Ann. Inst. Statist. Math.
29 (1977), Part A, 397-406

EXTENSION OF EDGEWORTH TYPE EXPANSION OF THE DISTRIBUTION OF THE SUMS OF I.I.D. RANDOM VARIABLES IN NON-REGULAR CASES

KEI TAKEUCHI AND MASAFUMI AKAHIRA*

(Received Oct. 21, 1976; revised July 7, 1977)

Abstract

The asymptotic expansions of the distributions of the sums of independent identically distributed random variables are given by Edgeworth type expansions when moments do not necessarily exist, but when the density can be approximated by rational functions.

1. Introduction

Asymptotic distributions for the sums of independent and identically distributed (i.i.d.) random variables has been extensively studied for many years ([1]). Suppose that $X_1, X_2, \cdots, X_n$ is a sequence of i.i.d. random variables with mean μ and the distribution function $F(x)$. Then the asymptotic expansion for the distribution of $Y_n=\left(\sum_1^n X_i-n\mu\right)\Big/\sqrt{n}$ up to the order $n^{-m/2+1}$ is given by the classical Edgeworth expansion when F has the moments up to the mth order as well as the continuous density ([2], [3]). The purpose of the present paper is to extend the result to the case when the moments do not necessarily exist assuming only that the density function $f(x)$ of F is approximated by some appropriate rational functions as $|x|\to\infty$. Put, for some real constants M_n and V_n (>0), $Z_n=\left(\sum_1^n X_i-M_n\right)\Big/V_n$. Then, as is well known, if a $G(x)$ is a limiting distribution of a sequence of distributions of Z's, it is necessarily a stable distribution. We shall show in the next section that if this is the case then under the assumption stated above V_n must be either n or $\sqrt{n\log n}$ or $\sqrt{n}$ and the corresponding leading term of the asymptotic expansion is the Cauchy, or the unsymmetric stable distribution with characteristic exponent 1, or the normal distribution, respectively. The second term is shown to be of order either

* Supported in part by the Sakkokai Foundation.

$1/n$, $(\log n)/n$, $1/\log n$ or $1/\sqrt{n}$ and the density function corresponding to the second term are also given. Generally the relative magnitudes of the successive terms are complicated and is difficult to write down explicitely but may be obtained for each particular case. The details of the computation are given in [4] and may be referred to for further examples. Further the similar results for fractional characteristic exponents are given in [5] (see also Cramér [6] for this case).

2. Results

We shall obtain the Edgeworth type expansions for the sums of random variables not necessarily with finite moments but with the density $f(x)$ which satisfies the following condition:

There exist two rational functions $g^+(x)$ and $g^-(x)$ such that for a positive integer m ($\geqq 2$) and some $0<\delta<1$

(a) $\lim_{x\to\infty} x^{m+1+\delta}|f(x)-g^+(x)|=0$;

(b) $\lim_{x\to-\infty} |x|^{m+1+\delta}|f(x)-g^-(x)|=0$.

Remark 1. m depends on the order of the expansion and should not be smaller than 2 in order that we have meaningful expansions.

Remark 2. It should be remarked that since the distribution has a density $f(x)$, we have

$$\sup_{|t|>\rho} |\phi(t)|<1$$

for any $\rho>0$.

We shall first show that the following proposition holds.

PROPOSITION. $g^+(x)$ and $g^-(x)$ may be taken of the form

$$g^+(x)=\sum_{j=2}^{m+1} \frac{\alpha_j}{(x+1)^j}\,; \qquad g^-(x)=\sum_{j=2}^{m+1} \frac{\beta_j}{(x-1)^j}\,,$$

where α_j and β_j are real constants.

PROOF. We first decompose $g^+(x)$ into the sums of partial fractions such as

$$g^+(x)=\sum_i \sum_j \frac{B_{ij}}{(x+D_i)^j}+\sum_i \sum_j \frac{E_{ij}}{(x^2+F_i x+G_i)^j}\,.$$

But each term in above expression can be approximated by a sum of the terms of the type $\beta_j/(x+1)^j$ up to order x^{-m-1} when x is large. Therefore $g^+(x)$ may be replaced by a sum of such terms. Similar is true of $g^-(x)$.

