The Cumulants of the Non-normal t Distribution

Hark Hwang*

1. Introduction

The simplicity and usefulness of the Student's "t" defined in case of a normally distributed sample is due to the independence of $\overline{X}$ and S.

However, in a variety of cases, it is necessary to test for the mean of a population which does not come from normal distribution and $\overline{X}$ and S are no longer independent.

Thus, the following question is of interest. If t_p^0 is the 100 p percentile from the normal tabulated "t" and if t is computed from non-normal sample data, what is the true probability, i.e., prob $[t \leq t_p^0]$?

A solution to this question consists of finding the cumulative distribution function of the non-normal t, this is done with an Edgeworth series which expresses $F(t)$ in terms of the moments of the non-normal distribution.

Geary [7] furnished an expression for the first four moments of t correct to n^{-2} with following way:

$$\text{Let} \qquad t=\frac{\sqrt{n}\ k_1}{\sqrt{k_2}}=\frac{\sqrt{n}\ k_1}{\sqrt{K_2}\left(1+\dfrac{k_2-K_2}{K_2}\right)^{\frac{1}{2}}}$$

using R.A. Fisher's notation k_i and K_2, the population variance. In order to find the first four moments of t (from zero origin) the denominator is expanded formally in power of $\dfrac{k_2-K_2}{K_2}$, which is of the order $n^{-\frac{1}{2}}$.

Then joint moments of k_1 and k_2 are substituted with joint cumulants of k_1 and k_2 and again joint cumulants are replaced by cumulants of the

* Assistant Professor, Korea Advanced Institute of Science.

population according to the simple formulae which Fisher has provided.

Later Geary [7] gave the expansion to the terms in n^{-2} of the first six cumulants L_i of t as follows.

$$L_1 \simeq -\frac{1}{n^{\frac{1}{2}}}\left\{\frac{\lambda_3}{2}+\frac{3}{16n}(2\lambda_3-2\lambda_5+5\lambda_3\lambda_4)\right\}+\cdots,$$

$$L_2 \simeq 1+\frac{1}{4}\left(8+7\lambda_3^2\right)n^{-1}+\left(6-2\lambda_4-\frac{3}{8}\lambda_3^2-\frac{45}{8}\lambda_3\lambda_5+\frac{177}{16}\lambda_3^2\lambda_4\right)n^{-2},$$

$$L_3 \simeq -2\lambda_3 n^{-\frac{1}{2}}-\left(9\lambda_3-3\lambda_5+\frac{15}{4}\lambda_3\lambda_4+\frac{83}{8}\lambda_3^3\right)n^{-\frac{3}{2}},$$

$$L_4 \simeq \left(6-2\lambda_4+12\lambda_3^2\right)n^{-1}+\left(54-18\lambda_4+4\lambda_6+75\lambda_3^2-63\lambda_3\lambda_5-6\lambda_4^2 +81\lambda_3^2\lambda_4+\frac{699}{8}\lambda_3^4\right)n^{-2}$$

$$L_5 \simeq -\left(60\lambda_3-6\lambda_5-20\lambda_3\lambda_4+105\lambda_3^3\right)n^{-\frac{3}{2}},$$

$$L_6 \simeq \left(240-120\lambda_4+577\frac{1}{2}\lambda_3^2+16\lambda_6-210\lambda_3\lambda_5-150\lambda_3^2\lambda_4+1200\lambda_3^4\right)n^{-2}.$$

where $\lambda_i = K_i/K_2^{i/2}$ and K_i are the cumulants of the population.

