BULLETIN OF THE
INSTITUTE OF MATHEMATICS
ACADEMIA SINICA
Volume 25, Number 4, December 1997

AN EXTENSION OF A STRONG LIMIT THEOREM ON RANDOM SELECTION

BY

LIU WEN (劉文) AND JIANG DONGMING (蔣東明)

Abstract. In this paper, we introduce likelihood ratio to be a measure of the deviation of the dependent sequence of discrete random variables, relative to the type of independence. By restricting the deviation, a subset of sample space is determined, and on this subset, a class of limit theorems represented by inequalities are given. Furthermore, by allowing the selection function to take values in the interval [0,1], the conception of random selection is generalized. A strong limit theorem on gambling system is a simple corollary of the conclusion of this paper.

1. Introduction. The notion of random selection originates from gambling. Consider a sequence of Bernoulli trials and suppose that at each trial the bettor has the free choice of whether or not to bet. A theorem on gambling system asserts that under any system the successive bets form a sequence of Bernoulli trial with unchanged probability for success. The importance of this statement was recognized by von Mises, who introduced the impossibility of a successful gambling system as a fundamental axiom. (cf.[3], p.91; [7], p.186). This topic was discussed still further by Kolmogorov in [9]. The connection between this topic and the martingale theory were discussed by several authors (cf.[2], p.316; [5], pp.299-302; [8], p.328; [24], p.452). The purpose of this paper is to extend the discussion to the case

Received by the editors July 19, 1994 and in revised form November 4, 1996.
AMS 1980 Subject Classification: 60F15, 60F99.
Key words and phrases: Limit theorem, likelihood ratio, random selection, gambling system.

of dependent random variables, in virtue of the notion of likelihood ratio and by using an analytic technique. We also extend the notion of random selection by allowing the selection function to take values in [0,1].

Let $S = \{t_0, t_1, \ldots\}$ be a countable (finite or denumerable) set, $\{X_n, n \geq 1\}$ a sequence of random variables taking values in S with the joint distribution

(1)

$$P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n) = p(x_1, \ldots, x_n) > 0 \ x_i \in S, \ 1 \leq i \leq n.$$

Let

$$\boldsymbol{\alpha}_i = (p_i(t_0), p_i(t_1), \ldots, p_i(t_n), \ldots), \quad p_i(t_n) > 0, \ i = 1, 2 \ldots \tag{2}$$

be a sequence of distributions on S. In order to indicate the deviation between $\{X_n, n \geq 1\}$ and a sequence of independent random variables with distribution (2), we introduce the likelihood ratio of $\{X_n, n \geq 1\}$, relative to the product distribution $\Pi_{i=1}^n p_i(x_i)$, as follows

$$\gamma_n(\omega) = \frac{\Pi_{i=1}^n p_i(X_i)}{p(X_1, \ldots, X_n)}. \tag{3}$$

where ω is a sample point, and $X_n(\omega)$ is denoted by X_n for the brevity. The product distribution $\Pi_{i=1}^n p_i(x_i)$ is called the independent reference of $\{X_i, 1 \leq i \leq n\}$. In order to extend the idea of random selection (cf.[22], p.277), we first give a set of functions $f_n(x_1, \ldots, x_n)$ defined on $S^n (n = 1, 2, \ldots)$ and taking values in the interval [0,1], which we call the [0,1]-valued selection functions. Then let

$$Y_{n+1} = f_n(X_1, \ldots, X_n), \quad Y_1 \equiv y_1, \quad 0 < y_1 < 1. \tag{4}$$

2. Main results.

Theorem 1. *Let $\{X_n, n \geq 1\}$ be a sequence of random variables with distribution (1), $c \geq 0$ be a constant, $t_n \in S, \gamma_n(\omega)$ and Y_n are defined respectively by (3) and (4). $\delta_i(\cdot)$ be the Kronecker delta function on S,i.e.*

$$(5) \qquad \delta_i(j) = \delta_{ij} \qquad (i, j \in S).$$

Let

$$(6) \qquad D(c) = \{\omega : \liminf_n (1/\sum_{i=1}^n Y_i) \ln \gamma_n(\omega) > -c, \sum_{i=1}^\infty Y_i = \infty\}.$$

Then

$$(7) \quad \limsup_n (1/\sum_{i=1}^n Y_i) \sum_{i=1}^n Y_i[\delta_{t_k}(X_i) - p_i(t_k)] \leq c + 2\sqrt{c} \quad \text{a.e., } \omega \in D(c);$$

When $0 < c < 1$,

$$(8) \quad \liminf_n (1/\sum_{i=1}^n Y_i) \sum_{i=1}^n Y_i[\delta_{t_k}(X_i) - p_i(t_k)] > -2\sqrt{c} \quad \text{a.e., } \omega \in D(c).$$

Remark 1. In order to explain the real meaning of the [0,1]-valued random selection we consider the following gambling model. Let $\{X_n, n \geq 1\}$ be independent random variables with distribution (2). Interpret X_n as the result of the n-th trial, and Y_n as the stake which the bettor puts down at the n-th trial, the type of which may change at each step. Let μ_n denote the gain of the bettor at the n-th trial. Suppose that by the gambling rules, $\mu_n = Y_n$ if the event $\{X_n = t_k\}$ occurs at the n-th trial; and $\mu_n = 0$ if $\{X_n = t_k\}$ does not occur. To sum up, $\mu_n = Y_n \delta_{t_k}(X_n)$. Let the entrance fee that the bettor pays at the n-th trial be μ'_n. Also suppose that μ'_n is proportional to Y_n, i.e., $\mu'_n = b_n Y_n$, where b_n is a constant. Thus the $\sum_{i=1}^n Y_i \delta_{t_k}(X_i)$ represents the total gain in the first n independent trials, $\sum_{i=1}^n b_i Y_i$ the accumulated entrance fees, and $\sum_{i=1}^n Y_i[\delta_{t_k}(X_i) - b_i]$ the accumulated net gain. The Corollary 2 of above theorem shows that if $b_n = p_n(t_k)(n = 1, 2, 3, \ldots)$, then the accumulated net gain in the first n independent trials is to be of smaller order of magnitude than the accumulated stakes $\sum_{i=1}^n Y_i$ as the later tends to infinite, and the formula (39) may be regarded as an extension of the classical definition of "fairness" of game of chance. The above theorem extends the discussion to the dependent case.

