ON THE REMAINDER TERM FOR THE CENTRAL LIMIT THEOREM

RYOICHI SHIMIZU

(Received Jan. 11, 1974; revised May 9, 1974)

Summary

An upper bound for the remainder term of the Edgeworth expansion for the distribution of the normalized sum of independent and identically distributed random variables is given in terms of 3rd and 4th order moments, together with the total variation of the probability density function of the underlying distribution.

1. Inequalities

Let F be an absolutely continuous distribution with the finite moments up to 4th order: $\alpha_1=0$, $\alpha_2=1$, α_3, and α_4, and with the density $f(x)$ which is of bounded variation: total variation of $f(x)=M<\infty$. Let $F_n(x)$ be the distribution function for the normalized sum of a sample of size n from F. Consider the remainder term,

$$(1)\qquad R_n(x)=F_n(x)-\Phi(x)-\frac{\alpha_3}{6\sqrt{n}}(1-x^2)\phi(x)$$

of the Edgeworth expansion for F_n. We shall show that the following inequalities hold.

$$(2)\qquad \begin{aligned}\sup|n\pi R_n(x)|\leqq\ & 2[2c(1-2c)^{-2}+e^{-n/2\alpha_4}]\alpha_4\\ &+[2\alpha_3^2/9+\sqrt{2\pi}\,|\alpha_3|\,|\alpha_4|/4\sqrt{n}]\\ &+\min\,[2(\sqrt{\alpha_4}\,M)^n,\ 6\alpha_4M^2L_0^n]\ ,\end{aligned}$$

where $c=c(\alpha_4)=\alpha_4\log(2\alpha_4/(2\alpha_4-1))-11/24\leqq 1/4$, $L_0=(M\sqrt{3\alpha_4/(k+1)})^{1/k}$, and $k=$maximum integer not greater than $3\alpha_4M^2$. When F is symmetric,

$$(3)\quad \sup|n\pi R_n(x)|\leqq[2c(1-2c)^{-2}+e^{-n/2\alpha_4}]\alpha_4+\min\,[(\sqrt{\alpha_4}\,M)^n,\ 3\alpha_4M^2L_0^n]\ .$$

A similar inequality for the density is given in Section 6.

2. An upper bound for $|R_n(x)|$

Let

$$g_n(t)=\int_{-\infty}^{\infty} e^{itx}dG_n(x)=e^{-t^2/2}\left(1-i\frac{\alpha_3}{6\sqrt{n}}t^3\right)$$

be the Fourier-Stiltjes transform of

$$G_n(x)=\Phi(x)+\frac{\alpha_3}{6\sqrt{n}}(1-x^2)\phi(x)\ ,$$

and let $\psi(t)$ be the characteristic function of F. Since,

$$(4)\qquad |\psi(t)|=\left|\int_{-\infty}^{\infty} e^{itx}f(x)dx\right|=\left|-\frac{1}{it}\int_{-\infty}^{\infty} e^{itx}df(x)\right|\leqq\frac{M}{|t|}\ ,\quad \text{for all } t\neq 0\ ,$$

the inversion formula gives, for any real x and y,

$$(5)\qquad R_n(x)-R_n(y)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{e^{-ity}-e^{-itx}}{it}\left[\psi^n\left(\frac{t}{\sqrt{n}}\right)-g_n(t)\right]dt\ .$$

It follows that

$$(6)\qquad \begin{aligned}|R_n(x)|&\leqq\sup_y|R_n(x)-R_n(y)|\\ &\leqq\frac{2}{\pi}\int_0^{\infty}t^{-1}\left|\psi^n\left(\frac{t}{\sqrt{n}}\right)-g_n(t)\right|dt\\ &\leqq\frac{2}{\pi}\int_0^{\infty}t^{-1}|\psi^n(t)-g_n(\sqrt{n}\,t)|dt\\ &=\frac{2}{\pi}(I_{1,n}+I_{2,n}+I_{3,n})\ ,\end{aligned}$$

where

$$I_{1,n}=\int_0^{\alpha_4^{-1/2}}t^{-1}|\psi^n(t)-g_n(\sqrt{n}\,t)|dt\ ,$$

$$I_{2,n}=\int_{\alpha_4^{-1/2}}^{\infty}t^{-1}|\psi^n(t)|dt\ ,\qquad\text{and}$$

