The distribution of statistics drawn from the Gram-Charlier A type population

Uranishi, Hisao
Marine Institute, Tokyo

https://doi.org/10.5109/12945

出版情報：統計数理研究. 4 (1/2), pp.1-14, 1950-12. Research Association of Statistical Sciences

THE DISTRIBUTION OF STATISTICS DRAWN FROM THE GRAM-CHARLIER TYPE A POPULATION[1]

Hisao URANISI

(Marine Institute, Tokyo)

1. Introduction The current theories of statistical inferences chiefly concern themselves with the sampling from normal populations. There are, however, some current confidences that the ordinary test such as depending on t-distribution or F-distribution will hold true at least approximately, provided that the distribution of the sampled population is nearly normal. In this note we shall assume the distribution density of a population Π is expanded in a Gram-Charlier series of A type

$$(1.01)\qquad f(x)dx=\frac{e^{-\frac{1}{2}x^2}}{\sqrt{2\pi}}\{1+a_3H_3(x)+a_4H_4(x)+a_5H_5(x)+\dots\}dx,$$

where $x=(x-M)/\sigma$, with population mean M and standard deviation σ, and we shall give the sampling distributions of the statistic

$$(1.02)\qquad t=\frac{\sqrt{n}\,(\bar{x}-M)}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2}},$$

where $(X_1, X_2,\dots, X_n)$ is a random sample of size n from the population and X denotes the sample mean.

Furthermore we shall discuss the distribution of the statistic F defined as the ratio of two independent sample variances. The results will yield us some picture about the effect of non-normality on these fundamental statistics.

2. We shall prove at first the following

LEMMA 1. *Let Π be a parent population whose distribution density function $f(x)$ is given by (1.01), where $\{H_m(x)\}$ $(m=3,4,5,\dots)$ is a series of Hermite polynomials and the coefficints $\{a_m\}$ $(m=3,4,5,\dots)$ are given by the procedure of Gram-Charlier expansion of A-type, that is,*

$$(2.01)\qquad \begin{aligned} a_3&=\frac{\mu_3}{3!\,\sigma^3}=\frac{1}{6}\sqrt{\beta_1}\\ a_4&=\frac{1}{4!}\left(\frac{\mu_4}{\sigma_6}-3\right)=\frac{1}{24}(\beta_2-3)\\ a_5&=\frac{\mu_5-10\mu_2\mu_3}{120\,\sigma^5}\end{aligned}$$

(1) Communicated at the Annual Meeting of Math. Soc. of Japan, 2rd, June, 1950

$$a_6 = \frac{\mu_6 - 15\mu_2\mu_4 + 30\mu_2^3}{720\ \sigma^6} = \frac{1}{720}(\beta_4 - 15\beta_2 + 30)$$

and etc, where $\mu_i (i=2,3,...)$ *is the i-th moments of the population* Π.

Let O_n: $(x_1, x_2, ..., x_n)$ *be a sample of size n drawn from the population* Π, *and let us put*

(2.02) $$y_0 = \frac{1}{\sqrt{n}}(x_1 + x_2 + ... + x_n)$$

(2.03) $$\chi^2 = \sum_{i=1}^{n} (x_i - \bar{x})^2.$$

Then the joint distribution density function of the statistic (y_0, χ^2) *is given by*

(2.04) $$p(y_0, \chi) = \left(\frac{1}{\sqrt{2\pi}}\right)^n \exp\left\{-\frac{1}{2}(y^2 + \chi_0^2)\right\} \chi^{n-2} \Phi(y_0, \chi, a_3, a_4, ...),$$

where

(2.05) $$\Phi(y_0, \chi, a_3, a_4, ...) = \int_{-\pi}^{\pi} d\theta_{n-} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \left[\prod_{i=1}^{n} \{1 + a_3 H_3(x_i) + a_4 H_4(x_i) + ...\}\right]$$
$$\cdot \cos^{n-3}\theta_1 \cos^{n-4}\theta_2 ... \cos\theta_{n-3}\, d\theta_1 ... d\theta_{n-3}.$$

PROOF: The sampling distribution density function of sample of size n drawn from the population Π is given by

(2.06) $$p(x_1, x_2, ..., x_n) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n} x_i^2\right\} \prod_{i=1}^{n} \left\{1 + \sum_{m=3}^{\infty} a_m H_m(x_i)\right\},$$

with the normalized variate x.

The proof of our Lemma will be readily seen from the transformation

(2.07) $$\begin{cases} x_1 = \frac{1}{\sqrt{n}} y_0 + \frac{1}{\sqrt{2\cdot 1}} y_1 + \frac{1}{\sqrt{3\cdot 2}} y_2 + ... + \frac{1}{\sqrt{n(n-1)}} y_{n-1} \\ x_2 = \frac{1}{\sqrt{n}} y_0 - \frac{1}{\sqrt{2\cdot 1}} y_1 + \frac{1}{\sqrt{3\cdot 2}} y_2 + ... + \frac{1}{\sqrt{n(n-1)}} y_{n-1} \\ x_3 = \frac{1}{\sqrt{n}} y_0 \qquad - \frac{2}{\sqrt{3\cdot 2}} y_2 + ... + \frac{1}{\sqrt{n(n-1)}} y_{n-1} \\ \dots\dots\dots\dots \\ x_n = \frac{1}{\sqrt{n}} y_0 \qquad - \frac{n-1}{\sqrt{n(n-1)}} y_{n-1} \end{cases}$$

and the adoption of the polar coordinates defined by

$$(2.08)\quad \begin{cases} y_1=\chi\cos\theta_1\cos\theta_2\ldots\ldots\cos\theta_{n-3}\ \cos\theta_{n-2} \\ y_2=\chi\cos\theta_1\cos\theta_2\ldots\ldots\cos\theta_{n-3}\sin\theta_{n-2} \\ y_3=\chi\cos\theta_1\cos\theta_2\ldots\ldots\sin\theta_{n-3} \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ y_{n-2}=\chi\cos\theta_1\ \sin\theta_2 \\ y_{n-1}=\chi\sin\theta_1, \end{cases}$$

where

$$(2.09)\qquad \chi^2=\sum_{i=1}^{n}(x_i-\bar{x})^2=\sum_{i=1}^{n-1}y_i^2$$

with $x>0$ and $-\pi/2<\theta_i<\pi/2$ $(i=1,2,\ldots,n-3)$, $-\pi<\theta_{n-2}<\pi$.