It should be mentioned that in the above remark the stake and entrance fee at the n-th trial are all random variables determined by $X_1, X_2,$

$\ldots, X_{n-1} (n \geq 2)$. For the discussion of nonrandom situation the readers may consult [1], [4], [6], [10], and [23].

Remark 2. Obviously $\gamma_n(\omega) \equiv 0$ *a.e.* if and only if $\{X_n, n \geq 1\}$ are independent random variables having the distribution (2), and it will be shown in (27) that in general case $\limsup_n (1/\sum_{i=1}^n Y_i) \ln \gamma_n(\omega) \leq 0$ *a.e.* in D. Hence $\gamma(\omega) = \liminf_n (1/\sum_{i=1}^n Y_i) \ln \gamma_n(\omega)$ can be used to measure the deviation between $\{X_n, n \geq 1\}$ and the independent reference with the distribution (2) on D. The smaller $|\gamma(\omega)|$ is, the smaller the deviation is. Roughly speaking, the first condition in (6) can be regarded as a restriction on the deviation between $\{X_n, n \geq 1\}$ and the independent reference with the distribution (2). The above theorem states that under such restriction the ratio $(1/\sum_{i=1}^n Y_i) \sum_{i=1}^n Y_i[\delta_{t_k}(X_i) - p_i(t_k)]$ is correspondingly restricted. Formulas (7) and (8) give, respectively, the upper and lower bounds of its superior and inferior limits with respect to c. As c is small, the absolute values of these bounds are small also. Summarily, the situations of this theorem are similar in some sense to those appeared in the theorem on the stability of solution of differential equation.

Remark 3. The above theorem is a kind of small deviation theorems, which investigate the approximation of the true joint distribution of $\{X_n\}$ by a product distribution (reference distribution). Though by the definition of a.e. convergence on a measurable set (7) and (8) hold trivially if $P(D(c)) = 0$, this case goes beyond the scope of small deviation. For this theorem to be of any use, one needs to choice proper reference product distribution and give some sufficient conditions to establish that $P(D(c)) > 0$. We shall discuss this problem in section 4.

Proof. Let $\Omega = [0, 1]$, the class of all Borel measurable set $\mathcal{F}$ and the Lebesgue measure P be the probability space to be considered. We first give, in the above probability space, a realization of the sequence of random variables with distribution (1). Split the interval $[0, 1)$ into countable right-semiopen intervals at the ratio

$$p(t_0) : p(t_1) : \ldots : p(t_i) : \ldots$$

and denote them by $d_{x_1}, (x_1 = t_0, t_1, \ldots)$. These intervals will be called d-intervals of first order. Suppose the $(n-1)$st $(n \geq 2)$ order d-intervals $d_{x_1 \ldots x_{n-1}}, (x_i \in S, 1 \leq i \leq n-1)$ have been defined, then split each right-semiopen d-intervals $d_{x_1 \ldots x_{n-1}}$ into countable right-semiopen intervals $d_{x_1 \ldots x_{n-1} x_n} (x_n = t_0, t_1, \ldots)$ at the ratio

$$p(x_1, \ldots, x_{n-1}, t_0) : p(x_1, \ldots, x_{n-1}, t_1) : \ldots : p(x_1, \ldots, x_{n-1}, t_i) : \ldots,$$

the d-intervals of the nth order are created. It is easy to see that for $n \geq 1$.

$$P(d_{x_1, \ldots, t_n}) = p(x_1, \ldots, x_n) \tag{9}$$

define. For $n \geq 1$, a random variable $X_n : [0,1) \to S$ as follows

$$X_n(\omega) = x_n \quad \text{if } \omega \in d_{x_1 \ldots x_n}. \tag{10}$$

Then we have

$$\{\omega : X_1 = x_1, \ldots, X_n = x_n\} = d_{x_1 \ldots x_n}; \tag{11}$$

$$P(d_{x_1 \ldots x_n}) = P(X_1 = x_1, \ldots, X_n = x_n) = p(x_1, \ldots, x_n). \tag{12}$$

Thus, $\{X_n, n \geq 1\}$ has distribution (1). It is obvious that $\gamma_n(\omega)$ and $Y_{n+1}(\omega)$ are constants in each n-th order d-interval. Moreover,

$$\gamma_n(\omega) = \frac{\Pi_{i=1}^n p_i(x_i)}{p(x_1 \ldots x_n)}, \quad \omega \in d_{x_1 \ldots x_n}. \tag{13}$$

Now, we shall give a proof for this theorem according to the above realization of $\{X_n, n \geq 1\}$. For the need of proof, we construct a auxiliary function. Let $\lambda > 0$ be a constant. We first define a sequence of random variables $\lambda_n(t_j, \omega)(t_j \in S, n = 1, 2, \ldots)$ by the following equations:

$$\frac{\lambda_n(t_k, \omega)[1 - p_n(t_k)]}{p_n(t_k)[1 - \lambda_n(t_k, \omega)]} = \lambda^{Y_n(\omega)}, \quad \omega \in [0,1) \tag{14}$$

that is,

(15) $$\lambda_n(t_k,\omega) = \frac{\lambda^{Y_n(\omega)} p_n(t_k)}{1+(\lambda^{Y_n(\omega)}-1)p_n(t_k)}, \omega \in [0,1)$$

and when $t_j \neq t_k$, let

(16) $$\lambda_n(t_j,\omega) = \frac{1-\lambda_n(t_k,\omega)}{1-p_n(t_k)} p_n(t_j); \quad \omega \in [0,1)$$

By (15),

(17) $$\frac{1-\lambda_n(t_k,\omega)}{1-p_n(t_k)} = \frac{1}{1+(\lambda^{Y_n}-1)p_n(t_k)}$$

It is easy to see that on each $(n-1)$st order d-interval (the interval [0,1) is named the 0-th order d-interval) $\lambda_n(t_j,\omega)$ is a constant and $\{\lambda_n(t_j,\omega), j = 0,1,\ldots\}$ is a probability distribution on S.