Now let us denote that

$$\phi_j(t)=\int_0^\infty \frac{(j-1)e^{itx}}{(1+x)^j}dx ,$$

then we have

$$\int_{-\infty}^0 \frac{(j-1)e^{itx}}{(1-x)^j}dx=\phi_j(-t)=\overline{\phi_j(t)} .$$

Thus if the density $f(x)$ satisfies the condition above, the characteristic function can be expressed as

$$(1) \qquad \phi(t)=\int e^{itx}f(x)dx=\sum_j \alpha_j\phi_j(t)+\sum_j \beta_j\phi_j(-t)+\psi(t)$$

and since

$$\int_0^\infty x^m|f(x)-g^+(x)|dx<\infty ; \qquad \int_{-\infty}^0 |x|^m|f(x)-g^-(x)|dx<\infty ,$$

the remainder term $\psi(t)$ can differentiated m times at $t=0$, so that we have

$$(2) \qquad \psi(t)=\sum_{p=0}^m c_p t^p+o(|t|^m) .$$

It is shown by Takeuchi and Akahira [4] that for small $|t|>0$ and $j\geqq 2$

$$(3) \qquad \begin{aligned}\phi_j(t)&=(j-1)\left(\int_{|t|}^\infty u^{-j}e^{-iu}du\right)|t|^{j-1}e^{-it}\\ &=(j-1)\left\{\sum_{k=0}^{j-2} i^k a_k|t|^k-\frac{i^{j-1}}{(j-1)!}|t|^{j-1}\log|t|+\alpha_j^*|t|^{j-1}\right.\\ &\qquad\left.+i\beta_j^*|t|^{j-1}-\frac{i^j}{j}|t|^j\right\}\left(1-it-\frac{1}{2}t^2+\frac{i}{6}t^3+\cdots\right)\\ &=1+\sum_{k=1}^m a_{jk}|t|^k+\sum_{k=1}^m b_{jk}t|t|^{k-1}+\sum_{k=j}^{m+1} c_{jk}|t|^{k-1}\log|t|\\ &\qquad+\sum_{k=j}^m d_{jk}t|t|^{k-1}\log|t|+o(|t|^m) ,\end{aligned}$$

where $a_k=-1/k!(k-j+1)$ and α_j^* and β_j^* are real constants equal to the real part and the imaginary part of

$$\int_1^\infty u^{-j}e^{iu}du+\int_0^1 u^{-j}\left\{e^{iu}-1-iu+\frac{u^2}{2}+\cdots-\frac{(iu)^{j-1}}{(j-1)!}\right\}du+\sum_{k=0}^{j-2}\frac{i^k}{k!(k-j+1)}$$

respectively, and a_{jk}, b_{jk}, c_{jk} and d_{jk} are certain complex numbers. Hence we have from (1), (2) and (3)

$$(4)\qquad \phi(t)=1+\sum_{k=1}^{m} A_k|t|^k+\sum_{k=1}^{m} B_k t|t|^{k-1}+\sum_j\sum_{k=j}^{m+1} c_{jk}|t|^{k-1}\log|t| +\sum_j\sum_{k=j}^{m} d_{jk}t|t|^{k-1}\log|t|+o(|t|^m)\,,$$

where A_k and B_k are certain complex numbers.

Let $X_1, X_2,\cdots, X_n,\cdots$ be a sequence of i.i.d. random variables with the characteristic function $\phi(t)$. We define $z_n=\left(\sum_{i=1}^{n} X_i-M_n\right)\Big/V_n$ whose characteristic function $\phi_n(t)$ is expressed as

$$\phi_n(t)=\mathrm{E}\,(e^{itZ_n})=\left\{\phi\left(\frac{t}{V_n}\right)\right\}^n \exp\left(-i\frac{M_n}{V_n}t\right)\,.$$

Since

$$\log\phi_n(t)=n\log\phi\left(\frac{t}{V_n}\right)-i\frac{M_n}{V_n}t\,,$$

it follows from (4) that

$$(5)\qquad \log\phi_n(t)=-i\frac{M_n}{V_n}t+n\left[\left\{\sum_{k=1}^{m} A_k\frac{|t|^k}{V_n^k}+\sum_{k=1}^{m} B_k\frac{t|t|^{k-1}}{V_n^k}+\sum_{k=1}^{m} C_k\frac{|t|^k}{V_n^k}\cdot(\log|t|-\log V_n)+\sum_{k=1}^{m} D_k\frac{t|t|^{k-1}}{V_n^k}(\log|t|-\log V_n) +\cdots\right\}-\frac{1}{2}\{\quad\}^2+\frac{1}{3}\{\quad\}^3-\cdots\right]\,.$$