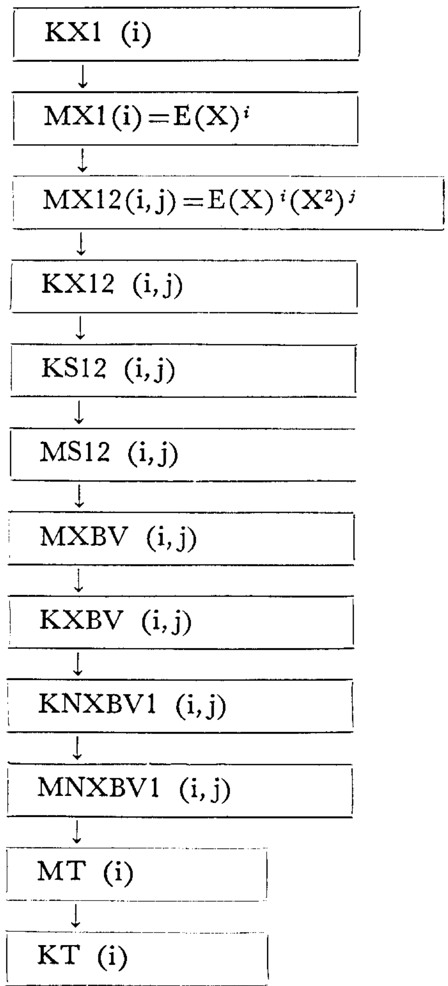
A program for obtaining the first eight cumulants of t in terms of the population cumulants using Algebraic Manipulation Package [10] [12] (ALMAP) of the University of Minnesota is established and Figure (1.1) shows the flow chart of the procedure for generating the cumulants of t in computer programming.

Notations for statistical variables in computer programming are made in such a way that each notation is easy enough to read. Here the initial letter K and M are adopted to represent the meaning of cumulant and moment respectively.

2. Moments and Cumulants Recursion Equation

As shown in Figure 1 the procedure involves the conversion from moments to cumulants and vice versa in univariate and bivariate case at various stages.

Figure 1. Flow Chart of the Procedure for Calculation of the Cumulants of t



Let m_j be the jth moment of a random variable X and K_j be the jth cumulant of X.

R.A. Fisher [4] first developed expansion of m_j in terms of K_i, $i \leq j$. Later M.G. Kendall [9] developed more explicitly the approach used by Fisher. The procedure for generating moments in terms of cumulants mentioned by them is that of equating coefficients in the power series expansions of the defining relationship between the respective generating functions, *viz.*,

$M(t)=\exp\ K(t)$.

From the same approach F.N. David and M.G. Kendall [3] developed the expression of m_j in terms of K_j using tables which express the relationship between two symmetric functions; the augmented monomial symmetric function and the power sum function. The procedures for generating the relationship between m_j and K_j given by the above is not oriented toward computer programming.

Another approach established by J.S. White [11] gives a recursion relation more adapted to computer programming and the result is following.

In univariate case

$$m_{j+1}=K_{j+1}+\sum_{i=1}^{j}\binom{j}{i}K_{j-i+1}m_i \tag{2.1}$$

Equation (2.1) may also be used as a recursion relation for expressing K_{j+1} in terms of m_i, $i\leq j+1$.

$$K_{j+1}=m_{j+1}-\sum_{i=1}^{j}\binom{j}{i}K_{j-i+1}m_i \tag{2.2}$$

In bivariate case with

$$M(s,t)=\sum m_{ij}\frac{s^i}{i!}\frac{t^j}{j!} \qquad \text{and}$$

$$K(s,t)=\log M(s,t)=\sum K_{ij}\frac{s^i}{i!}\frac{t^j}{j!}\ ,$$

$$\begin{aligned} m_{i+1,j+1}=&K_{i+1,j+1}+K_{i+1,0}\ m_{0,j+1}\\ &+\sum_{q=1}^{j}\binom{j+1}{q}K_{i+1,j+1-q}\ m_{o,q}\\ &+\sum_{p=1}^{i}\binom{i}{p}K_{i+1-p,j+1}\ m_{p,0}+K_{i+1-p,0}\ m_{p,j+1}\\ &+\sum_{p=1}^{i}\sum_{q=1}^{j}\binom{i}{p}\binom{j+1}{q}K_{i+1-p,j+1-q}m_{pq} \end{aligned} \tag{2.3}$$

where $\sum_{p=1}^{0}=0$, $\sum_{q=1}^{0}=0$ and

where m_{ij} is (i,j)th joint moment and K_{ij} is (i,j)th joint cumulant of two different random variables. The corresponding expression for the joint cumulant is obtained directly from (2.3) by solving for $K_{i+1,j+1}$.