$$I_{3,n}=\int_{\alpha_4^{-1/2}}^{\infty}t^{-1}|g_n(\sqrt{n}\,t)|dt\ .$$

If F is symmetric,

$$(7)\qquad |R_n(x)|=|R_n(x)-R_n(0)|=\frac{1}{2\pi}\left|\int_{-\infty}^{\infty}\frac{1-e^{-itx}}{it}\left[\psi^n\left(\frac{t}{\sqrt{n}}\right)-g_n(t)\right]dt\right|$$

$$=\frac{1}{2\pi}\left|\int_{-\infty}^{\infty}\frac{\sin tx}{t}\left[\phi^n\left(\frac{t}{\sqrt{n}}\right)-g_n(t)\right]dt\right|$$
$$\leqq\frac{1}{\pi}(I_{1,n}+I_{2,n}+I_{3,n}) .$$

3. Estimation of $I_{1,n}$

LEMMA 1. *For* $0\leqq t\leqq\alpha_4^{-1/2}$, $\phi(t)$ *can be put in the form,*

$$\phi(t)=\exp\left(-\frac{t^2}{2}-i\frac{\alpha_3}{6}t^3+A\alpha_4t^4\right) , \tag{8}$$

where $|A|\leqq c(\alpha_4)=\alpha_4\log(2\alpha_4/(2\alpha_4-1))-11/24$.

PROOF. We can write $\phi(t)$ in two ways:

$$\phi(t)=1-\frac{1}{2}\theta_1t^2 \tag{9}$$

$$=1-\frac{t^2}{2}-i\frac{\alpha_3}{6}t^3+\frac{\alpha_4}{24}\theta_2t^4 , \tag{10}$$

where $|\theta_1|\leqq1$, and $|\theta_2|\leqq1$. Put $\alpha(t)=\phi(t)-1$. In this case, because $|\alpha(t)|\leqq t^2/2\leqq\alpha_4^{-1}/2\leqq1/2$, for $0\leqq t\leqq\alpha_4^{-1/2}$, we have,

$$\log\phi(t)=\log(1+\alpha(t))=\alpha(t)+\alpha^2(t)\beta(t) ,$$

where, of course,

$$\beta(t)=-\frac{1}{2}+\frac{1}{3}\alpha(t)-\frac{1}{4}\alpha^2(t)+\cdots . \tag{11}$$

Therefore, (8) holds if we take

$$A=\frac{\theta_2}{24}+\alpha^2(t)\beta(t)\alpha_4^{-1}t^{-4}$$

with

$$|A|\leqq\frac{1}{24}+|\alpha(t)t^{-2}|^2\alpha_4^{-1}|\beta(t)|$$
$$\leqq\frac{1}{24}+\frac{1}{4}\alpha_4^{-1}\left(\frac{1}{2}+\frac{1}{3}|\alpha(t)|+\frac{1}{4}|\alpha(t)|^2+\cdots\right)$$
$$\leqq\frac{1}{24}+\alpha_4\left(\frac{1}{2}\left(\frac{1}{2\alpha_4}\right)^2+\frac{1}{3}\left(\frac{1}{2\alpha_4}\right)^3+\cdots\right)$$
$$=\alpha_4\log\frac{2\alpha_4}{2\alpha_4-1}-\frac{11}{24} .$$

q.e.d.

Note that $c(\alpha_4)$ is monotone decreasing and tends to 1/24 as α_4 tends to infinity, while $c(1)\leqq 1/4$ and $c(9/5)\leqq 1/7$.

It follows, from the lemma and the inequalities $|e^{x+iy}-1|\leqq|x+iy|\cdot e^{|x+iy|}$, $|e^{iy}-1-iy|\leqq y^2/2$ which are valid for any real x and y, that,

$$\begin{aligned}
(12)\qquad I_{1,n}&=\int_0^{\alpha_4^{-1/2}} t^{-1}e^{-nt^2/2}\left|e^{-i(\alpha_3/6)nt^3+A\alpha_4nt^4}-1+i\frac{\alpha_3}{6}nt^3\right|dt\\
&\leqq\int_0^{\alpha_4^{-1/2}} t^{-1}e^{-nt^2/2}|e^{A\alpha_4nt^4}-1|dt\\
&\quad+\int_0^{\alpha_4^{-1/2}} t^{-1}e^{-nt^2/2}\left|e^{-i(\alpha_3/6)nt^3}-1+i\frac{\alpha_3}{6}nt^3\right|dt\\
&\leqq cn\alpha_4\int_0^\infty t^3e^{-n(1-2c)t^2/2}dt+\frac{1}{2}\left(\frac{\alpha_3}{6}\right)^2n^2\int_0^\infty t^5e^{-nt^2/2}dt\\
&=2c(1-2c)^{-2}\alpha_4/n+\alpha_3^2/9n\ .
\end{aligned}$$