Similarly, split the interval [0,1], according to the ratio $\lambda_1(t_0,\omega) : \lambda_1(t_1,\omega) : \ldots$, into countable right-semiopen intervals, and denote them by $D_{x_1}(x_1 = t_0, t_1, \ldots)$, these intervals will be called D-intervals of first order. Proceeding inductively, suppose the $(n-1)$st order $(n \geq 2)$ D-interval $D_{x_1\ldots x_{n-1}}(x_i \in S, 1 \leq i \leq n-1)$ has been defined, then split $D_{x_1\ldots x_{n-1}}$ into countable right-semiopen intervals according to the ratio

$$\lambda_n(t_0,\omega) : \lambda_n(t_1,\omega) : \ldots, (\omega \in d_{x_1\ldots x_{n-1}}),$$

and denote then by $D_{x_1\ldots x_n}(x_n = t_0, t_1, \ldots)$. By induction, when $n \geq 1$, we have

(18) $$P(D_{x_1\ldots x_n}) = \Pi_{i=1}^n \lambda_i(x_i,\omega), \ \omega \in d_{x_1\ldots x_n}.$$

Let $d^-_{x_1\ldots x_n}$ and $d^+_{x_1\ldots x_n}$ be, respectively, the left and right end points of $d_{x_1\ldots x_n}$, define $D^-_{x_1\ldots x_n}$ and $D^+_{x_1\ldots x_n}$ similarly. Let Q be the set of end points of all d-intervals. Now we define an function f on Q as follows:

(19) $$f(d^-_{x_1\ldots x_n}) = D^-_{x_1\ldots x_n}, \ f(d^+_{x_1\ldots x_n}) = D^+_{x_1\ldots x_n}$$

If $\omega \in [0,1] - Q$, let

$$f(\omega) = \sup\{f(t), \ t \in Q \cap [0,\omega)\}$$

Obviously, f is a increasing function on [0,1]. Let

$$T_n(\lambda,\omega) = \frac{P(D_{x_1\ldots x_n})}{P(d_{x_1\ldots x_n})},\ \omega \in d_{x_1\ldots x_n}.$$

By (13) and (18)-(21),

$$\begin{aligned} T_n(\lambda,\omega) &= \frac{f(d^+_{x_1\ldots x_n}) - f(d^-_{x_1\ldots x_n})}{d^+_{x_1\ldots x_n} - d^-_{x_1\ldots x_n}} = \gamma_n(\omega)\Pi_{i=1}^n \frac{\lambda_i(x_i)}{p_i(x_i)} \\ &= \gamma_n(\omega)\Pi_{i=1}^n \Pi_{j\in S}\left[\frac{\lambda_i(j,\omega)}{p_i(j)}\right]^{\delta_j(x_i)}, \omega \in d_{x_1\ldots x_n}. \end{aligned} \tag{22}$$

By (10) and (14)-(17)

$$\begin{aligned} T_n(\lambda,\omega) &= \gamma_n(\omega)\Pi_{i=1}^n \Pi_{j\in S}\left[\frac{\lambda_i(j,\omega)}{p_i(j)}\right]^{\delta_j(X_i)} \\ &= \gamma_n(\omega)\Pi_{i=1}^n\left\{\left[\frac{\lambda_i(t_k,\omega)}{p_i(t_k)}\right]^{\delta_{t_k}(X_i)} \Pi_{j\neq t_k}\left[\frac{1-\lambda_i(t_k,\omega)}{1-p_i(t_k)}\right]^{\delta_j(X_i)}\right\} \\ &= \gamma_n(\omega)\Pi_{i=1}^n\left\{\left[\frac{\lambda_i(t_k,\omega)}{p_i(t_k)}\right]^{\delta_{t_k}(X_i)} \left[\frac{1-\lambda_i(t_k,\omega)}{1-p_i(t_k)}\right]^{1-\delta_{t_k}(X_i)}\right\} \\ &= \gamma_n(\omega)\Pi_{i=1}^n\left\{\left[\frac{i_i(t_k,\omega)(1-p_i(t_k))}{p_i(t_k)(1-\lambda_i(t_k,\omega))}\right]^{\delta_{t_k}(X_i)} \left[\frac{1-\lambda_i(t_k,\omega)}{1-p_i(t_k)}\right]\right\} \\ &= \gamma_n(\omega)\lambda^{\Sigma_{i=1}^n Y_i\delta_{t_k}(X_i)}\Pi_{i=1}^n \frac{1}{1+(\lambda^{Y_i}-1)p_i(t_k)},\ \omega\in[0,1) \end{aligned} \tag{23}$$

Let $A(t_k,\lambda)$ be the set of points of differentiability of f. Then $P(A(t_k, \lambda)) = 1$ by the theorem on the existence of the derivative of a monotone function. Let $\omega \in A(t_k,\lambda)$, $d_{x_1\ldots x_n}$ be the nth d-interval including ω. Then by a property of the derivative (cf. [3], p.423),

$$\lim_n T_n(\lambda,\omega) = \text{finite number},\ \omega \in A(t_k,\lambda). \tag{24}$$

Since $\sum_{i=1}^\infty Y_i = \infty$ when $\omega \in D(c)$, hence by (24),

$$\limsup_n (1/\sum_{i=1}^{\infty} Y_i)\ln T_n(\lambda,\omega) < 0, \quad \omega \in A(t_k,\lambda)\cap D(c). \tag{25}$$

Let

$$D = \{\omega : \sum_{i=1}^{\infty} Y_i = \infty\} \tag{26}$$

Since $T_n(1,\omega) = \gamma_n(\omega)$, by (24) we have

$$\limsup_n (1/\sum_{i=1}^{n} Y_i) \ln \gamma_n(\omega) \leq 0 \quad \text{a.e.}, \quad \omega \in D. \tag{27}$$

This is the reason that we assume $c \geq 0$ in the definition of $D(c)$. By (25), (23) and (6),

$$\begin{aligned} \limsup_n (1/\sum_{i=1}^{n} Y_i)\{\sum_{i=1}^{n} Y_i \delta_{t_k}(X_i) \ln \lambda - \sum_{i=1}^{n} \ln[1 + (\lambda^{Y_i} - 1)p_i(t_k)]\} \\ \leq c, \quad \omega \in A(t_k, \lambda) \cap D(c). \end{aligned} \tag{28}$$

Letting $\lambda > 1$, and dividing the two side of (28) by $\ln \lambda$, we have

$$\begin{aligned} \limsup_n (1/\sum_{i=1}^{n} Y_i)\left\{\sum_{i=1}^{n} Y_i \delta_{t_k}(X_i) - \sum_{i=1}^{n} \frac{\ln[1 + (\lambda^{Y_i} - 1)p_i(t_k)]}{\ln \lambda}\right\} \\ \leq c/\ln \lambda, \quad \omega \in A(t_k, \lambda) \cap D(c). \end{aligned} \tag{29}$$

In virtue of the property of superior limit:

