

Likelihood Ratio and Strong Limit Theorems for the Discrete Random Variable

Wenhan Li¹, Wei Wang¹, Zhiqiang Liu²

¹College of Mathematics and Physics, Shijiazhuang University of Economics, Shijiazhuang, China

²Science College, Beijing University of Civil Engineering and Architecture, Beijing, China

Email: lwh88@163.com, zhiqliu@163.com

Received June 26, 2012; revised July 26, 2012; accepted August 26, 2012

ABSTRACT

This in virtue of the notion of likelihood ratio and the tool of moment generating function, the limit properties of the sequences of random discrete random variables are studied, and a class of strong deviation theorems which represented by inequalities between random variables and their expectation are obtained. As a result, we obtain some strong deviation theorems for Poisson distribution and binomial distribution.

Keywords: Likelihood Ratio; Strong Limit Theorem; Moment Generating Function

1. Introduction

The analytic technique was proposed by Liu [1] to study the strong deviation theorems represented by inequalities, many comprehensive works may be found in Liu [2] and references therein. In references, Liu [3] discussed the strong deviation theorems for discrete random variables by using the generating function method. Liu [4] obtained some strong limit theorems for a multivariate function sequence of discrete random variables by using the concept of the conditional moment generating function and the concept of measure differentiation on a net. Yang and Yang [5] established a strong limit theorem of the Dubins-Freedman type for arbitrary stochastic sequences. Li, Chen and Zhang [6] and Li, Chen and Feng [7] studied the strong limit theorems of arbitrary dependent continuous random variables by using the analytic technique and the Laplace transform approach, and further extended the strong deviation theorems to the differential entropy for arbitrary dependent continuous information sources in more general settings. Yang [8] further studied the limit properties for Markov chains indexed by a homogeneous tree through the analytic technique.

The purpose of this paper is to establish a kind of strong limit theorems represented by inequalities with random bounds, and to extend the analytic technique proposed by Liu [9]. In the proof, the approach of applying the tool the notions of likelihood ratio and of moment generating function to the study of strong limit theorem of the sequences of random discrete variables is proposed. As a result, we obtain some strong deviation theorems

for Poisson distribution and binomial distribution.

Let (Ω, F, P) be a probability space, and let $\{X_n, n \geq 1\}$ which is defined in (Ω, F, P) be a sequence of random variables taking values in $S = \{1, 2, \dots\}$ with the joint distribution

$$P(X_1 = x_1, \dots, X_n = x_n) = f_n(x_1, \dots, x_n), \quad (1)$$
$$x_k \in S, 1 \leq k \leq n.$$

Let

$$P(X_k = x_k) = p_k(x_k), x_k \in S, k \geq 1 \quad (2)$$

In order to indicate the deviation between $\{X_n, n \geq 1\}$ and a sequence of independent random variables with the joint distribution

$$\pi_n(x_1 \cdots x_n) = \prod_{k=1}^n p_k(x_k) \quad (3)$$

we first introduce the following definition.

Definition 1. Let $\{X_n, n \geq 1\}$ be a sequence of random variables with the joint distribution (1), and $\pi_n(x_1, \dots, x_n)$ be defined by (3), let

$$Z_n(\omega) = \frac{f_n(X_1, X_2, \dots, X_n)}{\pi_n(X_1, X_2, \dots, X_n)} \quad (4)$$
$$r_n(\omega) = \log Z_n(\omega).$$

The random variable

$$r(\omega) = \limsup_{n \rightarrow \infty} \frac{1}{n} r_n(\omega) \quad (5)$$

is called the limit logarithmic likelihood ratio, relative to the product of marginal distribution of (3), of $X_n, n \geq 1$,

where \log is the natural logarithm, ω is the sample point, and X_k stands for $X_k(\omega)$.

Although $r(\omega)$ is not a proper metric between probability measures, we nevertheless think of it as a measure of “dissimilarity” between their joint distribution $f_n(x_1, \dots, x_n)$ and the product $\pi_n(x_1, \dots, x_n)$ of their marginals. Obviously, $r(\omega) = 0$ if and only if

$$f_n(x_1, \dots, x_n) = \prod_{k=1}^n p_k(x_k), \quad n \geq 1.$$

and it will be shown in (21) that $r(\omega) \geq 0$, a.s. in any case. Hence, $r(\omega)$ can be used as a random measure of the deviation between the true joint distribution $f_n(x_1, \dots, x_n)$ and the distribution (3) of independent type. In view of the above discussion of the limit logarithmic likelihood ratio, it is natural to think of $r(\omega)$ as a measure how far (the random deviation of) X_n is from being independent, how dependent they are. The smaller $r(\omega)$ is, the smaller the deviation is.