Indeed we shall have

$$(2.10)\qquad p(x_1,x_2,\ldots x_n)dx_1,\ dx_2\ldots dx_n$$
$$=\left(\frac{1}{\sqrt{2\pi}}\right)^n\exp\left\{-\frac{1}{2}(y_0^2+x^2)\right\}\chi^{n-2}\cos^{n-3}\theta_1\cos^{n-4}\theta_2\ldots\cos\theta_{n-3}\cdot$$
$$\prod_{i=1}^{n}\{1+\sum_{m=3}^{\infty}a_mH_m(x_i)\ d\chi d\theta_1 d\theta_2\ldots d\theta_{n-3},$$

which leads to (2.03).

It is to be noted that $\Phi(y_0,\chi,a_3,a_4,\ldots)$ is a polynomial of the degree n at most with respect to each a_m $(m=3,4,\ldots)$. Our chief concerns are to deal with the cases when the distribution of Π is nearly normal. This means that the coefficients $\{a_m\}$ $(m=3,4,\ldots)$ are very small. In what follows we shall adopt the notation

$$(2.11)\qquad I[\varphi(y_1,y_2,\ldots,y_m)]$$
$$=\int_{-\pi}^{\pi}d\theta_{n-}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\cdots\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\varphi(y_1,y_2,\ldots,y_n)\cos^{n-3}\theta_1\cos\theta_2^{n-4}\ldots\cos\theta_{n-3}d\theta_1\ldots d\theta_{n-3}.$$

Then we shall have, by direct calculations,

LEMMA 2.

$$(2.12)\qquad I\ [1]=\frac{(\sqrt{2\pi})^{n-1}}{2^{\frac{n-3}{2}}\Gamma\left(\frac{n-1}{2}\right)}=C,\ \textit{say}.$$

$$(2.13)\qquad I\ [y_1^{2r_1}y_2^{2r_2}\ldots y_{n-1}^{2r_{n-1}}]=C\chi^{2m}\frac{\prod_{i=1}^{n-1}\{1\cdot3\cdot5\ldots(2r_i-1)\}}{(n-1)(n+1)(n+3)\ldots(n-3+2m)},$$

where $m=r_1+r_2+\ldots+r_{n-1}$.

If at least one of r_i $(i=1,2,\ldots,n-1)$ *is odd, then*

$$(2,14)\qquad I\ [y_1^{r_1}y_2^{r_2}\ldots y_{n-1}^{r_{n-1}}]=0,$$

COROLLARY 1.

$$(2.15)\quad I\left[\sum_{i=1}^{n}(x_i-\bar{x})^{2m+1}\right]=0.$$

$$(2.16)\quad I\left[\sum_{i=1}^{n}(x_i-\bar{x})^{2m}\right]=C\chi^{2m}\left(\frac{n-1}{n}\right)^{m-1}\frac{1\cdot3\cdot5\ldots(2m-1)}{(n+1)(n+3)\ldots(n-3+2m)}.$$

LEMMA 3.

$$(2.17)\quad \Phi(y_0,\chi,0,0,0,\ldots)=C=\frac{(\sqrt{2\pi})^{n-1}}{2^{\frac{n-3}{2}}\Gamma\left(\frac{n-1}{2}\right)}$$

$$(2.18)\quad \left(\frac{\partial\Phi(y_0,\chi,a_3,a_4,\ldots)}{\partial a_3}\right)_0=C\left\{\frac{y_0^3}{\sqrt{n}}+\frac{3y_0}{\sqrt{n}}\chi^2-3\sqrt{n}\,y_0\right\}$$

$$(2.19)\quad \left(\frac{\partial\Phi(y_0,\chi,a_3,a_4,\ldots)}{\partial a_4}\right)_0$$
$$=C\left\{\frac{y_0^4}{n}+\frac{6y_0^2}{n}\chi^2+\frac{3(n-1)}{n(n+1)}\chi^4-6(y_0^2+\chi^2)+3n\right\}$$

$$(2.20)\quad \left(\frac{\partial\Phi(y_0,\chi,a_3,a_4,\ldots)}{\partial a_5}\right)_0=\frac{C}{\sqrt{n}}\left\{\frac{y_0^5}{n}+\frac{10y_0^3}{n}\chi^2\right.$$
$$\left.+\frac{15(n-1)y_0}{n(n+1)}\chi^4-10(y_0^3+3y_0\chi^2)+15ny_0\right\}$$

$$(2.21)\quad \left(\frac{\partial\Phi(y_0,\chi,a^3,a_4,\ldots)}{\partial a_6}\right)_0=C\left\{\frac{y_0^6}{n}+\frac{15y_0^4}{n^2}\chi^2+\frac{45(n-1)y_0^2}{n^2(n+1)}\chi^4\right.$$
$$+\frac{15(n-1)^2}{n^2(n+1)(n+3)}\chi^6-15\left(\frac{y_0^4}{n}+\frac{6y_0^2}{n}\chi^2\right.$$
$$\left.\left.+\frac{3(n-1)}{n(n+1)}\chi^4\right)+45(y_0^2+\chi^2)-15n\right\}$$