(30) $\limsup_n (a_n - b_n) \leq d$ implys $\limsup_n (a_n - c_n) \leq \limsup_n (b_n - c_n) + d,$

and the inequalities:

(31) $1 - 1/x \leq \ln x \leq x - 1 \ (x > 0), \ x^r - 1 \leq r(x-1) \ (x > 0, 0 \leq r \leq 1)$

we have, by (29)

$$\begin{aligned} &\limsup_n (1/\sum_{i=1}^{n} Y_i) \sum_{i=1}^{n} Y_i[\delta_{t_k}(X_i) - p_i(t_k)] \\ &\leq \limsup_n (1/\sum_{i=1}^{n} Y_i) \sum_{i=1}^{n} \left\{\frac{\ln[1 + (\lambda^{Y_i} - 1)p_i(t_k)]}{\ln \lambda} - Y_i p_i(t_k)\right\} + c/\ln \lambda \\ &\leq \limsup_n (1/\sum_{i=1}^{n} Y_i) \sum_{i=1}^{n} \left[\frac{(\lambda^{Y_i} - 1)p_i(t_k)}{1 - 1/\lambda} - Y_i p_i(t_k)\right] + c/(1 - 1/\lambda) \\ &\leq \limsup_n (1/\sum_{i=1}^{n} Y_i) \sum_{i=1}^{n} Y_i p_i(t_k)(\lambda - 1) + c + c/(\lambda - 1) \\ &\leq \lambda - 1 + c + c/(\lambda - 1), \ \omega \in A(t_k, \lambda) \cap D(c). \end{aligned} \tag{32}$$

It is easy to see that the function $g(\lambda)=\lambda-1+c+c/(\lambda-1)$, $\lambda>1$, attains its smallest value $g(1+\sqrt{c})=2\sqrt{c}+c$ at $\lambda=1+\sqrt{c}$ when $c>0$. Thus it follows from (32) that

$$\limsup_n (1/\sum_{i=1}^n Y_i)\sum_{i=1}^n Y_i[\delta_{t_k}(X_i)-p_i(t_n)] \tag{33}$$
$$\le c+2\sqrt{c},\quad \omega\in A(t_k,1+\sqrt{c})\cap D(c).$$

Since $P(A(t_k,1+\sqrt{c}))=1$, (7) holds by (33) when $c>0$. In the case $c=0$, choose $\lambda_i>1$ $(i=1,2,\ldots)$, such that $\lambda_i\to 1$ (as $i\to\infty$), and let

$$A(t_k)=\cap_{i=1}^{\infty}A(t_k,\lambda_i).$$

Then for all $i\ge 1$, we have by (32).

$$\limsup_n (1/\sum_{i=1}^n Y_i)\sum_{i=1}^n Y_i[\delta_{t_k}(X_i)-p_i(t_k)]<\lambda_i-1,\ \omega\in A(t_k)\cap D(0). \tag{34}$$

Since $\lambda_i-1\to 0$ and $P(A(t_k)=1$. (7) follows from (34) as $c=0$.

Let $0<\lambda<1$, and divide the two sides of (28) by $\ln\lambda$. In virtue of the property of inferior limit

$$\liminf_n (a_n-b_n)\ge d \quad \text{implys } \liminf_n (a_n-c_n)\ge \liminf_n (b_n-c_n)+d \tag{35}$$

and by using the inequalities (31), we have

(36)

$$\liminf_n (1/\sum_{i=1}^n Y_i)\sum_{i=1}^n Y_i[\delta_{t_k}(X_i)-p_i(t_k)]$$
$$\ge \liminf_n (1/\sum_{i=1}^n Y_i)\sum_{i=1}^n \left\{\frac{\ln[1+s(\lambda^{Y_i}-1)p_i(t_k)]}{\ln\lambda}-Y_ip_i(t_k)\right\}+c/\ln\lambda$$
$$\ge \liminf_n (1/\sum_{i=1}^n Y_i)\sum_{i-1}^n \left[\frac{(\lambda^{Y_i}-1)p_i(t_k)}{1-1/\lambda}-Y_ip_i(t_k)\right]+c/(1-1/\lambda)$$
$$\ge \lambda-1+c+c/(\lambda-1),\ \omega\in A(t_k,\lambda)\cap D(c).$$

It is clear that the function $h(\lambda)=\lambda-1+c+c/(\lambda-1)$, $0<\lambda<1$, attains its greatest value $h(1-\sqrt{c})=2\sqrt{c}$ at $\lambda=1-\sqrt{c}$. Consequently, we have by (36),

$$\liminf_n (1/\sum_{i=1}^n Y_i)\sum_{i=1}^n Y_i[\delta_{t_k}(X_i) - p_i(t_k)] \tag{37}$$
$$\geq -2\sqrt{c},\ \omega \in A(t_k, 1-\sqrt{c}) \cap D(c).$$

Since $P(A(t_n, 1-\sqrt{c})) = 1$, (8) follows from (37) when $0 < c < 1$. Imitating the proof of (7) with $c = 0$, it can be shown that (8) also holds for $c = 0$.

Remark. In the above proof the analytic technique proposed by the first author was used, the crucial part of which is the application of Lebesgue's theorem on differentiability of monotone functions to the study of a.e. convergence (see [11] and [12]). In [13]-[20] this technique was used to obtain a class of strong limit theorems, some of which were represented by inequalities, and it was extended by considering the random selection system in [21]. In the proof of Theorem 1 these techniques were further spreaded by using the [0,1]-valued selection functions in place of the 0,1-valued ones in the random selection system.

Corollary 1. *Under the assumptions of theorem* 1, *we have*

$$\lim_n (1/\sum_{i=1}^n Y_i)\sum_{i=1}^n Y_i[\delta_{t_k}(X_i) - p_i(t_k)] = 0 \quad a.e.,\ \omega \in D(0). \tag{38}$$

Proof. Let $c = 0$, (38) follows from (7) and (8).

Corollary 2. *Let* $\{X_n, n > 1\}$ *be independent random variables with the distribution* (2), D *defined by* (26). *Then*

$$\lim_n (1/\sum_{i=1}^n Y_i)\sum_{i=1}^n Y_i[\delta_{t_k}(X_i) - p_i(t_k)] = 0 \quad a.e.,\ \omega \in D. \tag{39}$$

Proof. In this case. $\gamma_n(\omega) \equiv 1$ and $D(0) = D$, hence (39) follows from (38).