4. Estimation of $I_{2,n}$

Inequality (4) gives,

$$(13)\qquad I_{2,n}\leqq\int_{\alpha_4^{-1/2}}^\infty (M/t)^n t^{-1}dt=(\sqrt{\alpha_4}\,M)^n/n\ .$$

This estimate of $I_{2,n}$ is useful only when $\sqrt{\alpha_4}\,M\leqq 1$. In order to derive another estimate which covers the case $\sqrt{\alpha_4}\,M>1$, we need the following

LEMMA 2. *The total variation of* $f^{*n}\leqq\sqrt{3/(n+1)}\,M$.

PROOF. Let $U_M(x)$ be the density of the uniform distribution on the interval $(-1/M, 1/M)$. According to Rogozin [3], the left-hand side of the above inequality is bounded by the total variation $2U_M^{*n}(0)$ of the density $U_M^{*n}(x)$. For $n\leqq 3$, therefore, the lemma is clear from $U_M^{*1}(0)=U_M^{*2}(0)=M/2$ and $U_M^{*3}(0)=3M/8$. If $n\geqq 4$, using the inversion formula, we obtain,

$$\begin{aligned}
0\leqq U_M^{*n}(0)&=\frac{1}{2\pi}\int_{-\infty}^\infty(\sin tM/tM)^n dt\\
&\leqq\frac{M}{\pi}\int_0^\infty\left|\frac{\sin t}{t}\right|^n dt\\
&\leqq\frac{M}{\pi}\left\{\int_0^{\pi/2}e^{-nt^2/6}dt+\int_{\pi/2}^\infty t^{-n}dt\right\}\\
&\leqq\sqrt{\frac{d_n}{n+1}}\,\frac{M}{2}
\end{aligned}$$

where

$$d_n=\left\{\frac{2}{\sqrt{2\pi}}\sqrt{\frac{3(n+1)}{n}}+\frac{\sqrt{n+1}}{n-1}\left(\frac{2}{\pi}\right)^n\right\}^2 .$$

We have only to show that $d_n \leqq 3$. But d_n is decreasing and $d_4 \leqq 2.781$. q.e.d.

If k is the greatest integer less than or equal to $3\alpha_4 M$, then

$$k \leqq 3\alpha_4 M^2 , \tag{15}$$

and

$$L \equiv \sqrt{\frac{3\alpha_4}{k+1}}M<1 . \tag{16}$$

Since for $n=1, 2, \cdots$, $|\phi^n(t)|=|\phi^k(t)|^{n/k} \leqq \{L\alpha_4^{-1/2}/|t|\}^{n/k}$ (using Lemma 2 and an inequality similar to (4)), we obtain,

$$I_{2,n} \leqq \int_{\alpha_4^{-1/2}}^{\infty} (L\alpha_4)^{n/k} t^{-1-n/k} dt=\frac{k}{n} L^{n/k} \leqq \frac{3\alpha_4 M^2}{n} L_0^n , \tag{17}$$

where $L_0=L^{1/k}<1$. Since both inequalities (13) and (17) hold for any case, we can use as the estimate of $I_{2,n}$, the smaller one.

$$I_{2,n} \leqq \min\left((\sqrt{\alpha_4}\, M)^n/n,\ 3\alpha_4 M^2 L_0^n/n\right) . \tag{18}$$

5. Estimation of $I_{3,n}$ and the proof of (2) and (3)

$I_{3,n}$ is directly evaluated as,

$$I_{3,n} \leqq \int_{\alpha_4^{-1/2}}^{\infty} t^{-1} e^{-nt^2/2}\left(1+\frac{|\alpha_3|}{6} nt^3\right)dt \leqq \alpha_4 e^{-n/2\alpha_4}/n+\sqrt{2\pi}\,|\alpha_3|\alpha_4/4n^{3/2} . \tag{19}$$

The desired inequalities are obtained from (12), (18) and (19).