It is noted that since $\phi_n(t)$ is a characteristic function B_1 is a pure imaginary if not equal to zero, and that term can be cancelled out by an appropriate choice of M_n. Without loss of generality we put $B_1=0$.

In order to obtain V_n and the leading term of $\phi_n(t)$ it is enough to consider the following cases:

(i) $A_1\neq 0$,
(ii) $A_1=0,\ A_2\neq 0$,
(iii) $A_1=A_2=0,\ C_1=D_1=0,\ C_2>0,\ D_2=0$,
(iv) $A_1=A_2=0,\ C_1>0,\ D_1=0$.

Since the asymptotic distribution of Z_n is stable, it is seen from (5) that each case is as follows:

Case (i). $V_n=n$ necessarily holds and then $C_1=0$. If $D_1=0$, then the leading term of $\phi_n(t)$ is given by $e^{A_1|t|}$. If $D_1\neq 0$, then M_n must be $O(n\log n)$. When $D_1\neq 0$, we have the characteristic function $\phi_n(t)$ with the leading term of the type $\exp\{A_1|t|+D_1 t\log|t|\}$ which is the unsymmetric stable law of characteristic exponent 1.

Case (ii). $V_n=n$ necessarily holds and then $B_2=C_1=C_2=D_1=D_2=0$.

Hence the leading term of $\phi_n(t)$ is given by $e^{A_2t^2}$.

Case (iii). $V_n=\sqrt{n\log n}$ necessarily holds. Hence the leading term of $\phi_n(t)$ is given by $e^{-C_2t^2}$.

Case (iv). $V_n=n\log n$ and the leading term of $\phi_n(t)$ must be $e^{-C_1|t|}$. If $C_1>0$, then F belongs to the domain of attraction of the Cauchy distribution, and the density $f(x)$ is expressed as

$$f(x)=\begin{cases}\alpha(1+x)^{-2}+o(x^{-2}) & \text{as } x\to\infty\,,\\ \beta(1-x)^{-2}+o(x^{-2}) & \text{as } x\to-\infty\,,\end{cases}$$

with $\alpha+\beta>0$. But this implies that F belongs to the domain of normal attraction of the stable law with characteristic exponent 1 (see [1], p. 181), which contradicts the assumption $V_n=n\log n$. Thus the case (iv) can be excluded from our consideration, except for the trivial case $C_1=0$.

In other cases the leading term of $\phi_n(t)$ is equal to 1 which is the degenerate case.

When the leading term is obtained in each case, the remaining terms may be arranged in the order of magnitude. A few examples are in order to show the point.

Example 1. Let X_i's $(i=1,2,\cdots)$ be i.i.d. random variables with the density function $f(x)$ given by

$$f(x)=\frac{1}{2(x^2+1+\sqrt{x^2+1})}\,.$$

Since $f(x)$ is a symmetric function and $f(x)\sim 1/2x^2$ as $x\to\infty$, $F(y)$ belongs to the domain of normal attraction of a stable law with characteristic exponent 1 ([1]).

Note that X_i can be expressed as $X_i=(1/(1-U_i)-1/U_i)/2$, where U_i is distributed uniformly in (0, 1). Putting

$$g(x)=\frac{1}{2(1+x^2)(1+x^2+\sqrt{1+x^2})}\,;$$

$$h(x)=\frac{1}{2(1+x^2)^{3/2}}-\frac{1}{2(1+|x|)^3}$$

we have

$$f(x)=\frac{1}{2(1+x^2)}-\frac{1}{2(1+x^2)^{3/2}}+g(x) \tag{6}$$
$$=\frac{1}{2(1+x^2)}-\frac{1}{2(1+|x|)^3}-h(x)+g(x)\,.$$