Remark. If $p_i(t_k) \equiv p(t_k)$ $(i = 1, 2, \ldots)$, then (39) can be rewritten as follows:

$$\lim_n (1/\sum_{i=1}^n Y_i) \sum_{i=1}^n Y_i \delta_{t_k}(X_i) = p(t_k) \quad \text{a.e.}, \omega \in D. \tag{40}$$

If $f_n(x_1, \dots, x_n)$ in (4) is restricted to take values in $\{0, 1\}$, then (40) is the well known theorem on gambling system (cf. [7], p.186).

Corollary 3. *Let* $t_k \in S, \sum_{i=1}^\infty p_i(t_k) = \infty, c \geq 0$ *be a constant. And let*

$$D(t_k, c) = \{\omega : \liminf_n [1/\sum_{i=1}^n p_i(t_k)] \ln \gamma_n(\omega) \geq -c\}. \tag{41}$$

Then

$$\begin{aligned} &\limsup_n [1/\sum_{i=1}^n p_i(t_k)] \sum_{i=1}^n p_i(t_k)[\delta_{t_k}(X_i) - p_i(t_k)] \\ &\leq c + 2\sqrt{c} \ \textit{a.e.}, \ \omega \in D(t_k, c); \end{aligned} \tag{42}$$

When $0 \leq c < 1$,

$$\begin{aligned} &\liminf_n [1/\sum_{i=1}^n p_i(t_k)] \sum_{i=1}^n p_i(t_k)[\delta_{t_k}(X_i) - p_i(t_k)] \\ &\geq -2\sqrt{c} \quad \textit{a.e.}, \ \omega \in D(t_k, c). \end{aligned} \tag{43}$$

Proof. Let $f_{i-1}(x_1, \dots, x_{i-1}) \equiv p_i(t_k)$ $(i \geq 2)$, (42) and (43) follow immediately from theorem 1.

Corollary 4. *Under the assumptions of Corollary* 3, *we have*

$$\lim_n [1/\sum_{i=1}^n p_i(t_k)] \sum_{i=1}^n p_i(t_k)[\delta_{t_k}(X_i) - p_i(t_k)] = 0 \quad \textit{a.e.}, \ \omega \in D(t_k, 0), \tag{44}$$

Corollary 5. *Let* m *be a positive integer,* $u_i \in S$, $1 \leq i \leq m$, $S_n(u_1, \dots, u_m; \omega)$ *be the number of* $(u_1, \dots, u_m)$ *in the sequence* $(X_1, \dots, X_m)$, $(X_2, \dots, X_{m+1}), \dots, (X_{n-m+1}, \dots, X_n)$ $(n \geq m)$, *denote* $\delta_{u_1}(i_1) \cdots \delta_{u_n}(i_m)$ *by* $\delta_{u_1 \cdots u_m}(i_1, \dots, i_m)$, *that is*

$$\delta_{u_1 \cdots u_m}(i_1, \dots, i_m) = \begin{cases} 1, \textit{if } (i_1, \dots, i_m) = (u_1, \dots, u_m), \\ 1, \textit{if } (i_1, \dots, i_m) \neq (u_1, \dots, u_m). \end{cases}$$

Also let $t_k \in S$, $c \geq 0$ be constants, $D(u_1, \dots, u_m; c)$ the set of sample points satisfying the following conditions:

$$\lim_n S_n(u_1, \dots, u_m; \omega) = \infty; \tag{45}$$

$$\liminf_n [1/S_n(u_1, \dots, u_m; \omega)] \ln \gamma_n(\omega) > -c. \tag{46}$$

Then

$$\begin{aligned} &\limsup_n [1/S_n(u_1, \dots, u_m; \omega)][S_n(u_1, \dots, u_m, t_k; \omega) \\ &- \sum_{i=m+1}^{n} p_i(t_k)\delta_{u_1 \cdots u_m}(X_{i-m}, X_{i-m+1}, \dots, X_{i-1})] \leq c + 2\sqrt{c} \\ &\quad a.e., \ \omega \in D(u_1, \dots, u_m; c). \end{aligned} \tag{47}$$

When $0 \leq c < 1$,

$$\begin{aligned} &\liminf_n [1/S_n(u_1, \dots, u_m; \omega)][S_n(u_1, \dots, u_m, t_k; \omega) \\ &- \sum_{i=m+1}^{n} p_i(t_k)\delta_{u_1 \cdots u_m}(X_{i-m}, X_{i-m+1}, \dots, X_{i-1})] \geq -2\sqrt{c} \\ &\quad \text{a.e.}, \ \omega \in D(u_u, \dots, u_m; c). \end{aligned} \tag{48}$$

Proof. In theorem 1, let $\gamma_n = 0$ and

$$f_i(x_1, \dots, x_i) \equiv 0 \quad \text{if } 0 \leq i \leq m-1;$$

$$f_{i-1}(x_1, \dots, x_{i-1}) = \delta_{u_1 \cdots u_m}(x_{i-m}, \dots, x_{i-1}), \quad \text{if } i \geq m+1.$$

Then

$$Y_i \equiv 0, \quad \text{if } 1 \leq i \leq m;$$

$$Y_i = \delta_{u_1 \dots u_m}(X_{i-m}, \dots, X_{i-1}), \quad \text{if } i \geq m+1.$$

Now we have

$$\begin{aligned} \sum_{i=1}^{n} Y_i \delta_{t_k}(X_i) &= \sum_{i=m+1}^{n} \delta_{u_1 \cdots u_m}(X_{i-m}, \dots, X_{i-1}\delta_{t_k}(X_i) \\ &= S_n(u_1, \dots, u_m, t_k; \omega), \end{aligned}$$

$$\sum_{i=1}^{n} Y_i = \sum_{i=m+1}^{n} \delta_{u_1\cdots u_m}(X_{i-m},\ldots,X_{i-1}) = S_{n-1}(u_1,\ldots,u_m,\omega).$$

Noticing that $S_n(u_1,\ldots,u_m;\omega) - S_{n-1}(u_1,\ldots,u_m;\omega) = 0$ or 1 (47) and (48) can be derived respectively from (7) and (8).