6. Remainder term for density and some special cases

The same method as above is applied to evaluate the remainder term,

$$S_n(x)=f_n(x)-\phi(x)-\frac{\alpha_3}{6\sqrt{n}}(x^3-3x)\phi(x)$$

for the density $f_n(x)=\sqrt{n}\, f^{*n}(\sqrt{n}\, x)$.

Using the inversion formula we obtain,

$$|S_n(x)|=\left|\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}\left(\phi^n\left(\frac{t}{\sqrt{n}}\right)-g_n(t)\right)dt\right|$$

$$\leqq \frac{\sqrt{n}}{\pi} \int_{-\infty}^{\infty} |\phi^n(t) - g_n(\sqrt{n}\, t)| dt , \qquad \text{for } n \geqq 2 ,$$

and

$$\text{(20)} \qquad \sup |(n-2)\pi S_n(x)|$$
$$\leqq \sqrt{2\pi}\left[\frac{3}{2}c(1-2c)^{-5/2}\alpha_4 + \frac{15}{154}\alpha_3^2\right] + \left(\alpha_4^{3/2} + \frac{1}{6}|\alpha_3|\right)\Big/\sqrt{n}$$
$$+\min\,[(\sqrt{\alpha_4}\,M)^{n-1}M,\ 3\alpha_4 M^3 L_0^{n-1}] , \qquad n \geqq 3 .$$

An important special case is when F is unimodal with the mode at $x=0$. The corresponding characteristic function $\phi(t)$ can be written in the form,

$$\phi(t) = t^{-1}\int_0^t \xi(\tau)d\tau$$

where $\xi(t)$ is a characteristic function. Multiplying both sides by t and differentiating at $x=0$, we have (if $\phi^{(2k+1)}(0)$ exists),

$$(j+1)\phi^{(j)}(0) = \xi^{(j)}(0) , \qquad j=0,\cdots,2k .$$

The well known sequence of inequalities for absolute moments, viz

$$|\xi^{(2)}(0)|^{1/2} \leqq |\xi^{(4)}(0)|^{1/4} \leqq \cdots$$

implies,

$$\text{(21)} \qquad [(2j+1)\alpha_{2j}]^{1/2j} \geqq [(2j-1)\alpha_{2j-2}]^{1/(2j-2)} , \qquad \text{or}$$
$$\alpha_{2j} \geqq \frac{1}{2j+1}[(2j-1)\alpha_{2j-2}]^{j/(j-1)} , \qquad j=2,\cdots,k .$$

Therefore, if 5th moment does exist, then $\alpha_4 \geqq 9/5$, $c(\alpha_4) \leqq c(9/5) \leqq 1/7$, and $c(1-2c)^{-2} \leqq 7/25$. If F is a symmetric unimodal distribution with the finite 5th moment, and if $\sqrt{\alpha_4}\,M \leqq 1$, (3) reduces to

$$\sup |n\pi R_n(x)| \leqq (39/25)\alpha_4 + (\sqrt{\alpha_4}\,M)^n .$$

Example. For the symmetric uniform distribution with the variance 1 ($f(x)=\sqrt{3}/6$, if $|x| \leqq \sqrt{3}$, $=0$, otherwise), $\alpha_4 = 12/5$, $M = 2f(0) = 1/\sqrt{3}$, and (3) becomes,

$$\sup R_n(x) \leqq R_n \equiv [2.4(2c(1-2c)^{-2} + e^{-5n/24}) + 0.9^n]/n\pi .$$

A computer calculation shows that $R_2=0.5027$, $R_3=0.2946$, $R_4=0.1957$, $R_5=0.1398$, $R_{10}=0.0447$, $R_{15}=0.0227$, $R_{20}=0.01466$, $R_{30}=0.00866$, and $R_{50}=0.004969$.

THE INSTITUTE OF STATISTICAL MATHEMATICS

REFERENCES

[1] Feller, W. (1966). *An Introduction to Probability Theory and its Applications II,* John Wiley, New York.

[2] Gnedenko, B. V. and Kolmogorov, A. N. (1967). *Limit Distributions for Sum of Independent Random Variables* (English translation), Addison-Wesley, Cambridge, Mass.

[3] Rogozin, B. A. (1965). On the maximum of the density of the sum of random variables with unimodal distribution, *Litovsk. Mat. Sb.*, **5**, 499-503, English translation: *Selected Transl. Math. Statist. and Prob.*, **9**, 69-74, 1970.