If we have $m=2$, $0<\delta<1$ and $g^{\pm}(x)=1/2(1+x^2)-1/2(1+|x|)^3$, then (a) and (b) are satisfied. First we obtain

$$\int_0^x y^2h(y)dy=\frac{1}{2}\log(\sqrt{x^2+1}+x)-\frac{x}{2\sqrt{1+x^2}}-\frac{1}{2}\log(1+x)$$
$$-\left(\frac{1}{1+x}-1\right)+\frac{1}{(1+x)^2}-1$$
$$=\frac{1}{2}\log\{(\sqrt{x^2+1}+x)/(x+1)\}-\frac{x}{2\sqrt{1+x^2}}-\frac{x}{(1+x)^2}.$$

Hence it follows that

$$\int_{-\infty}^{\infty} x^2h(x)dx=\log 2 .$$

Since

$$\int_{-\infty}^{\infty} h(x)dx=1/2 ; \qquad \int_{-\infty}^{\infty} xh(x)dx=0 ,$$

it is seen that

$$(7)\qquad \int_{-\infty}^{\infty} e^{itx}h(x)dx=\frac{1}{2}+\frac{\log 2}{2}t^2+o(t^2) .$$

Since

$$\int_{-\infty}^{\infty} g(x)dx=2-\frac{\pi}{2} ; \qquad \int_{-\infty}^{\infty} x^2g(x)dx=\frac{\pi}{2}-1 ,$$

we have

$$(8)\qquad \int_{-\infty}^{\infty} e^{itx}g(x)dx=2-\frac{\pi}{2}-\frac{1}{2}\left(\frac{\pi}{2}-1\right)t^2+o(t^2) .$$

It follows by Takeuchi and Akahira [4] that

$$(9)\qquad \int_{-\infty}^{\infty}\frac{e^{itx}}{(1+|x|)^3}dx=1+t^2\log|t|-\gamma t^2+o(t^2) ,$$

where γ is some constant. From (6)–(9) we obtain

$$\phi(t)=\int_{-\infty}^{\infty} e^{itx}f(x)dx$$
$$=\frac{\pi}{2}e^{-|t|}-\frac{1}{2}-\frac{t^2}{2}\log|t|+\frac{\gamma}{2}t^2-\frac{1}{2}-\frac{\log 2}{2}t^2+2-\frac{\pi}{2}$$
$$-\frac{1}{2}\left(\frac{\pi}{2}-1\right)t^2+o(t^2)$$
$$=1-\frac{\pi}{2}|t|-\frac{t^2}{2}\log|t|+\frac{1}{2}(1-\gamma-\log 2)t^2+o(t^2) .$$

Letting $\phi_n(t)=\mathrm{E}[\exp itZ_n]$ with $Z_n=\sum_1^n X_i/n$, we obtain

$$\log\phi_n(t)=n\log\phi\left(\frac{t}{n}\right)$$
$$=-\frac{\pi}{2}|t|-\frac{t^2}{2n}(\log|t|+\log n)+\frac{1}{2}(1-\gamma-\log 2)\frac{t^2}{n}$$
$$-\frac{\pi^2}{8n}t^2+o\left(\frac{1}{n}\right).$$

Since

$$\phi_n(t)=e^{-(\pi/2)|t|}\left[1-\frac{\log n}{2n}t^2-\frac{1}{2n}t^2\log|t|\right.$$
$$\left.+\frac{1}{8n}\{4(1-\gamma-\log 2)-\pi^2\}t^2\right]+o\left(\frac{1}{n}\right),$$

the asymptotic expansion of the distribution of Z_n agrees essentially with the Cauchy distribution. The second term in the expansion has the form

$$\frac{1}{n}(a\log n+b)f_1(x)+cf_2(x),$$

where $f_1(x)$ and $f_2(x)$ are the functions that are Fourier transforms of

$$t^2e^{-(\pi/2)|t|} \quad\text{and}\quad (t^2\log|t|)e^{-(\pi/2)|t|}$$

respectively. Hence $f_1(x)$ and $f_2(x)$ are equal to $-g''(x)$ and $-h''(x)$, respectively, where

$$g(x)=\frac{1}{\pi(c'^2+x^2)};$$

$$h(x)=\frac{1}{2\pi(1+x^2)}\{-\log(1+x^2)-2x\tan^{-1}x+C''\}$$

with certain constants c' and C''.