Corollary 6. *Under the assumptions of Corollary* 5, *we have*

$$\lim_n [1/S_n(u_1,\ldots,u_m;\omega)][S_n(u_1,\ldots,u_m,t_k;\omega) - \sum_{i=m+1}^{n} p_i(t_k)\delta_{u_1\cdots u_m}(X_{i-m},X_{i-m+1},\ldots,X_{i-1})] = 0 \tag{49}$$

$$a.e.,\ \omega \in D(u_1,\ldots,u_m;0).$$

3. A special case of independent reference. In Theorem 1 the estimations of the bounds of the superior and inferior limits of $(1/\sum_{i=1}^{n} Y_i) \sum_{i=1}^{n} Y_i[\delta_{t_k}(X_i) - p_i(t_k)]$ are given, where t_k is a individual value of X_i. In some cases of independent reference the above technique can be used to give estimations of corresponding bounds of $(1/\sum_{i=1}^{n} Y_i)\sum_{i=1}^{n} Y_i(X_i - m_i)$, where m_i is the expectation of the distribution (2). As an example, the case that (2) is Poisson distribution will be studied in the following theorem.

Theorem 2. *Let S be the set of nonnegative integers, $\{X_n, n \geq 1\}$ a sequence of random variables with distribution* (1) . *Y_n defined by* (4), *$\lambda_i > 0 (i = 1,2,\ldots)$ constants, $g(\lambda_i,j) = (1/j!)e^{-\lambda_i}\lambda_i^j (j = 0,1,2,\ldots)$ a Poisson distribution with parameter $\lambda_i, \gamma_n(\omega)$ be the likelihood ratio of $\{X_i, 1 \leq i \leq n\}$ relative to the product distribution $\Pi_{i=1}^{n} g(\lambda_i,x_i)$ $(x_i \in S)$, that is*

$$\gamma_n(\omega) = [\Pi_{i=1}^{n} g(\lambda_i,X_i)]/p(X_i,\ldots,X_n). \tag{50}$$

And let $M > 0$, $0 \leq c < 1$ be constants. $D(c)$ the set of sample points satisfying the following conditions:

$$\sum_{i=1}^{\infty} Y_i = \infty; \tag{51}$$

$$\limsup_n (1/\sum_{n=1}^n Y_i) \sum_{i=1}^n \lambda_i \le M; \tag{52}$$

$$\liminf_n (1/\sum_{i=1}^n Y_i) \ln \gamma_n(\omega) \ge -c. \tag{53}$$

Then

$$\limsup_n (1/\sum_{i=1}^n Y_i) \sum_{i=1}^n Y_i(X_i - \lambda_i) \le 2\sqrt{Mc} + c \quad \text{a.e., } \omega \in D(c); \tag{54}$$

when $0 \le c < M$,

$$\liminf_n (1/\sum_{i=1}^n Y_i) \sum_{i=1}^n Y_i(X_i - \lambda_i) \ge -2\sqrt{Mc} \quad \text{a.e., } \omega \in D(c). \tag{55}$$

Proof. Use the realization of $\{X_n, n \ge 1\}$, which was created in the proof of theorem 1 (let $t_k = k, k = 0, 1, \ldots,$). Let $\mathcal{A}$ denotes the collection of d-intervals of all orders, and $\lambda > 0$, $0 \le y_1 \le 1$ be constants. Denote

$$y_n = f_{n-1}(x_1, \ldots, x_{n-1}) = Y_n(\omega), \quad \omega \in d_{x_1,\ldots,x_{n-1}}, \ n \ge 2. \tag{56}$$

Define a set function on $\mathcal{A}$ as follows:

$$\mu(d_{x_1\ldots x_n}) = \lambda^{\sum_{i=1}^n x_i y_i} \Pi_{i=1}^n [\exp(-\lambda^{y_i}\lambda_i)\lambda_i^{x_i}/x_i!]; \tag{57}$$

$$\mu([0,1)) = \sum_{x_1=0}^{\infty} \mu(d_{x_1}). \tag{58}$$

It is easy to see that μ is an additive set function on $\mathcal{A}$. Hence there exits an increasing function f_λ defined on [0,1] such that, for any $d_{x_1\ldots x_n}$,

$$\mu_n(d_{x_1\ldots x_n}) = f_\lambda(d^+_{x_1\ldots x_n}) - f_\lambda(d^-_{x_1\ldots x_n}). \tag{59}$$

Let

$$T_n(\lambda,\omega) = \frac{f_\lambda(d^+_{x_1\ldots x_n}) - f_\lambda(d^-_{x_1\ldots x_n})}{(d^+_{x_1\ldots x_n}) - (d^-_{x_1\ldots x_n})} = \frac{\mu(d_{x_1\ldots x_n})}{p(x_1\ldots x_n)}, \quad \omega \in d_{x_1\ldots x_n}. \tag{60}$$

Let $A(\lambda)$ be the set of all points of differentiability of f_λ. Imitating (25), we have

$$\limsup_n (1/\sum_{i=1}^n Y_i) \ln T_n(\lambda,\omega) \leq 0, \quad \omega \in A(\lambda) \cap D(c). \tag{61}$$

By (10), (50) and (56)-(60).

$$\ln T_n(\lambda,\omega) = \sum_{i=1}^n X_i Y_i \ln \lambda + \sum_{i=1}^n \lambda_i (1-\lambda^{Y_i}) + \ln \gamma_n(\omega). \tag{62}$$

By (61), (60) and (53),

$$\begin{aligned} &\limsup_n (1/\sum_{i=1}^n Y_i) \Big[\sum_{i=1}^n X_i Y_i \ln \lambda - \sum_{i=1}^n \lambda_i (\lambda^{Y_i} - 1) \Big] \leq c, \\ &\omega \in A(\lambda) \cap D(c). \end{aligned} \tag{63}$$

Letting $\lambda > 1$, dividing the two sides of (63) by $\ln \lambda$, and using the property of superior limit (30), it can be obtained that

$$\begin{aligned} &\limsup_n (1/\sum_{i=1}^n Y_i) \sum_{i=1}^n Y_i (X_i - \lambda_i) \\ &< \limsup_n (1/\sum_{i=1}^n Y_i) \sum_{i=1}^n \Big[\frac{Y_i(\lambda^{Y_i}-1)}{\ln \lambda} - Y_i \lambda_i \Big] + c/\ln\lambda, \\ &\omega \in A(\lambda) \cap D(c). \end{aligned} \tag{64}$$

By (31), (64) and (52),

$$\begin{aligned} &\limsup_n (1/\sum_{i=1}^n Y_i) \sum_{i=1}^n Y_i (X_i - \lambda_i) \\ &\leq \limsup_n (1/\sum_{i=1}^n Y_i) \sum_{i=1}^n \Big[\frac{(\lambda-1) Y_i \lambda_i}{1-1/\lambda} - Y_i \lambda_i \Big] + \frac{c}{1-1/\lambda} \\ &\leq (\lambda-1)M + c + c/(\lambda-1), \ \omega \in A(\lambda) \cap D(c). \end{aligned} \tag{65}$$

It is easy to see that when $c > 0$, the function $g(\lambda) = (\lambda-1)M + c + c/(\lambda-1)$, $\lambda > 1$, attains, at $\lambda = 1 + \sqrt{c/M}$, its smallest value $g(1+\sqrt{c/M}) = 2\sqrt{Mc} + c$. Hence, imitating the proof of (8), (54) holds by (65) when $c \geq 0$.