$$(2.22)\quad \left(\frac{\partial^2\Phi(y_0,\chi,a_3,a_4,\ldots)}{\partial a_3^2}\right)_0=C\left\{\frac{(n-1)y_0^6}{n^2}+\frac{3(2n-5)}{n^2}\chi^2\right.$$
$$+\frac{9(n^2-4n+5)y_0^2}{n^2(n+1)}\chi^4-\frac{3(3n^2-6n+5)}{n^2(n+1)(n+3)}\chi^6$$
$$\left.-6\left(\frac{n-1}{n}y_0^4+\frac{3(n-2)}{n}y_0^2\chi^2-\frac{3(n-1)}{n(n+1)}\chi^4\right)+9(n-1)y_0^2-9\chi^2\right\}$$

$$(2.23)\quad \left(\frac{\partial^2\Phi(y_0,\chi,a_3,a_4,\ldots)}{\partial a_3\,\partial a_4}\right)_0=\frac{C}{\sqrt{n}}\left\{\frac{(n-1)y_0^7}{n^2}+\frac{3(3n-7)y_0^5}{n^2}\chi^2\right.$$
$$+\frac{3(7n^2-30n+35)y_0^3}{n^2(n+1)}\chi^4+\frac{3(3n^3-21n^2+45n-35)y_0}{n^2(n+1)(n+3)}\chi^6$$
$$-3\left(\frac{3(n-1)y_0^5}{n}+2\frac{(7n-15)y_0^3}{n}\chi^2+\frac{3(3n^2-14n+15)y_0}{n(n+1)}\chi^4\right)$$
$$\left.+21(n-1)y_0^3+9(3n-7)y_0\chi^2-9n(n-1)y_0\right\}$$

and

$$(2.24)\quad \left(\frac{\partial^2\Phi(y_0,\chi,a_3,a_4,\dots)}{\partial a_4^2}\right)_0$$

$$= C\Big\{\frac{n-1}{n^3}y_0^8+\frac{4(3n-7)}{n^3}y_0^6\chi^2+\Big(\frac{36}{n}+\frac{6(n-1)}{n^2(n+1)}-\frac{210(n-1)}{n^3(n+1)}\Big)y_0^4\chi^4$$

$$+\Big(\frac{36(n-1)}{n^2(n+1)}+\frac{96(n-2)}{n^2(n+1)(n+3)}-\frac{420(n-1)^2}{n^3(n+1)(n+3)}\Big)y_0^2\chi^6$$

$$+\frac{3(3n^4-12n^3+42n^2-60n+35)}{n^3(n+1)(n+3)(n+5)}\chi^8$$

$$+12\Big(\frac{y_0^6}{n^2}+\frac{15y_0^4}{n^2}\chi^2+\frac{45(n-1)y_0^2}{n^2(n+1)}\chi^4+\frac{15(n-1)^2}{n^2(n+1)(n+3)}\chi^6\Big)$$

$$-12\Big(\frac{y_0^4}{n}+\frac{6y_0^2\chi^2}{n}+\frac{3(n-1)\chi^4}{n(n+1)}\Big)(y_0^2+\chi^2)+36(y_0^2+\chi^2)^2$$

$$+\frac{6(n-7)}{n}y_0^4+\frac{36(n-7)}{n}y_0^2\chi^2+\frac{18(n-1)(n-7)}{n(n+1)}\chi^4$$

$$-36(n-1)(y_0^2+\chi^2)+9n(n-1)\Big\},$$

where suffix o *means that we put all* a_m $(m=3,4,\dots)$ *as zeros.*

In combination of these three Lemmas, we shall reach

THEOREM 1. *The probability density function* $p(t)$ *of the statistic* t *defined by* (1.02) *from a parent population* Π *whose distribution is* (1.01) *can be approximated in the following formula:*

$$(2.25)\quad p(t)=p_0(t)+a_3p_3(t)+a_4p_4(t)+a_5p_5(t)+a_6p_6(t)+\dots$$
$$+\frac{a_3^2}{2}p_{3,3}(t)+a_3a_4p_{3,4}(t)+\frac{a_4^2}{2}p_{4,4}(t)+\dots,$$

where $p_0(t)$ *concides with the t-distribution density function with degree of freedom* $n-1$ *from a normal population and*

$$(2.26)\quad p_3(t)=\frac{t}{\sqrt{2n\pi}}\left\{-\frac{2n-1}{\left(1+\frac{t^2}{n-1}\right)^{\frac{n+1}{2}}}+\frac{2(n+1)}{\left(1+\frac{t^2}{(n-1)}\right)^{\frac{n+3}{2}}}\right\}$$

$$(2.27)\quad p_4(t)=\frac{-2}{\sqrt{n-1}\,B\left(\frac{n-1}{2},\frac{1}{2}\right)}\left\{\frac{n-1}{\left(1+\frac{t^2}{n-1}\right)^{\frac{n}{2}}}\right.$$
$$\left.-\frac{2(n+2)}{\left(1+\frac{t^2}{n-1}\right)^{\frac{n+2}{2}}}+\frac{(n+2)(n+4)}{(n+1)\left(1+\frac{t^2}{n-1}\right)^{\frac{n+4}{2}}}\right\}$$

$$(2.28)\quad p_5(t)=\frac{3t}{n\sqrt{2n\pi}}\left\{\frac{2n^2-2n+1}{\left(1+\frac{t^2}{n-1}\right)^{\frac{n+1}{2}}}-\frac{4(n+1)(n-2)}{\left(1+\frac{t^2}{n-1}\right)^{\frac{n+3}{2}}}+\frac{2(n+3)(n-4)}{\left(1+\frac{t^2}{n-1}\right)^{\frac{n+5}{2}}}\right\}$$

$$(2.29)\quad p_6(t)=\frac{8}{n\sqrt{n-1}\,B\left(\frac{n-1}{2},\frac{1}{2}\right)}\left\{\frac{(n-1)(2n-1)}{\left(1+\frac{t^2}{n-1}\right)^{\frac{n}{2}}}-\frac{6(n-1)(n+2)}{\left(1+\frac{t^2}{n-1}\right)^{\frac{n+2}{2}}}\right.$$