Example 2. Let X_i's $(i=1,2,\cdots)$ be i.i.d. random variables with the normal density with mean 0 and variance 1. Let $Y_i=1/X_i$ $(i=1,2,\cdots)$. Then the density function $f(y)$ of Y_i is given by

$$f(y)=\frac{1}{\sqrt{2\pi}}\frac{1}{y^2}e^{-1/2y^2}.$$

Since $f(y)$ is a symmetric function and $f(y)\sim 1/(\sqrt{2\pi}y^2)$ as $y\to\infty$, $F(x)$ belongs to the domain of normal attraction of a stable law with the characteristic exponent 1 ([1]). Put

$$g(y)=f(y)-\frac{\sqrt{2}}{\sqrt{\pi}\,(1+2y^2)}\,. \tag{10}$$

If we have $m=2$, $0<\delta<1$ and $g^{\pm}(x)=2/(\sqrt{\pi}\,(1+2x^2))$, then (a) and (b) are satisfied. First we have

$$\int_{-\infty}^{\infty} g(y)dy=1-\sqrt{\pi}\ ; \tag{11}$$

$$\int_{-\infty}^{\infty} yg(y)dy=0\ . \tag{12}$$

Further we obtain

$$\begin{aligned}\int_{-\infty}^{\infty} y^2g(y)dy&=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\left(e^{-1/2y^2}-\frac{2y^2}{1+2y^2}\right)dy\\&=\sqrt{\frac{2}{\pi}}\int_0^{\infty}\left\{\frac{1}{x^2}e^{-x^2/2}-\frac{2}{x^2(2+x^2)}\right\}dx\\&=\sqrt{\frac{2}{\pi}}\int_0^{\infty}(2u)^{-3/2}\left(e^{-u}-\frac{1}{1+u}\right)du\\&=\frac{1}{2\sqrt{\pi}}\left\{-2\Gamma\left(\frac{1}{2}\right)+2\int_0^{\infty}\frac{u^{-1/2}}{(1+u)^2}du\right\}\\&=-1+\frac{1}{\sqrt{\pi}}\int_0^{\infty}v^{1/2}(1-v)^{-1/2}dv\\&=-1+\frac{1}{\sqrt{\pi}}B\left(\frac{3}{2},\frac{1}{2}\right)\\&=-1+\frac{\sqrt{\pi}}{2}\,.\end{aligned} \tag{13}$$

From (10)-(13) we have

$$\begin{aligned}\phi(t)&=\int_{-\infty}^{\infty}e^{ity}f(y)dy\\&=\int_{-\infty}^{\infty}\frac{\sqrt{2}\,e^{ity}}{\sqrt{\pi}\,(1+2y^2)}dy+\int_{-\infty}^{\infty}e^{ity}g(y)dy\\&=\sqrt{\pi}\int_{-\infty}^{\infty}\frac{\exp\,(i(t/\sqrt{2}\,)x)}{\pi(1+x^2)}dx+\int_{-\infty}^{\infty}e^{ity}g(y)dy\\&=\sqrt{\pi}\,\exp\left(-\frac{|t|}{\sqrt{2}}\right)+1-\sqrt{\pi}-\frac{1}{2}\left(-1+\frac{\sqrt{\pi}}{2}\right)t^2+o(t^2)\\&=1-\sqrt{\frac{\pi}{2}}|t|+\frac{t^2}{2}+o(t^2)\,.\end{aligned}$$

Letting $\phi_n(t)=\mathrm{E}\,[\exp\,(itZ_n)]$ with $Z_n=\sum_1^n X_i/n$, we obtain

$$\log \phi_n(t) = n \log \phi\left(\frac{t}{n}\right) = -\sqrt{\frac{\pi}{2}}|t| + \left(\frac{\pi}{4} + \frac{1}{2}\right)\frac{t^2}{n} + o\left(\frac{1}{n}\right).$$

Since

$$\phi_n(t) = \left(\exp\left(-\sqrt{\frac{\pi}{2}}|t|\right)\right)\left\{1 + \left(\frac{\pi}{4} + \frac{1}{2}\right)\frac{t^2}{n}\right\} + o\left(\frac{1}{n}\right),$$

the asymptotic expansion of the distribution of Z_n agrees essentially with the case when the density function is $\sqrt{2}/\{\sqrt{\pi}(\pi + 2x^2)\}$ and the next term is $(a + bx^2 + cx^4)/\{\sqrt{2\pi}(1 + x^2)\}^3$.