Letting $0 < \lambda < 1$, dividing the two sides of (63) by $\ln \lambda$ and using the property of inferior limit (35), we have

$$\begin{aligned} &\liminf_n (1/\sum_{i=1}^n Y_i) \sum_{i=1}^n Y_i(X_i - \lambda_i) \\ (66) \qquad &\geq \liminf_n (1/\sum_{i=1}^n Y_i) \sum_{i=1}^n \left[\frac{\lambda_i(\lambda^{Y_i} - 1)}{\ln \lambda} - Y_i \lambda_i \right] + c/\ln \lambda, \\ &\omega \in A(\lambda) \cap D(c). \end{aligned}$$

By using the inequalities (31), it follows from (66) and (52) that

$$\begin{aligned} &\liminf_n (1/\sum_{i=1}^n Y_i) \sum_{i=1}^n Y_i(X_i - \lambda_i) \\ (67) \qquad &\geq \liminf_n (1/\sum_{i=1}^n Y_i) \sum_{i=1}^n \left[\frac{(\lambda - 1) Y_i \lambda_i}{\cdot 1 - 1/\lambda} - Y_i \lambda_i \right] + c/(\lambda - 1) \\ &\geq (\lambda - 1)M + c/(\lambda - 1), \quad \omega \in A(\lambda) \cap D(c). \end{aligned}$$

It is clear that the function $h(\lambda) = (\lambda - 1)M - c/(\lambda - 1)$, $0 < \lambda < 1$, attains, at $\lambda = 1 - \sqrt{c/M}$, its greatest value $h(1 - \sqrt{c/M}) = -2\sqrt{Mc}$ when $0 < c < M$. Consequently imitating the proof of (9), (55) holds by (67) when $0 \leq c < M$.

Remark. It is easy to see that the fact that the Poisson distribution depends on a parameter plays an important role in the above proof . This is the reason that we choose the Poisson distribution as the reference in Theorem 2. Similarly, the other discrete distributions depending on a parameter, such as the geometrical distribution and Pascal distribution, may also be chosen as the reference to obtain the corresponding formulas. We shall discuss this problem in another article.

Corollary 1. *Under the assumptions of Theorem 2, we have*

$$(68) \qquad \lim_n (1/\sum_{i=1}^n Y_i) \sum_{i=1}^n Y_i(X_i - \lambda_i) = 0 \quad \text{a.e.,} \quad \omega \in D(0).$$

Proof. Putting $c = 0$, (68) follows from (54) and (55).

Corollary 2. *Let $\{X_n, n \geq 1\}$ be independent random variables having Poisson distributions with parameters $\{\lambda_n, n \geq 1\}$, D^* the set of sample points satisfying* (51) *and* (52). *Then*

$$\lim_n (1/\sum_{i=1}^n Y_i) \sum_{i=1}^n Y_i(X_i - \lambda) = 0 \quad \text{a.e.,} \quad \omega \in D^*. \tag{69}$$

proof. In this case, $\gamma_n(\omega) \equiv 1$, and $D(0) = D^*$. Hence, (69) follows from (68).

Now, we shall provide a nontrivial example for which the condition $P(D(c)) > 0$ can be satisfied.

4. Application to the approximation of Markov chains.

Finally we give some applications of the above theorems to the independent approximation of Markov chains.

Theorem 3. *Let $\{X_n, n \geq 1\}$ be a nonhomogeneous Markov chain with state space $S = \{0, 1, 2, \ldots\}$, initial distribution*

$$q(i) = P(X_1 = i) > 0, \quad i \in S \tag{70}$$

and transition probabilities

$$p_n(i,j) = P(X_n = j | X_{n-1} = i) > 0,\ i,j \in S,\ n \geq 2, \tag{71}$$

and $\gamma_n(\omega)$ and Y_n defined by (3) *and* (4) *respectively. Let $k \in S$, and $S_n(k, \omega)$ be the number of k in the sequence $X_1, X_2, \ldots, X_n$, that is,*

$$S_n(k, \omega) = \sum_{i=1}^n \delta_k(X_i). \tag{73}$$

If for all $i, j \in S$,

$$\liminf_n p_n(j)/p_n(i,j) \geq d \quad \textit{uniformly,} \tag{74}$$

where $0 < d \leq 1$ is a constant, then

$$\limsup_n (1/n)[S_n(k,\omega) - \sum_{i=1}^n p_i(k)] \le 2\sqrt{-\ln d} - \ln d \quad \text{a.e.}; \tag{75}$$

and when $d > e^{-1}$,

$$\liminf_n (1/n)[S_n(k,\omega) - \sum_{i=1}^n p_i(k)] \ge -2\sqrt{-\ln d} \quad \text{a.e.} \tag{76}$$

Proof. In this case we have

$$p(X_1,\ldots,X_n) \doteq q(X_i)\Pi_{i=2}^n p_i(X_{i-1},X_i); \tag{77}$$

$$\gamma_n(\omega) = [\Pi_{i=1}^n p_i(X_i)]/[q(X_1)\Pi_{i=2}^n p_i(X_{i-1},X_i)]. \tag{78}$$

Moreover, for arbitrary sequence of positive numbers $\{a_n, n \ge 1\}$ we have

$$\liminf_n \sqrt[n]{a_n} \ge \liminf_n a_n/a_{n-1}. \tag{79}$$

We have by (77)-(79) and (74),

$$\begin{aligned} \liminf_n [\gamma_n(\omega)]^{1/n} &> \liminf_n [\gamma_n(\omega)/\gamma_{n-1}(\omega)] \\ &= \liminf_n p(X_n)/p_n(X_{n-1},X_n) \ge d \end{aligned} \tag{80}$$

implying that

$$\liminf_n (1/n)\ln \gamma_n(\omega) \ge \ln d. \tag{81}$$

Thus, if put $y_1 = 1$ and $f_i(x_1,\ldots,x_{i-1}) \equiv 1$ $(i \ge 2)$ in (4), then $Y_n \equiv 1$, $D(-\ln d) = [0,1)$. And consequently, (75) and (76) can be established by Theore 1.