$$\left.+\frac{3(2n-3)(n+2)(n+4)}{(n+1)\left(1+\frac{t^2}{n-1}\right)^{\frac{n+4}{2}}}-\frac{2(n-2)(n+2)(n+4)(n+6)}{(n+1)(n+3)\left(1+\frac{t^2}{n-1}\right)^{\frac{n+6}{2}}}\right\}$$

$$(2.30)\quad p_{3,3}(t)=\frac{2}{n\sqrt{n-1}\,B\left(\frac{n-1}{2},\frac{1}{2}\right)}\left\{\frac{(n-1)(2n^2-3n+4)}{\left(1+\frac{t^2}{n-1}\right)^{\frac{n}{2}}}-\frac{3(2n^3-3n^2+16)}{\left(1+\frac{t^2}{n-1}\right)^{\frac{n+2}{2}}}\right.$$

$$\left.+\frac{3(n+2)(2n^3-3n^2-8n+48)}{(n+1)\left(1+\frac{t^2}{n-1}\right)^{\frac{n+4}{2}}}-\frac{(n+2)(n+4)(2n^3-3n^2-20n+96)}{(n+1)(n+3)\left(1+\frac{t^2}{n-1}\right)^{\frac{n+6}{2}}}\right\}$$

$$(2.31)\quad p_{3,4}(t)=\frac{t}{n^2\sqrt{2n\pi}}\left\{\frac{(n-1)(4n^3-6n^2-4n+15)}{\left(1+\frac{t^2}{n-1}\right)^{\frac{n+1}{2}}}\right.$$

$$-\frac{6(n+1)(2n^3+2n^2-27n+45)}{\left(1+\frac{t^2}{n-1}\right)^{\frac{n+3}{2}}}+\frac{6(n+3)(2n^3+7n^2-67n+120)}{\left(1+\frac{t^2}{n-1}\right)^{\frac{n+5}{2}}}$$

$$\left.-\frac{4(n+5)(n^3+6n^2-61n+120)}{\left(1+\frac{t^2}{n-1}\right)^{\frac{n+7}{2}}}\right\}$$

and

$$(2.32)\quad p_{4,4}(t)=\frac{4}{n^2\sqrt{n-1}\,B\left(\frac{n-1}{2},\frac{1}{2}\right)}\left\{\frac{(n-1)(n^3+6n^2-13n+12)}{\left(1+\frac{t^2}{n-1}\right)^{\frac{n}{2}}}\right.$$

$$-\frac{4(n+2)(n^3+10n^2-33n+36)}{\left(1+\frac{t^2}{n-1}\right)^{\frac{n+2}{2}}}$$

$$+\frac{6(n+2)(n+4)(n^3+14n^2-57n+72)}{(n+1)\left(1+\frac{t^2}{n-1}\right)^{\frac{n+4}{2}}}$$

$$-\frac{4(n+2)(n+4)(n+6)(n^3+18n^2-85n+120)}{(n+1)(n+3)\left(1+\frac{t^2}{n-1}\right)^{\frac{n+6}{2}}}$$

$$\left.+\frac{(n+2)(n+4)(n+6)(n+8)(n^3+22n^2-117n+180)}{(n+1)(n+3)(n+5)\left(1+\frac{t^2}{n-1}\right)^{\frac{n+8}{2}}}\right\}.$$

Fnrthermore, for any assigned positive number τ,

$$(2.33)\quad P(\tau)\equiv P_r\{t>\tau\}$$

$$=2\left\{\int_\tau^\infty p_0(t)dt+a_4\int_\tau^\infty p_4(t)dt+a_6\int_\tau^\infty p_6(t)dt+\dots\right\}$$

$$+a_3^2\int_\tau^\infty p_{3,3}(t)dt+a_4^2\int_\tau^\infty p_{4,4}(t)dt+\ldots\ldots$$

It is to be remarked that each of the functions (2.26) (2.32) is an odd or even function of t according to the fact the associated index (the sum of two indexes when there are two indexes) is odd or even. This fact simplifies the approximation of the integral $P(\tau)$ as given in (2.33).

We shall note also

COROLLARY 1. *Under the hypothesis to Theorem* 1, *we shall have the following approximate formula*

$$P(\tau)\equiv Pr\{t>\tau\} \cong 2\int_\tau^\infty p_0(t)dt+\frac{\beta_1}{36}\int_\tau^\infty p_{3,3}(t)dt. \tag{2.34}$$

3. The distributions of χ^2 and F.

In view of Lemma, we shall have

THEOREM 2. *Under the hypothesis to Theorem* 1 *the probability density function* $p(\chi^2)$ *of the statistic* χ^2 *defined in* (2.03) *become*

$$\begin{aligned} p(\chi^2)=p_0(\chi^2)\Big\{1&+\frac{3(n-1)^2a_4}{n}\Big(1-\frac{2\chi^2}{n-1}+\frac{\chi^4}{(n-1)(n+1)}\Big)\\ &-\frac{15(n-1)^3a_6}{n^2}\Big(1-\frac{3\chi^2}{n-1}+\frac{3\chi^4}{(n-1)(n+1)}-\frac{\chi^6}{(n-1)(n+1)(n+3)}\Big)\\ &+\frac{3(n-1)(3n^2-6n+5)a_3^2}{2n^2}\Big(1-\frac{3\chi^2}{n-1}+\frac{3\chi^4}{(n-1)(n+1)}\\ &-\frac{\chi^6}{(n-1)(n+1)(n+3)}\Big)+\frac{3(n-1)(3n^4-12n^3+42n^2-60n+35)a_4^2}{2n^3}\\ &\cdot\Big(1-\frac{4\chi^2}{n-1}+\frac{6\chi^4}{(n-1)(n+1)}-\frac{4\chi^6}{(n-1)(n+1)(n+3)}\\ &+\frac{\chi^8}{(n-1)(n+1)(n+3)(n+5)}\Big)+\ldots\ldots\Big\} \end{aligned} \tag{3.01}$$

where $p(\chi^2)$ *is the usnal* χ^2*-distribution with degree of freedom* $n-1$ *from the normal distribution*

$$p_0(\chi^2)=\frac{1}{2^{\frac{n-1}{2}}\Gamma\left(\frac{n-1}{2}\right)}\exp\left\{-\frac{1}{2}\chi^2\right\}\{\chi^2\}^{\frac{n-1}{2}-1}. \tag{3.02}$$