Example 3. X_i's ($i = 1, 2, \cdots$) are distributed according to t-distribution with 2-degrees of freedom i.e. with the density

$$f(x) = \frac{1}{2(1 + x^2)^{3/2}}.$$

Note that the distribution has no finite variance, still it belongs to the domain of attraction of the normal. First we have

$$f(x) = h(x) + \frac{1}{2(1 + |x|)^3},$$

where $h(x)$ is given in Example 1.

The conditions (a) and (b) are satisfied with $m = 2$, $0 < \delta < 1$ and

$$g^{\pm}(x) = \frac{1}{2(1 + |x|)^3}.$$

From Example 1 we obtain

$$\phi(t) = \int_{-\infty}^{\infty} e^{itx} f(x)dx = 1 + \frac{1}{2}t^2 \log|t| - \frac{\log 2 + \gamma}{2}t^2 + o(t^2),$$

where γ is some constant. Letting $\phi_n(t) = \mathrm{E}[\exp(itZ_n)]$ with $Z_n = \sum_1^n X_i / \sqrt{n \log n}$, we have

$$\begin{aligned}\log \phi_n(t) &= n \log \phi\left(\frac{t}{\sqrt{n \log n}}\right) \\ &= \frac{1}{2}\frac{t^2 \log|t|}{\log n} - \frac{1}{4}\left\{1 + \frac{\log\log n}{\log n} + \frac{2(\log 2 + \gamma)}{\log n}\right\}t^2 + o\left(\frac{1}{\log n}\right) \\ &= e^{-(c_n/4)t^2}\left\{1 + \frac{1}{2}\frac{t^2 \log|t|}{\log n} + o\left(\frac{1}{n \log n}\right)\right\},\end{aligned}$$

where

$$c_n = 1 + \frac{\log\log n}{\log n} + \frac{2(\log 2 + \gamma)}{\log n}.$$

The leading term of the characteristic function $\phi_n(t)$ is equal to $e^{-t^2/4}$ which corresponds to the normal distribution. And the second term $g(x)$ is the Fourier transform of

$$\frac{1}{2}(t^2 \log|t|)e^{-t^2/4} ;$$

that is,

$$g(x) = -\pi G''(x) ,$$

where

$$G(x) = -\frac{2}{\sqrt{\pi}} e^{-x^2} \left\{ \int_0^x \left(\int_0^x e^{-x^2} dx \right) e^{x^2} dx + C \right\}$$

with

$$C = -\frac{1}{4\sqrt{\pi}} \int_{-\infty}^{\infty} (\log|t|) e^{-t^2/4} dt .$$

Acknowledgement

The authors wish to thank the referee for his kind comments.

UNIVERSITY OF TOKYO
UNIVERSITY OF ELECTRO-COMMUNICATIONS

REFERENCES

[1] Gnedenko, B. V. and Kolmogorov, A. N. (1954). *Limit Distributions for Sums of Independent Random Variables*, Addison-Wesley, Cambridge, Massachussets, (Translated from Russian).

[2] Ibragimov, I. A. and Linnik, Yu. V. (1971). *Independent and Stationary Sequences of Random Variables*, Wolters-Noordhoff, the Netherlands, (Translated from Russian).

[3] Petrov, V. V. (1975). *Sums of Independent Random Variables*, Springer-Verlag, (Translated from Russian).

[4] Takeuchi, K. and Akahira, M. (1976). On Gram-Charlier-Edgeworth type expansion of the sums of random variables (I), *Rep. Univ. Electro-Comm.* 27-1, 95–115.

[5] Takeuchi, K. and Akahira, M. (1976). On Gram-Charlier-Edgeworth type expansion of the sums of random variables (II), *Rep. Univ. Electro-Comm.* 27-1, 117–123.

[6] Cramér, H. (1963). On asymptotic expansions for sums of independent random variables with a limiting stable distribution, *Sankhyā*, Ser. A, 25, 13–24.