Since zero subscripts are not allowed in programming in FORTRAN, some care is required in programming. The univariate and bivariate recursions are easily adapted to computer programming and the FORTRAN subroutine VK2M, VM2K corresponding to (2.1) and (2.2) respectively and VK2M2 corresponding to (2.3) and VM2K2 are developed[8].

3. Cumulants of t

3-1 The Joint Moments of $S1$ and $S2$

Let X, $X_i (i=1, 2, \cdots, N)$ be independent identically distributed random variables with cumulants $K_i = KX1(i)$.

The moments of X, i.e., $m_i = MX1(i)$ are derived from the cumulants of X using the subroutine VK2M.

Now consider the joint distribution of X and X^2.

The joint moments of X and X^2 are

$$MX12\ (i,j) = E(X^i (X^2)^j) = MX1(i+2j) \qquad (3.1)$$

Knowing $MX12(i,j)$, subroutine VM2K2 yields the joint cumulants of X and X^2, $KX12\ (i,j)$.

Let $S1 = \sum X_i$ and $S2 = \sum X_i^2$, then the joint moment generating function of ($S1$, $S2$) is

$$M_{s1,s2}(u,v) = E(e^{uS_1+vS_2}) = E(e^{uX+vX^2})^N$$
$$= (e^{K_{X,X^2}(u,v)})^N = e^{K_{S1,S2}(u,v)} \qquad (3.2)$$

Here $K_{S1,S2}(u,v)$ is the joint cumulant generating function of ($S1$,$S2$).

Thus $K_{S1,S2}(u,v) = NK_{X,X^2}(u,v)$ (3.3)

Hence $KS12\ (i,j) = NKX12(i,j)$ (3.4)

where $KS12(i,j)$ is the joint cumulant of $S1$ and $S2$. The subroutine VK2M2 yields the joint moment of $S1$ and $S2$, $MS12(i,j)$.

3-2 Product Cumulants of Sample Mean and Sample Variance

The sample variance is defined as $\bar{X}=\frac{S1}{N}=XB$ and the sample variance is

$$V=\frac{1}{N-1}\left(S2-\frac{S1}{N}\right) \tag{3.5}$$

We also introduce $N1=N-1$, then the joint moment of SB and V are

$$MXBV(i,j)=E(XB)^i(V)^j=\sum_{p=0}^{j}\binom{j}{p}\frac{(-1)^p}{N1^j}MS12(2p+i,j-p) \tag{3.6}$$

The subroutine $VM2K2$ gives the joint cumulant of XB and V, $KXBV(i,j)$ from $MXBV(j,j)$.

According to the semivariant property of the cumulants, the joint cumulants of $\sqrt{N}XB$ and $(V-1)$, $KNXBV1\ (i,j)$ are obtained easily which are input data of $VK2M2$ to calculate the joint moment of $\sqrt{N}XB$ and $(V-1)$.

3-3 Moments of t and Truncation in Computation

The moment of t, i.e., $MT(i)$ in programming notation is an infinite series of $1/N$.

We limit our interest to the first eight moments of t and to those terms of degree less than or equal to 3 in $1/N$.

The moment of t is

$$\begin{aligned} MT(i)&=E\left(\frac{\sqrt{N}\bar{X}}{S}\right)^i \\ &=\sum_{j=0}^{\infty}\binom{-\frac{i}{2}}{j}E(\sqrt{N}\,\bar{X})^i\,(S^2-1)^j \\ &=\sum\binom{-\frac{i}{2}}{j}E(NXB)^i\,(V1)^j=\sum\binom{-\frac{i}{2}}{j}MNXBV1(i,j) \end{aligned}$$

where $MNXBV1(i,j)$ is the (i,j)th moment of $\sqrt{N}\,\bar{X}$ and (S^2-1).