Furthermore we shall add

THEOREM 3. *Let* O_n: $(x_1, x_2,\ldots, x_n)$ *be a random sample of size* n *from a parent population* Π *with distribution function assumed in Theorem* 1. *Let* O_m': $(x_1', x_2',\ldots, x_m')$ *be a random sample of size* m *from another population* Π' *with distribution function which can also be expanded into* *a Gram-Charlier series of the type*

$$(3.03)\qquad g(x')dx'=\left(\frac{1}{\sqrt{2\pi}}\right)^{n}\exp\left\{-\frac{1}{2}\chi'^{2}\right\}\left\{1+\sum_{k=1}^{\infty}b_{k}H_{k}(\chi')\right\},$$

where $x'=(x'-M')/\alpha'$, M' *and* σ' *being the population mean and standard deviation of* Π'.

Let us define F *by*

$$(3.04)\qquad F=\frac{\frac{1}{n-1}\sum_{i=1}^{n}(x_{i}-\bar{x})^{2}}{\frac{1}{n-1}\sum_{i=1}^{m}(x_{j}'-\bar{x}')^{2}}$$

with sample means X *and* X' *of* O_n *and* O_m' *respectively.*

Then the probability density function $p(F)$ *of the statistic* F *becomes*

$$(3.05)\qquad p(F)=\frac{\left(\frac{n-1}{m-1}\right)^{\frac{n-1}{2}}F^{\frac{n-1}{2}-1}}{B\left(\frac{n-1}{2},\frac{m-1}{2}\right)\left(1+\frac{n-1}{m-1}F\right)^{\frac{n+m-2}{2}}}\cdot[1+A_{1}B_{1}+A_{2}B_{2}+A_{3}B_{3}+A_{4}B_{4}],$$

where we put for the moment

$$(3.051)\qquad A_{1}=\frac{3(n-1)^{2}}{n}a_{4}$$

$$(3.052)\qquad B_{1}=1-\frac{2(n+m-2)\frac{n-1}{m-1}F}{(n-1)\left(1+\frac{n-1}{m-1}F\right)}+\frac{(n-m)(n+m-2)\left(\frac{n-1}{m-1}F\right)^{2}}{(n-1)(n+1)\left(1+\frac{n-1}{m-1}F\right)^{2}}$$

$$(3.053)\qquad A_{2}=\frac{-30(n-1)^{3}a^{6}+3(n-1)(3n^{2}-6n+5)a_{3}^{2}}{2n^{2}}$$

$$(3.054)\qquad B_{2}=1-\frac{3(n+m-2)\frac{n-1}{m-1}F}{(n-1)\left(1+\frac{n-1}{m-1}F\right)}+\frac{3(n+m-2)(n+m)\left(\frac{n-1}{m-1}F\right)^{2}}{(n-1)(n-1)\left(1+\frac{n-1}{m-1}F\right)^{2}}$$
$$-\frac{(n+m-2)(n+m)(n+m-2)}{(n-1)(n+1)(n+3)}\cdot\frac{\left(\frac{n-1}{m-1}F\right)^{3}}{\left(1+\frac{n-1}{m-1}F\right)^{3}}$$

$$(3.055)\qquad A_{3}=\frac{3(m-1)^{2}}{m}b_{4}$$

$$(3.056)\qquad B_{3}=1-\frac{2(m+n-2)}{(m-1)\left(1+\frac{n-1}{m-1}F\right)}+\frac{(m+n-2)(m+n)}{(n-1)(m+1)\left(1+\frac{n-1}{m-1}F\right)^{2}}$$

$$(3.057)\qquad A_{4}=+\frac{-30(m-1)^{3}b_{6}+3(m-1)(3m^{2}-6m+5)b_{3}^{3}}{2m^{2}}$$

$$(3.058)\qquad B_4=1-\frac{3(m+n-2)}{(m-1)\left(1+\frac{n-1}{m-1}F\right)}+\frac{3(n+n-1)(m+n-2)}{(m-1)(m+1)\left(1+\frac{n-1}{m-1}F\right)^2}$$
$$-\frac{(m-n-2)(n+m-2)(m+n)}{(m-1)(m+1)(m+3)\left(1+\frac{n-1}{m-1}F\right)^3}.$$

4. Numerical considerations

The effects of non-normality on the t-distribution will be discussed in this § in view of our Theorem 1.

