YOKOHAMA MATHEMATICAL
JOURNAL VOL. 35, 1987

# CONVERGENCE OF ENTROPY IN THE CENTRAL LIMIT THEOREM

By

SEIJI TAKANO

(Received June 2, 1987)

## §1. Introduction and results.

Let $\{S_n\}$ be a sequence of integer valued random variables and let $H_n$ denote the entropy of $S_n$ ([4]):

$$H_n=-\sum_i q_n(i)\log q_n(i),$$

where $q_n(i)$ is the distribution of $S_n$:

$$q_n(i)=P(S_n=i), \qquad n\geq 1,$$

and the base of the logarithm is 2. Suppose $\sigma_n^2=\operatorname{Var} S_n<\infty$ and define $\Delta_n$ by

$$\Delta_n=1/2\log(2\pi e\sigma_n^2)-H_n.$$

The condition $\Delta_n\to 0$ can be used to prove the convergence in law of $(S_n-ES_n)/\sigma_n$ to $N(0, 1)$ ([1], [2], [5]). In [2] Linnik proved that this type of condition holds for sums of independent random variables having bounded densities and used it for a proof of the central limit theorem. In this paper we restrict ourselves to sums of i.i.d. integer valued random variables and consider condition and rate of convergence of $\Delta_n$. Our proof is an application of the local limit theorem for lattice random variables.

Throughout the rest let $S_n=X_1+\cdots+X_n$, where $\{X_n\}$ is a sequence of i.i.d. integer valued random variables with finite variance $\sigma^2>0$. Let

$$\phi_n(i)=(\sigma n^{1/2})^{-1}\phi((i-n\mu)/(\sigma n^{1/2})), \qquad n\geq 1$$

where $\mu=EX_1$ and

$$\phi(x)=(2\pi)^{-1/2}\exp(-x^2/2).$$

Note that $\Delta_n$ can be expressed in terms of $\phi_n$:

$$\Delta_n=\sum_i q_n(i)\log[q_n(i)/\phi_n(i)].$$

An integer valued random variable $X$ is called to have a span $h\geq 1$ if there exists an integer $r$ such that $X-r$ is an integral multiple of $h$ with probability one. If maximal span of $X$ is one, then $X$ is called strongly aperiodic. The main results we are going to prove are as follows.

**Theorem 1.** *If $X_1$ is strongly aperiodic, then $\Delta_n \to 0$, $(n\to\infty)$. If $X_1$ has the maximal span $h>1$, then $\Delta_n \to \log h$, $(n\to\infty)$.*

**Theorem 2.** *If $X_1$ is strongly aperiodic and $E|X_1|^3<\infty$, then $\Delta_n=o(n^{-1/2+\varepsilon})$, $(n\to\infty)$ for every $\varepsilon>0$.*

## § 2. Proofs.

**Lemma 1.**

$$|\sum_i \phi_n(i)-1|=O(n^{-1}).$$

**Proof.** For fixed $n$, let $t_i=(i-n\mu)/(\sigma n^{1/2})$ and let $J_i$ denote the interval $[t_i-(2\sigma n^{1/2})^{-1},\ t_i+(2\sigma n^{1/2})^{-1})$. Then we have

$$(1)\qquad |\sum_i \phi_n(i)-1|=\left|\sum_i \int_{J_i}(\phi(t_i)-\phi(t))dt\right| \le \sum_i \left|\int_{J_i}(\phi(t_i)-\phi(t))dt\right|.$$

Using the expression

$$\phi(t)=\phi(t_i)+(t-t_i)\phi'(t_i)+(t-t_i)^2/2\,\phi''(t_i+\theta(t-t_i)),\qquad 0<\theta<1,$$

we obtain

$$(2)\qquad \left|\int_{J_i}(\phi(t_i)-\phi(t))dt\right|\le M_i/(24\sigma^3 n^{3/2})$$

where

$$M_i=\sup_{t\in J_i}|\phi''(t)|.$$

It is easy to see that $n^{-1/2}\sum_i M_i$ converges to $\sigma\int_{-\infty}^{\infty}|\phi''(t)|\,dt<\infty$ and therefore

$$M=\sup_n n^{-1/2}\sum_i M_i<\infty.$$

Thus by (1) we have

$$|\sum_i \phi_n(i)-1|\le M/(24\sigma^3 n).$$

This proves the lemma.

