = P(X_{2i+1} \leq c)$, is the distribution function of $t_{2i+1}/\sqrt{(2i+1)}$.

Example:

Suppose that it is required to find the distribution of $T = (5/12)\, X_1 + (1/3)\, X_3 + (1/4)\, X_8$. The characteristic function of X_8 is approximated as :

$$\phi(\theta\,;8) \approx \hat{\alpha}_8\, \phi(\theta\,;7) + (1 - \hat{\alpha}_8)\, \phi(\theta\,;9)$$

$$\text{where } \hat{\alpha}_8 = \frac{[\mu_2(7) - \mu_2(9)][\mu_2(8) - \mu_2(9)]}{[\mu_2(7) - \mu_2(9)]^2} = 5/12$$

hence

$$\phi(\theta\,;8) \approx e^{-|\theta|} \left\{ \frac{5}{12}\frac{3!}{6!} \sum_{k=0}^{3} \frac{(6-k)!}{k!(3-k)!}(2|\theta|)^k + \frac{7}{12}\frac{4!}{8!} \sum_{k=0}^{4} \frac{(8-k)!}{k!(4-k)!}(2|\theta|)^k \right\}$$

$$= e^{-|\theta|}\, [1 + |\theta| + (10/24)\,|\theta|^2 + (1/12)\,|\theta|^3 + (1/180)\,|\theta|^4]$$

Now, the characteristic function of T is approximated as

$$\Phi_T(\theta) \approx \phi(5/12\,\theta\,;1)\ \phi(1/3\,\theta\,;3)\ \phi(1/4\,\theta\,;8)$$

$$= e^{-|\theta|}\, (1 + 1/3\,|\theta|)(1 + 1/4\,|\theta| + 10/384\,|\theta|^2 + 1/768\,|\theta|^3 + 1/46080\,|\theta|^4$$

$$+ 1/138240\,|\theta|^5)$$

so that
$\lambda' = (1, 7/12, 504/4608 , 8832/884736 , 48384/106168320 , 1/138240 , 0 , 0 , 0 , ...)$
and
$\eta' = \lambda' Q^{-1}$
$= (5/12, .2851563 , .1686198 , .086263 , .0364583 , .0068359, 0 , 0 , 0, ...)$
Thus
$P(T \leq c) \approx 5/12\, P(t_1 \leq c) + (.28551563)\, P(t_3 \leq c) + (.1686198)\, P(t_5 \leq c)$

$+ (.086263)\, P(t_7 \leq c) + (.0364583)\, P(t_9 \leq c) + (.0068359)\, P(t_{11} \leq c)$

In computing $P (t_n \leq c)$,for odd n, we use the expressions (Johnson and Kotz (1970))

(I) $P(t_1 \leq c) = 1/2 + 1/\pi \tan^{-1}$,

(ii) $P(t_n \leq c) = 1/2 + 1/\pi [u + \{\cos u + \frac{2}{3} \cos^3 u + ... + \frac{(2)(4)...(n-3)}{(3)(5)...(n-2)} \cos^{n-2} u\} \sin u]$

where $u = \tan^{-1} (c/\sqrt{n})$

3.Adequcy of the Approximation

The proposed method demonstrated in this article may be used to obtain the (approximate) distribution of any arbitrary linear combination of student's random variable with any degrees of freedom. No other published worke exists to compare it with the proposed method. However, for the case of the difference (or the sum) of two independent t-variables, Ghosh (1975) gave the exact distribution function of $Z_{n,n} = t_n - t_n$ and a fairly extensive table of values of $P (Z_{n,n} \leq c)$ for n = 1 to n = 20 . In the following we shall compare the approximate probability using (2.10) with the exact one given by Ghosh (1975) for n = 4 , 6, 8 and 10. Table 2 summarize the result of this comparison.

ISSR, Cairo Univ., Vol.,41, No.1, 1997, PP. 51-59

Table 2 : Adequcy of approximation (2.10) to P($Z_{n,n} \le c$)

C	n = 4		n = 6		n = 8		n = 10	
	Exact	App.	Exact	App.	Exact	App.	Exact	App.
5	.618	.625	.625	.626	.628	.629	.630	.630
1.0	.724	.736	.736	.739	.742	.743	.745	.746
1.5	.810	.822	.625	.828	.832	.833	.837	.837
2.0	.873	.884	.889	.891	.897	.898	.902	.903
2.5	.916	.925	.932	.934	.940	.941	.945	.945
3.0	.945	.952	.960	.961	.967	.967	.970	.970
3.5	.964	.968	.976	.977	.982	.982	.985	.983
4.0	.976	.979	.986	.986	.990	.990	.992	.992
4.5	.984	.985	.992	.992	.995	.995	.996	.996
5.0	.989	.990	.995	.995	.997	.997	.998	.998
5.5	.992	.993	.997	.997	.999	.998	.999	.999
6.0	.994	.994	.998	.998	.999	.999	.999	.999

The implication of Table 2 is obvious. The approximation turns out to be reasonably good even for small n, and gets better as n increases.

REFERENCES

Chapman, D.G., (1950) Some Two-Sample Tests, Annalas of Mathematical Statistics, 21 pp.601-606

Fisher , R. A. and Yates, F. (1957) Statistical Tables, New York: Hafner publishing co.

Ghosh, K. ,(1975) On the distribution of the difference of two t- variables , JASA, 70 pp.463-467.

Johnson N.L. and Kotz,S. (1970) Continuous Univariate distribution-2 (New York, Wiley)

Saleh, A.A (1994) Approximating the characteristic function of the student's t distribution, The Egyptian Statistical Journal, 38 pp. 177-195.

Walker, G.A. and Saw, J.G.(1978) The Distribution of linear combinations of t-variables, JASA 73 , pp. 876-878.