We shall enunciate the following

LEMMA 4. *Let us put*

$$(4.01)\qquad u=\left(1+\frac{t^2}{n-1}\right)^{-1}.$$

Then we have

$$(4.02)\qquad 2\int_\tau^\infty p_0(t)dt=I_u\left(\frac{n-1}{2},\frac{1}{2}\right)$$

$$(4.03)\qquad 2\int_\tau^\infty p_4(t)dt=\frac{-2(n-1)}{n}\left\{nI_u\left(\frac{n-1}{2},\frac{1}{2}\right)\right.$$
$$\left.-(2n+4)I_u\left(\frac{n+1}{2},\frac{1}{2}\right)+(n+4)I_u\left(\frac{n+3}{2},\frac{1}{2}\right)\right\}$$

$$(4.04)\qquad 2\int_\tau^\infty p_6(t)dt=\frac{8(n-1)}{n^2}\left\{n(2n-1)I_u\left(\frac{n-1}{2},\frac{1}{2}\right)-6(n-1)(n+2)I_u\left(\frac{n+1}{2},\frac{1}{2}\right)\right.$$
$$\left.+3(n+4)(2n-3)I_u\left(\frac{n+3}{2},\frac{1}{2}\right)-2(n-2)(n+6)I_u\left(\frac{n+5}{2},\frac{1}{2}\right)\right\}$$

$$(4.05)\qquad \int_\tau^\infty p_{3,3}(t)dt=\frac{n-1}{n^2}\left\{(2n^3-3n_2+4n)\,I_u\left(\frac{n-1}{2},\frac{1}{2}\right)\right.$$
$$-3(2n^3-3n^2+16)I_u\left(\frac{n+1}{2},\frac{1}{2}\right)+3(2n^3-3n^2-8n+48)\cdot$$
$$\left.\cdot I_u\left(\frac{n+3}{2},\frac{1}{2}\right)-(2n^3-3n^2-20n+96)\,I_u\left(\frac{n+5}{2},\frac{1}{2}\right)\right\},$$

and

$$(4.06)\qquad \int_\tau^\infty p_{4,4}(t)dt=\frac{2(n-1)}{n^3}\left\{n(n^3+6n^2-13n+12)I_u\left(\frac{n-1}{2},\frac{1}{2}\right)\right.$$
$$-4(n+2)(n^3+10n^2-33n+36)I_u\left(\frac{n+1}{2},\frac{1}{2}\right)$$
$$+6(n+4)(n^3+14n^2-57n+72)I_u\left(\frac{n+3}{2},\frac{1}{2}\right)$$

$$-4(n+6)(n^3+18n^2-85n+120)I_u\left(\frac{n+5}{2},\frac{1}{2}\right)$$

$$+(n+8)(n^3+22n^2-117n+180)I_u\left(\frac{n+7}{2},\frac{1}{2}\right)$$

$$(4.07)\quad \int_\tau^\infty p_3(t)dt=\frac{1}{\sqrt{2n\pi}}\left\{-(2n-1)u^{\frac{n-1}{2}}+2(n-1)u^{\frac{n+1}{2}}\right\}$$

$$(4.08)\quad \int_\tau^\infty p_5(t)dt=\frac{3}{n\sqrt{2n\pi}}\left\{(2n^2-2n+1)u^{\frac{n-1}{2}}\right.$$

$$\left.-4(n^2-3n+2)u^{\frac{n+1}{2}}+(2n^2-10n+8)u^{\frac{n+3}{2}}\right\},$$

and

$$(4.09)\quad \int_\tau^\infty p_{3,4}(t)dt=\frac{n-1}{n^2\sqrt{2n\pi}}\left\{(4n^3-6n^2-4n-15)u^{\frac{n-1}{2}}\right.$$

$$-6(2n^3+2n^2-27n+45)\cdot u^{\frac{n+1}{2}}+6(2n^3+7n^2-67n+120)u^{\frac{n+3}{2}}$$

$$\left.-4(n^3+6n-61n+120)u^{\frac{n+5}{2}}\right\}.$$

LEMMA 5. *Under the hypothesis to Theorem 2, we shall have*

$$(4.10)\quad \int_F^\infty p(F)dF=I_u\left(\frac{m-1}{2},\frac{n-1}{2}\right)+\frac{3(n-1)^2a_4}{n}\left\{I_u\left(\frac{m-1}{2},\frac{n-1}{2}\right)\right.$$

$$\left.-2I_u\left(\frac{m-1}{2},\frac{n-1}{2}\right)+I_u\left(\frac{m-1}{2},\frac{n+3}{2}\right)\right\}$$

$$+\frac{-30(n-1)^3a_6+3(n-1)(3n^2-6n+5)a_3^2}{2n^2}\left\{I_u\left(\frac{m-1}{2},\frac{n-1}{2}\right)\right.$$

$$\left.-3I_u\left(\frac{m-1}{2},\frac{n+1}{2}\right)+3I_u\left(\frac{m-1}{2},\frac{n+3}{2}\right)-I_u\left(\frac{m-1}{2},\frac{n+5}{2}\right)\right\}$$

$$+\frac{3(m-1)^2b_4}{m}\left\{I_u\left(\frac{m-1}{2},\frac{n-1}{2}\right)-2I_u\left(\frac{m+1}{2},\frac{n-1}{2}\right)+I_u\left(\frac{m+3}{2},\frac{n-1}{2}\right)\right\}$$

$$+\frac{-30(m-1)^3b_6+3(m-1)(3m^2-6m+5)b_3^2}{2m^2}\left\{I_u\left(\frac{m-1}{2},\frac{n-1}{2}\right)\right.$$

$$\left.-3I_u\left(\frac{m+1}{2},\frac{n-1}{2}\right)+3I_u\left(\frac{m+3}{2},\frac{n-1}{2}\right)-I_u\left(\frac{m+5}{2},\frac{n-1}{2}\right)\right\}.$$

In view of these two lemmas we have computed the integrals which we mention in Table 1 and 2 for the cases when $n=5$, 10, 15 and 21.

Under our hypothesis to the distribution of parent populations the effects of non-normality may be summarized as follows:

[1] The effect of non-normality on the distribution of $|t|$. Here we shall assume that distribution of parent population belongs to some class of Pearson type. We have calculated the following examples by virtue of Table 1.