In order to avoid unnecessary computation of any terms of $MNXBV1(i,j)$ which would generate terms of degree higher than 3 in $1/N$ is needed.

First, the minimum degree of $1/N$ in $KXBV(i,j)$ is easily defined if we follow "the Rule 10" from Kendall [9] mentioned in Section 4 of chapter 3 and the formula $K(12^j)=K_{j+2/N^j}$ provided by Fisher [4].

According to these rules

$$\text{Min}_{1/N}(K(1^i2^j))=i+j-1 \quad \text{from } K(12^j)=K_{j+2/N^j}$$

$$\text{Min}_{1/\sqrt{N}}(K(1^i2^j))=2i+2j-2$$

$$\text{Min}_{1/\sqrt{N}}(KNXBV1(i,j))=i+2j-2$$

where $K(1^i,2^j)$ is (i,j)th cumulant of sample mean and sample variance and $KNXBV1(i,j)$ is the (i,j)th cumulant of $\sqrt{n}\,\bar{X}$ and S^2-1.

Second the degree of $1/N$ in $MNXBV1\ (i,j)$ is drawn from $KNXBV1\ (i,j)$ and the moment cumulant relation indicates that for each i, $i=1,2,\cdots,8$ we need to include the first six terms of $MXBV1\ (i,j)$ when i is an even number and the first five terms of $MNXBV1\ (i,j)$ where i is an odd number to meet our desired accuracy in $1/N$.

With $VM2K$, $MT(i)$ are converted to the cumulants of T, $KT(i)$.

4. Results and Discussion

The first eight cumulants of t are listed at table (4.1) in tabulated form. As an example, the first cumulant of t reads as follow.

$$KT(1)=\frac{1}{2\sqrt{N}}(-K_3)+\frac{1}{16(\sqrt{N})^3}(-6K_3+6K_5-15K_3K_4)$$
$$+\frac{1}{256(\sqrt{N})^5}\Big(-196K_3-120K_5-340\,K_3K_4+1120\,K_3^3$$
$$-80K_7+280K_3K_6+420K_4K_5-945K_3K_4^2\Big)$$

It is noted that, to the approximation used, the expression involves only the first eight cumulants of the population.

Computer output of moments of t after being expressed in terms of degree of freedom instead of the number of sample is in agreement with those from standard normal population when any terms having positive exponent of K_i $(i=3,4,\cdots,8)$ are eliminated.

Table 4.1. Cumulants of the Non-normal *t* Tabulated Form

	KT(1)			*KT*(3)			*KT*(5)		*KT*(7)
	1/2 $\sqrt{N}$	1/16 $\sqrt{N}$**3	1/256 $\sqrt{N}$**5	1 $\sqrt{N}$	1/8 $\sqrt{N}$**3	1/64 $\sqrt{N}$**5	2 $\sqrt{N}$**3	1/32 $\sqrt{N}$**5	1/8 $\sqrt{N}$**5
*K*3	−1	−6	−196	−2	−72	−2668	−60	−24240	−45360
*K*5		6	−120		24	216	6	3600	5040
*K*3**K*4		−15	−340		−30	−2232	20	−3440	16800
*K*3**3			1120		−83	3118	−105	−21120	−108360
*K*7			−80			−240		−480	−160
*K*3**K*6			280			280		−2480	−4368
*K*4**K*5			420			840		−440	−2240
*K*3**K*4**2			−945			−945		5520	3360
*K*3**2**K*5						4986		38880	46620
*K*3**3**K*4						−8973		−50940	−2520
*K*3**5								−30609	−135135