**Proof of Theorem 1.** Let $K=\log e$. Suppose that $X_1$ is aperiodic. By Lemma 1 we have

$$(3)\qquad K^{-1}\sum_i q_n(i)\log[q_n(i)/\phi_n(i)]\ge\sum_i q_n(i)[1-\phi_n(i)/q_n(i)] =1-\sum_i \phi_n(i)=O(n^{-1}).$$

On the other hand by the local limit theorem for i. i. d. integer ualued random variables ([3], p, 187), we have

$$(4) \qquad \delta_n=\sup_i n^{1/2}|q_n(i)-\phi_n(i)| \longrightarrow 0 \quad \text{as} \quad n\to\infty.$$

Hence we can choose a sequence $\{b_n\}$ satisfying

$$0<b_n \longrightarrow \infty \quad \text{and} \quad \delta_n \exp(b_n^2/2) \longrightarrow 0, \qquad (n\to\infty).$$

Denoting by $B_n$ the set $\{i: |i-n\mu|<n^{1/2}\sigma b_n\}$, we have as $n\to\infty$

$$(5) \qquad \begin{aligned} K^{-1}\sum_{i\in B_n} q_n(i)\log[q_n(i)/\phi_n(i)] &\leq \sum_{i\in B_n} q_n(i)[q_n(i)/\phi_n(i)-1] \\ &\leq \sup_{i\in B_n}|q_n(i)/\phi_n(i)-1| \\ &\leq \sup_{i\in B_n}\phi_n(i)^{-1}\sup_i|q_n(i)-\phi_n(i)| \\ &\leq (2\pi\sigma^2)^{1/2}\delta_n \exp(b_n^2/2) \longrightarrow 0. \end{aligned}$$

It follows from (4) that

$$(6) \qquad C=\sup_{n,i} n^{1/2}q_n(i)<\infty.$$

Hence we have

$$\begin{aligned} K^{-1}\log[q_n(i)/\phi_n(i)] &\leq \ln[(2\pi\sigma^2)^{1/2}C\exp((i-n\mu)^2/(2n\sigma^2))] \\ &\leq (i-n\mu)^2/(2n\sigma^2)+C_0, \end{aligned}$$

where $C_0=\max\{0, 1/2\ln(2\pi\sigma^2C^2)\}$, and therefore

$$(7) \qquad \begin{aligned} K^{-1}\sum_{i\notin B_n} q_n(i)\log[q_n(i)/\phi_n(i)] &\leq \sum_{i\notin B_n} q_n(i)[(i-n\mu)^2/(2n\sigma^2)+C_0] \\ &= E[\{(S_n-n\mu)^2/(2n\sigma^2)+C_0\}I\{|S_n-n\mu|>n^{1/2}\sigma b_n\}]. \end{aligned}$$

By the central limit theorem $(S_n-n\mu)/(n^{1/2}\sigma)\xrightarrow{d}N(0, 1)$. In view of the relation $E\{(S_n-n\mu)^2/(n\sigma^2)\}=1$, random variables $(S_n-n\mu)^2/(2n\sigma^2)+C_0$, $n\geq 1$, are uniformly integrable. Hence the right side of (7) converges to zero as $n\to\infty$. Combined with (3) and (5) this proves $\Delta_n\to 0$, $(n\to\infty)$.

Next, we suppose that $X_1$ has the span $h>1$. Then there exists an integer $r$ such that $Y_n=(X_n-r)/h$ are integer valued strongly aperiodic. The entropy of $T_n=Y_1+\cdots+Y_n$ is equal to $H_n$ and Var $Y_1=\sigma^2/h^2$. Hence from the first part of the proof we have

$$\Delta_n=\log h+1/2\log(2\pi e n\sigma^2/h^2)-H_n \longrightarrow \log h, \qquad (n\to\infty).$$

For $a>0$ and $n\geq 1$ let

$$A_n(a)=\{i: |(i-n\mu)/(n^{1/2}\sigma)|<(2a\ln n)^{1/2}\}.$$

**Lemma 2.** *If $X_1$ is strongly aperiodic and $E|X_1|^3<\infty$, then for every $a>0$*

$$\sup_{i\in A_n(a)} |q_n(i)/\phi_n(i)-1| = O(n^{-1/2+a}).$$

**Proof.** It is known that under the assumption

$$\sup_i |q_n(i)-\phi_n(i)| = O(n^{-1})$$

holds ([3], p. 197). Hence

$$\sup_{i\in A_n(a)} |q_n(i)/\phi_n(i)-1| \le \sup_{i\in A_n(a)} \phi_n(i)^{-1} \sup |q_n(i)-\phi_n(i)|$$
$$= O(n^{1/2+a})O(n^{-1}) = O(n^{-1/2+a}).$$

**Lemma 3.** *If $X_1$ is strongly aperiodic and $E|X_1|^3<\infty$, then*

(8) $$P(|S_n-n\mu|/(n^{1/2}\sigma)>(2a\ln n)^{1/2}) = O(n^{-a}(\ln n)^{-1/2})$$

*and*

(9) $$E((S_n-n\mu)^2/(n^{1/2}\sigma)I\{|S_n-n\mu|>\sigma(2an\ln n)^{1/2}\} = O(n^{-a}(\ln n)^{1/2}).$$

**Proof.** Let us write $G_n(x)=P(|S_n-n\mu|>n^{1/2}\sigma x)$. Then by the non-uniform form of Berry-Essen inequality ([3], p. 125)

(10) $$|G_n(x)-G(x)| \le C_1/[n^{1/2}(1+|x|^3)]$$

for some constant $C_1>0$, where $G(x)=\int_{|y|>x}\phi(y)dy$. Hence

$$G_n((2a\ln n)^{1/2}) \le G((2a\ln n)^{1/2}) + C_1/[a^{1/2}(2n\ln n)^{3/2}]$$
$$\le n^{-a}(\pi a\ln n)^{-1/2} + C_1 n^{-1/2}(2a\ln n)^{-3/2}$$
$$= O(n^{-a}(\ln n)^{-1/2}).$$

This proves (8).