TABLE 1. Table for calculating approximate value of $Pr.\{|t|>t_0\}$.

$$Pr.\{|t|>t_0\}=2\int_{t_0}^{\infty}p_0(t)dt+\frac{\beta_2-3}{12}\int_{t_0}^{\infty}p_4(t)dt+\frac{\beta_4-15\beta_2+30}{360}\int_{t_0}^{\infty}p_6(t)dt+\frac{\beta_1}{36}\int_{t_0}^{\infty}p_{3,3}(t)dt$$

$$+\frac{(\beta_2-3)^2}{576}\int_{t_0}^{\infty}p_{4,4}(t)dt, \qquad u=\left(1+\frac{t_0^2}{n-1}\right)^{-1}.$$

Multiplier		β_2-3	$\frac{1}{10}(\beta_4-15\beta_2+30)$	β_1	$(\beta_2-3)^2$
n	u	$\int_{t_0}^{\infty}p_4(t)dt/12$	$\int_{t_0}^{\infty}p_6(t)dt/36$	$\int_{t_0}^{\infty}p_{3,3}(t)dt/36$	$\int_{t_0}^{\infty}p_{4,4}(t)dt/576$
	0.05	−0.000 282	0.000 650	0.000 729	—
	0.10	−0.001 008	0.002 228	0.002 605	0.000 272
	0.20	−0.003 113	0.006 303	0.008 168	0.000 556
	0.30	−0.005 177	0.009 328	0.014 021	0.000 479
	0.40	−0.006 197	0.009 667	0.018 374	0.000 068
5	0.50	−0.005 524	0.006 813	0.020 233	−0.000 439
	0.60	−0.002 846	0.001 252	0.019 311	−0.000 753
	0.70	+0.001 677	−0.005 670	0.015 923	−0.000 671
	0.80	+0.007 155	−0.011 830	0.010 876	−0.000 183
	0.90	+0.011 206	−0.014 120	0.005 335	+0.000 436
	0.95	+0.010 721	−0.012 031	0.002 785	+0.000 579
	1.00	0.000 000	0.000 000	0.000 000	0.000 000
	0.05	—	—	—	—
	0.10	−0.000 005	—	—	—
	0.20	−0.000 093	0.000 191	0.000 486	0.000 028
	0.30	−0.000 445	0.000 811	0.002 235	0.000 081
	0.40	−0.001 195	0.001 877	0.005 800	0.000 090
10	0.50	−0.002 205	0.002 872	0.010 625	−0.000 029
	0.60	−0.002 912	0.002 877	0.014 915	−0.000 250
	0.70	−0.002 329	0.001 085	0.016 191	−0.000 382
	0.80	+0.000 578	−0.002 375	0.012 567	−0.000 163
	0.90	+0.005 555	−0.005 458	0.004 642	+0.000 387
	0.95	+0.007 201	−0.005 357	0.000 625	+0.000 561
	1.00	0.000 000	0.000 000	0.000 000	0.000 000
	0.05	—	—	—	—
	0.10	—	—	—	—
	0.20	−0.000 002	—	—	—
	0.30	−0.000 029	0.000 053	0.000 211	0.000 008
	0.40	−0.000 163	0.000 257	0.001 109	0.000 024
15	0.50	−0.000 549	0.000 724	0.003 491	0.000 022
	0.60	−0.001 244	0.001 294	0.007 556	−0.000 051
	0.70	−0.001 860	0.001 322	0.011 705	−0.000 182
	0.80	−0.001 146	−0.000 048	0.012 168	−0.000 176
	0.90	+0.002 541	−0.002 466	0.005 706	+0.000 184
	0.95	+0.004 889	−0.002 959	0.000 908	+0.000 370
	1.00	0.000 000	0.000 000	0.000 000	0.000 000
	0.05	—	—	—	—
	0.10	—	—	—	—
	0.20	—	—	—	—
	0.30	−0.000 001	0.000 001	0.000 007	—
	0.40	−0.000 013	0.000 020	0.000 118	0.000 003
21	0.50	−0.000 087	0.000 115	0.000 723	0.000 008
	0.60	−0.000 357	0.000 378	0.002 683	0.000 001
	0.70	−0.000 929	0.000 715	0.006 481	−0.000 060
	0.80	−0.001 282	0.000 507	0.009 831	−0.000 127
	0.90	+0.000 736	−0.000 996	0.006 531	+0.000 053
	0.95	+0.003 127	−0.001 664	0.001 625	+0.000 223
	1.00	0.000 000	0.000 000	0.000 000	0.000 000

TABLE 2. Table for calculating the approximate value of $Pr.\{t>t_0\}$.

$$Pr.\{t>t_0\}=\frac{1}{2}Pr.\{|t|>t_0\}+\frac{\sqrt{\beta_1}}{6}\int_{t_0}^{\infty}p_3(t)dt+\frac{\mu_5-10\mu_2\mu_3}{120\sigma^5}\int_{t_0}^{\infty}p_5(t)dt$$

$$+\frac{\sqrt{\beta_1}(\beta_2-3)}{144}\int_{t_0}^{\infty}p_{3,4}(t)dt\ , \qquad u=\left(1+\frac{t_0^2}{n-1}\right)^{-1}.$$