Definition 2. Let $\{X_n, n \geq 1\}$ be a sequence of random variables taking values in S , and $p_k(x_k), k = 1, 2, \dots, n$, be the marginal distribution of $f_n(x_1, \dots, x_n)$, let moment generating function and expectation as follows:

$$M_k(t) = \sum_{i=0}^{\infty} p_k(i) e^{ti}, \quad (6)$$

and

$$EX_k = m_k = \sum_{i=0}^{\infty} i p_k(i), \quad (7)$$

In this paper, we assume that there exists $t_0 > 0$, such that

$$M_k(t) < \infty, \quad t \in [-t_0, t_0], \quad k = 1, 2, \dots, n.$$

2. Main Results

Theorem 1. Let $\{X_n, n \geq 1\}$ be a sequence of random variables taking values in S , $r(\omega), p_k(i), M_k(t), m_k$ be given as above, and $M_k(t)$ is defined in $[-t_0, t_0]$, let

$$D = \{\omega : r(\omega) < \infty\}, \quad P(D) = 1. \quad (8)$$

Then

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n (X_k - m_k) \geq \alpha(r(\omega)) \text{ a.s.} \quad (9)$$

where

$$\alpha(x) = \sup\{\varphi(t, x), -t_0 < t < 0\}, \quad (10)$$

$$\varphi(t, x) = \liminf_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \left[\frac{\ln M_k(t)}{t} - m_k \right] + x/t, \quad (11)$$

$$-t_0 < t < 0,$$

and then

$$\alpha(x) \leq 0; \quad \lim_{x \rightarrow 0^+} \alpha(x) = \alpha(0) = 0. \quad (12)$$

Proof. For arbitrary $t \in [-t_0, t_0]$, let

$$g_k(t, i) = p_k(i) e^{ti} / M_k(t), \quad i \in S, \quad (13)$$

Then

$$\sum_{i \in S} g_k(t, i) = 1. \quad (14)$$

Let

$$\begin{aligned} q_n(t; x_1, \dots, x_n) &= \prod_{k=1}^n g_k(t, x_k) = \prod_{k=1}^n \left[e^{tx_k} p_k(x_k) / M_k(t) \right] \\ &= \exp\left(t \sum_{k=1}^n x_k\right) \cdot \prod_{k=1}^n p_k(x_k) / \prod_{k=1}^n M_k(t). \end{aligned} \quad (15)$$

By (14), it is easy to see that $q_n(t, x_1, \dots, x_n)$ is an n multivariate distribution function, let

$$t_n(t, \omega) = \frac{q_n(t, X_1, \dots, X_n)}{f_n(X_1, \dots, X_n)}. \quad (16)$$

By [10], we have that $t_n(t, \omega)$ is a nonnegative supermartingale by Doob’s martingale convergence theorem, $t_n(t, \omega)$ converges a.s. an integrable random variable T . Hence there exists $A(t) \in F$ such that $P(A(t)) = 1$.

So we have

$$\lim_{n \rightarrow \infty} t_n(t, \omega) < \infty, \quad \omega \in A(t). \quad (17)$$

This implies that

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \ln t_n(t, \omega) \leq 0, \quad \omega \in A(t). \quad (18)$$

By (4), (15) and (18), we obtain

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{1}{n} \left\{ \sum_{k=1}^n \ln M_k(t) + t \sum_{k=1}^n X_k - \ln r_n(\omega) \right\} &\leq 0, \\ \omega &\in A(t), \end{aligned} \quad (19)$$

Let $t = 0$, we obtain

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \ln r_n(\omega) \geq 0, \quad \omega \in A(0) \quad (20)$$

that is

$$r(\omega) \geq 0, \quad \omega \in A(0). \quad (21)$$

Let $-t_0 \leq t < 0$. From (6) and (19), we have

$$\begin{aligned} t \liminf_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n X_k &\leq \limsup_{n \rightarrow \infty} \frac{1}{n} \ln M_k(t) + r(\omega), \\ \omega &\in A(t). \end{aligned} \quad (22)$$

By the property of the inferior limit

$$\begin{aligned} \liminf_{n \rightarrow \infty} (a_n - b_n) &\geq d \\ \Rightarrow \liminf_{n \rightarrow \infty} (a_n - c_n) &\geq \liminf_{n \rightarrow \infty} (b_n - c_n) + d \end{aligned}$$

and dividing the two sides of (22) by t , we obtain

$$\begin{aligned} & \liminf_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n (X_k - m_k) \\ & \geq \liminf_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \left[\frac{\ln M_k(t)}{t} - m_k \right] \\ & + r(\omega)/t, \omega \in A(t) \end{aligned} \tag{23}$$