Since

$$E(X^2 I\{|X|>x\}) = x^2 P(|X|>x) + 2\int_x^\infty yP(|X|>y)dy,$$

the left side of (9) is written as

$$(2a\ln n)G_n((2a\ln n)^{1/2}) + 2\int_{(2a\ln n)^{1/2}}^\infty yG_n(y)dy.$$

Because of (10) this is dominated by

$$(2a\ln n)G_n((2a\ln n)^{1/2}) + 2\int_{(2a\ln n)^{1/2}}^\infty yG(y)dy + 2C_1 n^{-1/2}\int_{(2a\ln n)^{1/2}}^\infty y^{-2}dy.$$

By (8) the first term is $O(n^{-a}(\ln n)^{1/2})$. Since $yG(y)<2\phi(y)$, the second term is dominated by $2G((2a\ln n)^{1/2})=O(n^{-a}(\ln n)^{-1/2})$. Obviously the last term is $O(n^{-1/2})$. Thus we have (9).

**Proof of Theorem 2.** For any $\varepsilon$ with $0<\varepsilon<1/2$, let $k=[1/\varepsilon-1]$, and let $a_j=j\varepsilon/2$, for $1\leq j\leq k$. Then it is easy to verify that $0<a_1<\varepsilon$, $0<a_j-a_{j-1}<\varepsilon$ for $2\leq j\leq k$, and $1/2-\varepsilon<a_k\leq(1-\varepsilon)/2$.

Let $E_1=A_n(a_1)$, $E_j=A_n(a_j)\cap A_n(a_{j-1})^c$ for $2\leq j\leq k$ and $E_{k+1}=A_n(a_k)^c$. It follows from Lemma 2 that

$$K^{-1}\sum_{i\in E_1}q_n(i)\log[q_n(i)/\phi_n(i)]\leq\sum_{i\in E_1}q_n(i)\sup_{i\in E_1}|q_n(i)/\phi_n(i)-1|$$
$$\leq O(n^{-1/2+a_1})=o(n^{-1/2+\varepsilon}),\qquad(n\to\infty). \tag{11}$$

By Lemma 2 and Lemma 3, we have for $2\leq j\leq k$

$$K^{-1}\sum_{i\in E_j}q_n(i)\log[q_n(i)/\phi_n(i)]\leq\sum_{i\in E_j}q_n(i)\sup_{i\in E_j}|q_n(i)/\phi_n(i)-1|$$
$$\leq O(n^{-1/2+a_j})o(n^{-a_{j-1}})$$
$$\leq o(n^{-1/2+a_j-a_{j-1}})=o(n^{-1/2+\varepsilon}),\qquad(n\to\infty). \tag{12}$$

Finally from Lemma 3 and (6), we have

$$K^{-1}\sum_{i\in E_{k+1}}q_n(i)\log[q_n(i)/\phi_n(i)]\leq\sum_{i\in E_{k+1}}q_n(i)\ln[C(\phi_n(i)n^{1/2})^{-1}]$$
$$\leq\sum_{i\in E_{k+1}}q_n(i)[C_0+(i-n\mu)^2/(2n\sigma^2)]$$
$$\leq O(n^{-a_k}(\ln n)^{-1/2})+O(n^{-a_k}(\ln n)^{1/2})$$
$$=o(n^{-1/2+\varepsilon}),\qquad(n\to\infty). \tag{13}$$

In view of (3), the inequalities (11), (12) and (13) prove that

$$\Delta_n=\sum_i q_n(i)\log[q_n(i)/\phi_n(i)]=o(n^{-1/2+\varepsilon}),\qquad(n\to\infty).$$

**Remark.** At first the author proved Theorem 2 by direct calculation using Stirling's formula. The original proof was elementary but lengthy ([6]). The idea of the present proof is due to Prof. T. Mori. The author is very grateful to him for his kind advice in preparing this paper in the present form.

## References


[1] Barron, A.R.: Entropy and central limit theorem, *Ann. of Prob.*, 14 (1986), 336-342.

[2] Linnik, Yu.V.: An information theoretic proof of the central limit theorem with the Lindeberg condition, *Theory of Prob. Appl.*, 4 (1959), 288-298.

[3] Petrov, V.V.: *Sums of Independent Random Variables*, Springer, New York, (1975).

[4] Shannon, C.E. and W. Weaver: *The Mathematical Theory of Communication*, University of Illinois Press (1949).

[5] Shimizu, R.: *Central Limit Theorem*, Kyoiku Shuppan, Tokyo, (1973). (in Japanese).

[6] Takano, S.: On the entropy and the limit theorem, *7-th Symposium of Japan Operations Research Society*, Sept, (1979). (in Japanese).

Department of Mathematics
Yokohama National University
Tokiwadai, Hodogaya-ku
Yokohama
240 Japan