	KT(2)			*KT*(4)			*KT*(6)		*KT*(8)
	1/4 *N*	1/16 *N***2	1/128 *N***3	1 *N*	1/8 *N***2	1/9 *N***3	1 *N***2	1/8 *N***3	1 *N***3
1	8	96	2304	6	432	2928	240	34560	25200
*K*3**2	7	−6	3468	12	600	4872	840	109080	141120
*K*4		−32	−256	−2	−144	−976	−120	−17280	−16800
*K*3**K*5		−90	3048		−504	−1176	−210	−45720	−40320
*K*3**2**K*4		177	4860		648	11522	−150	103800	−8400
*K*3**4			−9680		699	−9942	1200	34020	244440
*K*6			512		32	288	16	3456	2688
*K*4**2			−768		−48	−912		−7680	−1400
*K*3**K*7			984			944		5736	1568
*K*5**2			750			732		4740	1680
*K*3**K*4**K*5			−7380			−5310		−9300	12040
*K*3**7**K*6			−2420			−1140		9780	15120
*K*3**2**K*4**2			10935			5175		−25290	−21840
*K*4**K*6						256		768	−560
*K*3**3**K*5						−9576		−198540	−158760
*K*3**4**K*4						16461		263700	65940
*K*8						−48		−384	−132
*K*4**3						−240			560
*K*3**6								103245	282240

In *KT* (2), add 1.

The interest thing is the third term of the sixth cumulant by Geary [7] is 577.5 λ_3^2 while the coefficient of λ_3^2 in the computer result shows 840.

ABSTRACT

The use of the statistic $t=\sqrt{n}(x-\mu)/S$, where $\bar{X}=\sum X_i/n$, $\mu=E(X_i)$, $S^2=\sum(X_i-\bar{X})^2/(n-1)$ in statistical inference is usually done under the assumption of normality of the population. If the population is not normally distributed the tabulated values of student t are no longer valid.

The moments of t are obtained as a power series in $1/\sqrt{n}$ whose coefficients are functions of the cumulants of X. The cumulants are obtained from the moments in the usual manner. The first eight cumulants of t are given up to terms of order $1/n^3$. These results extend those of Geary [7] who gave the first six cumulants of t to order $1/n^2$.

REFERENCES

[1] Bartlett, M. S., "The Effect of Non-normality on the t Distribution." *Proceedings of the Cambridge Philosophical Society,* 31, 223-231, 1935.

[2] Cook, M.B., "Bi-variate k-statistics and Cumulants of Their Joint Sampling Distribution," *Biometrika,* 38, 179, 1951.

[3] David, F.N., Kendall, M.G. and Barton, D.E., *Symmetric Functions and Allied Tables,* Cambridge University Press, 1966.

[4] Fisher, R.A., "Moments and Product Moments of Sampling Distributions." *Proceedings of the London Mathematical Society,* Series 2, Vol. 30, 1928.

[5] Gayen, A.K., "The Distribution of Student's t in Random Samples of Any Size Drawn from Non-normal Universe," *Biometrika,* 36, 363-369, 1949.

[6] Geary, R.C., "The Distribution of Student's t Ratio for Non-normal Samples," Supplement to the *Journal of the Royal Statistical Society,* Series

B, 3, 178-184, 1936.

[7] Geary, R.C., "Testing for Normality," *Biometrika*, 34, p. 209, 1947.

[8] Hwang, H., "The Non-normal *t* Distribution," Ph.D. Thesis, University of Minnesota, 1975.

[9] Kendall, M.G. and Stuart, A., *The Advanced Theory of Statistics*, Vol. 1, Hafner Publishing Company, New York, 1958.

[10] Richard, D. Heath, *FORPREP/ALMAP*, University Computer Center, University of Minnesota.

[11] White, J.S., *Moments, Cumulants and FORMAC*, Mathematics Department, Research Laboratories, General Motors Corporation, 1967.

[12] Zimmerman, C.D., "Extended Fortran Algebraic Manipulator with Applications to Linear Problems of Physics," Ph.D. Thesis, University of Minnesota, 1969.