Multiplier		1	$\sqrt{\beta_1}$	$\frac{\mu_5}{10\sigma^5}-\sqrt{\beta_1}$	$\sqrt{\beta_1}(\beta_2-3)$
n	u	$\int_{t_0}^{\infty}p_0(t)dt$	$\int_{t_0}^{\infty}p_3(t)dt/6$	$\int_{t_0}^{\infty}p_5(t)dt/12$	$\int_{t_0}^{\infty}p_{3,4}(t)dt/144$
5	0.05	0.000 477	−0.000 639	0.000 860	0.000 142
	0.10	0.001 941	−0.002 438	0.003 236	0.000 458
	0.20	0.008 065	−0.008 802	0.011 318	0.001 115
	0.30	0.018 921	−0.017 663	0.021 934	0.001 261
	0.40	0.035 242	−0.027 594	0.032 942	0.000 620
	0.50	0.058 058	−0.037 170	0.042 372	−0.000 743
	0.60	0.088 904	−0.044 960	0.048 428	−0.002 509
	0.70	0.130 287	−0.049 539	0.049 481	−0.004 184
	0.80	0.186 951	−0.049 480	0.044 075	−0.005 191
	0.90	0.270 735	−0.043 354	0.030 927	−0.004 952
	0.95	0.335 090	−0.037 571	0.021 093	−0.004 206
	1.00	0.500 000	−0.029 735	0.008 921	−0.002 974
10	0.05	—	−0.000 001	0.000 001	—
	0.10	0.000 004	−0.000 012	0.000 016	0.000 006
	0.20	0.000 101	−0.000 232	0.000 288	0.000 083
	0.30	0.000 662	−0.001 036	0.001 192	0.000 247
	0.40	0.002 561	−0.004 017	0.004 242	0.000 530
	0.50	0.007 478	−0.009 292	0.008 921	0.000 419
	0.60	0.018 394	−0.017 309	0.014 906	−0.000 396
	0.70	0.040 563	−0.027 032	0.020 476	−0.001 932
	0.80	0.083 925	−0.035 434	0.022 785	−0.003 532
	0.90	0.171 718	−0.036 645	0.018 217	−0.003 817
	0.95	0.254 323	−0.031 715	0.012 193	−0.002 942
	1.00	0.500 000	−0.021 026	0.003 154	−0.001 183
15	0.05	—	—	—	—
	0.10	—	—	—	—
	0.20	0.000 002	−0.000 005	0.000 006	0.000 003
	0.30	0.000 027	−0.000 077	0.000 086	0.000 033
	0.40	0.000 213	−0.000 501	0.000 504	0.000 134
	0.50	0.001 095	−0.002 012	0.001 798	0.000 272
	0.60	0.004 282	−0.005 863	0.004 570	0.000 193
	0.70	0.014 036	−0.013 290	0.008 812	−0.000 518
	0.80	0.041 209	−0.023 762	0.012 860	−0.001 924
	0.90	0.116 389	−0.031 203	0.012 547	−0.002 821
	0.95	0.202 566	−0.028 773	0.008 837	−0.002 279
	1.00	0.500 000	−0.017 168	0.001 717	−0.000 668
21	0.05	—	—	—	—
	0.10	—	—	—	—
	0.20	—	—	—	—
	0.30	0.000 001	−0.000 002	0.000 002	0.000 002
	0.40	0.000 012	−0.000 038	0.000 037	0.000 017
	0.50	0.000 117	−0.000 298	0.000 254	0.000 074
	0.60	0.000 793	−0.001 491	0.001 089	0.000 160
	0.70	0.004 161	−0.005 328	0.003 227	0.000 021
	0.80	0.018 452	−0.014 021	0.006 699	−0.000 812
	0.90	0.075 820	−0.025 296	0.008 600	−0.001 953
	0.95	0.158 576	−0.026 062	0.006 640	−0.001 776
	1.00	0.500 000	−0.014 509	0.001 036	−0.000 411

Here $t_{n-1}(\alpha)$ denotes the α-significance level of the ordinary t-distribution with degree of freedom $n-1$.

TABLE 3. $pr\{|t|>t_{n-1}(\alpha)\}\equiv\alpha'$

Parent population	$\alpha=1\%$	$\alpha=5\%$
$\beta_1=0$, $\beta_2=4$, $\beta_3=40$	1.3% ($n=5$) 1.1% ($n=10$) 1.0% ($n=15$) 1.0% ($n=21$)	5.4% ($n=5$) 4.9% ($n=10$) 4.9% ($n=15$) 4.9% ($n=21$)
$\beta_1=0$, $\beta_2=2.5$, $\beta_3=9$	1. 2% ($n=5$) 1. 1% ($n=10$) 1. 1% ($n=15$) 1.05% ($n=21$)	5. 5% ($n=5$) 5. 2% ($n=10$) 5. 1% ($n=15$) 5.05% ($n=21$)
$\beta_1=1$ $\beta_2=3$ $\beta_3=6$ $\beta_4=15$	1.6%~ 1.8% (n=5,10, 15,21)	6.0%~ 6.6% (n=5,10, 15,21)
$\beta_1=0.5$, $\beta_2=2.5\sim3.5$	1.5% (n=5,10, 15,21)	5.5%~ 6.2% (n=5,10, 15,21)

The effects of non-normality may be recognised to be chiefly due to the skewness β_1 so far as these examples show.

[2] The effect of non-normality on t-distribution.

Some authors have already remarked in view of sampling experiments that the skewness of t-distribution from non-normal distribution manifests itsself in the reverse direction to that of the parent population. This assertation may be also verified in our Theorem 2, so far as our hypothesis remain true.

The effect of skewness becomes far more significant in this case than the distribution of t, as we may observe from Table 2.

These effects may be seen from the following Table 4.

TABLE 4. $pr\{t>t_{n-1}(\alpha)\}\equiv\alpha'$ (n=5, 10,15, 21)

Parent population	$\alpha=0.5\%$	$\alpha=2.5\%$
$\sqrt{\beta_1}=\sqrt{0.5}$ $\beta_2=3.5$ $\beta_3\doteqdot5$ $\beta_4\doteqdot26.4$	about 0.2%	1.5%

Parent population	$\alpha=0.5\%$	$\alpha=2.5\%$
$\sqrt{\beta_1}=-0.5$ $(\beta_1=0.5)$ $\beta_2=3.5$ $\beta_3\doteqdot 5$ $\beta_4\doteqdot 26.4$	about 1.1%	3.9%~4.5%
$\beta_1=\sqrt{0.5}$ $\beta_2=2.5$ $\beta_3\doteqdot 2.78$ $\beta_4\doteqdot 9.88$	about 0.1%	0.8%~1.5%
$\beta_1=\sqrt{0.5}$ $\beta_2=2.5$ $\beta_3\doteqdot 2.78$ $\beta_4\doteqdot 9.88$	1.2% ~1.5%	4.1%~5.5%