Remark. By putting $y_1 = 1$, $f_i(x_1,\ldots,x_{i-1}) \equiv 1$, $i \ge 2$, (27) implies that

$$\limsup_n (1/n)\ln \gamma_n(\omega) \le 0. \tag{82}$$

This together with (81) implies $d \le 1$.

Corollary 1. *If the condition* (74) *of Theorem* 3 *is replaced by the following one:*

$$d_n = \inf\{p_n(j)/p_n(i,j), i,j \in S\}; \tag{83}$$

$$d = \liminf_n d_n > 0, \tag{84}$$

then

$$\limsup_n (1/n)[S_n(k,\omega) - \sum_{i=1}^n p_i(k)] \le 2\sqrt{-\ln d} - \ln d \quad \text{a.e.}; \tag{85}$$

and when $d > e^{-1}$,

$$\liminf_n (1/n)[S_n(k,\omega) - \sum_{i=1}^n p_i(k)] \ge -2\sqrt{-\ln d} \quad \text{a.e.} \tag{86}$$

Proof. Obviously (83) and (84) imply (74). Hence (85) and (86) follow from Theorem 3 directly.

Corollary 2. *Under the conditions of Theorem* 3 *or Corollary* 1,*if* $d = 1$, *then*

$$\lim_n (1/n)[S_n(k,\omega) - \sum_{i=1}^n p_i(k) = 0 \quad \text{a.e.} \tag{87}$$

Proof. Letting $d = 1$ in (75), (76), (85) and (86). Corollary 2 follows.

Theorem 4. *Let the distribution* (2) *in the definition of* $\gamma_n(\omega)$ *be the Poisson distribution* $g(\lambda_i, j) = (1/j!)e^{-\lambda_i}\lambda_i^j$, *where* $0 < \lambda_i \le M$. *Then under the conditions of Theorem* 3 *or its Corollary* 1 *we have*

$$\limsup_n (1/n)\sum_{i=1}^n (X_i - \lambda_i) < 2\sqrt{-M\ln d} - \ln d \quad \text{a.e.}; \tag{88}$$

and when $e^{-M} < d \le 1$,

$$\liminf_n (1/n)\sum_{i=1}^n (X_i - \lambda_i) > -2\sqrt{-M\ln d} \quad \text{a.e.} \tag{89}$$

Proof. Imitating the proof of Theorem 2, (88) and (89) can be established by Theorem 2.

Acknowledgment. The authors are thankful to the referee for his valuable suggestions.

References

1. A. Adler and A. Rosalsky, *On the Chow-Robbins "fair" games problems* , Bull, Inst. Math. Acad. Sinica. **17** (1989), 211-227.

2. R. B. Ash, *Real Analysis and Probability*, Academic Press, New York, (1970).

3. P. Billingsley, *Probability and Measure*, Wiley, New York, (1986).

4. Y. S. Chow and H. Robbins, *On sums of independent random variables with infinite moments and "fair" games*, Proc. Nat. Acad, Sci. U. S. A., **47** (1961), 330-335.

5. J. L. Doob, *Stochastic Processes*, Wiley, New York, 1953.

6. W. Feller, *Note on the law of large numbers and "fair" game*, Ann. Math. Statist., **16** (1945), 301-304.

7. W. Feller, *An introduction to Probability Theory and Its Applications*, vol. I, 2nd ed., Wiley, New York, 1957.

8. S. Karlin and H. M. Taylor, *A First Course in Stochastic Processes*, 2nd ed., Academic Press, New York, 1975.

9. A. N. Kolmogorov, *On the logical fundation of probability theory*, Lecture Notes in Mathematics, 1021, 1-5, Springer, New York, 1982.

10. K. H. Lin, T. G. Chen and L. H. Yang, *On the "fair" games problem for the weighted generalized Petersburg games*, Chinese J. Math. **21** (1993), 21-31.

11. W. Liu, *Relative entropy densities and a class of limit theorems of the sequences of m-valued random variables*, Ann. Probab. **18** (1990), 829-839.

12. W. Liu, *Analytic technique to prove Borel strong law of large numbers*, Amer. Math. Monthly **98** (1991), 146-148.

13. W. Liu, *Some strong limit theorems relative to the geometric average of random transition probabilities of arbitrary finite nonhomogeneous Markov chains*, Stat. Probab. Letts., **21** (1994), 77-83.

14. W. Liu and G. Liu, *A class of strong laws for functionals of countable nonhomogeneous Markov chains*, Stat. Probab. Letts., **22** (1995), 87-96.

15. W. Liu and W. Yang, *A limit theorem for the entropy densities of nonhomogeneous Markov information source*, Stat. Probab. Letts., **22** (1995), 295-301.

16. G. Liu and W. Liu, *On the strong law of large numbers for functionals of countable nonhomogeneous Markov chains*, Stochastic Process. Appl., **50** (1994), 375-391.

17. W. Liu and W. Yang, *An extension of Shannon-McMillan theorem and Some limit properties for nonhomogeneous Markov chains*, Stochastic Process. Appl., **61** (1996), 129-145.

18. W. Liu and Z. Li, *A strong limit theorem for frequency of m-tuple of the sequence of integer-valued random variables*,Chinese J. Math., **23** (1995), 87-91.

19. W. Liu and Z. Cheng, *The logarithmic likelihood ratio and an analytic approach to prove strong limit theorems for discrete random variables*, Chinese J. Math., **23** (1995), 371-382.

20. W. Liu, *A strong limit theorem for arbitrary countable nonhomogeneous Markov chains*, Chinese J. Math., **24** (1996), 211-222.

21. W. Liu and Z. Wang, *A strong limit theorem on random selection for countable nonhomogenous Markov chains*, Chinese J. Math., **24** (1996), 187-197.

22. M. Loéve, *Porbability Theory I*, 4th ed., Springer, New york, 1977.

23. A. Rosalsky, *On the Lin-Chen-Yang Solution to the "fair" game problem*, Chinese J. Math, **22** (1994), 385-394.

24. A. N. Shiryayev, *Probability*, Springer, New York, 1984.

Department of Mathematics, Hebei University of Technology, Tianjin 300130, China.