Let Q^- be the set of rational numbers in the interval $[-t_0, 0)$ and let $A^* = \bigcap_{t \in Q^-} A(t)$, then $P(A^*) = 1$. By (23), we have

$$\begin{aligned} & \liminf_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n (X_k - m_k) \\ & \geq \liminf_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \left[\frac{\ln M_k(t)}{t} - m_k \right] + \frac{r(\omega)}{t}, \end{aligned} \tag{24}$$

$$\omega \in A^*, \forall t \in Q^-$$

By (11), (21) and (24), we have

$$\begin{aligned} & \liminf_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n (X_k - m_k) \geq \varphi(t, r(\omega)), \\ & \omega \in A^*, \forall t \in Q^-. \end{aligned} \tag{25}$$

Because $\varphi(t)$ is a continuous function with respect to t on the interval $[-t_0, 0]$ and $\lim_{x \rightarrow \infty} \varphi(t, x) = -\infty$. By (11) and (21), for every $\omega \in A^* \cap A(0)$, take $t_n(\omega) \in Q^-$, such that

$$\lim_{n \rightarrow \infty} \varphi[t_n(\omega), r(\omega)] = \alpha(r(\omega)). \tag{26}$$

By (26), we have

$$\begin{aligned} & \liminf_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n (X_k - m_k) \geq \varphi[t_n(\omega)], \\ & \omega \in A^* \cap A(0), \end{aligned} \tag{27}$$

By (26) and (27), we have

$$\begin{aligned} & \liminf_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n (X_k - m_k) \geq \alpha(r(\omega)), \\ & \omega \in A^* \cap A(0). \end{aligned} \tag{28}$$

Since $P(A^* \cap A(0)) = 1$, (9) holds by (28). By Jensen Inequality, we obtain

$$M_k(t) = E(e^{tX_k}) \geq e^{tE(X_k)} = e^{tm_k}. \tag{29}$$

By (11) and (29), we have

$$\varphi(t, x) \leq \frac{x}{t} \leq 0, t_0 < t < 0. \tag{30}$$

then

$$\alpha(x) \leq 0. \tag{31}$$

By L'hospital rule, we have

$$\lim_{t \rightarrow 0^-} \frac{1}{t} \ln M_k(t) = \lim_{t \rightarrow 0^-} \frac{M_k(t)}{M_k(t)} = m_k. \tag{32}$$

By (10) and (32), we have

$$\alpha(0) = 0. \tag{33}$$

By (10) and (11), we have

$$\alpha(x) \geq \varphi(-\sqrt{x}, x) = \frac{[M_k(-\sqrt{x}) + x]}{-\sqrt{x}} - m_k, \tag{34}$$

It is easy to see that

$$\lim_{x \rightarrow 0^+} \varphi(-\sqrt{x}, x) = 0. \tag{35}$$

By (31), (34) and (35), we have

$$\lim_{x \rightarrow 0^+} \alpha(x) = 0. \tag{36}$$

Then (12) follows from (34) and (36).

Theorem 2. Under the conditions of the Theorem 1, then

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n (X_k - m_k) \leq \beta(r(\omega)) \text{ a.s.} \tag{37}$$

where

$$\beta(x) = \inf \{ \varphi(t, x), 0 < t < t_0 \}; \tag{38}$$

$$\varphi(t, x) = \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \left[\frac{\ln M_k(t)}{t} - m_k \right] + x/t, \tag{39}$$

$$0 < t < t_0;$$

and then

$$\beta(x) \geq 0; \lim_{x \rightarrow 0^+} \beta(x) = \beta(0) = 0. \tag{40}$$

The proof of Theorem 2 is similar to that of Theorem 1 and hence is omitted here.

Corollary 1. Let $\{X_n, n \geq 1\}$ are independent random variables, then

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n [X_k - m_k] = 0, \text{ a.s.} \tag{41}$$

It is easy to see that (41) can be obtained by (9) and (37).

Corollary 2. Under the conditions of the Theorem 1 and Theorem 2, if $\{X_n, n \geq 1\}$ are independent and identically binomial distribution random variables with parameters m and p , then

$$\begin{aligned} & \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n (X_k - m_k) \\ & < mpe + \frac{r(\omega) - mp(1+t)}{t}, \end{aligned} \tag{42}$$

$$0 < t < 1 < t_0, \text{ a.s.,}$$

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n (X_k - m_k) > \frac{r(\omega)}{t}, \quad (43)$$

$-t_0 < t < 0$, a.s.

Proof. Here, we only proof (42) and (43) is omitted. The moment generating function of binomial distribution is defined by $M_k(t) = (q + pe^t)^m$. Using inequality $\ln t \leq t - 1 (t > 0)$ and $e^t < et (t < 1)$, we have

$$\begin{aligned} & \frac{\ln M_k(t)}{t} - m_k + \frac{x}{t} \\ &= \frac{m \ln(q + pe^t)}{t} - mp + \frac{x}{t} \leq \frac{mp(e^t - 1) + x}{t} - mp \quad (44) \\ &= \frac{mpe^t + x - mp}{t} - mp < \frac{mpet + x - mp}{t} - mp \end{aligned}$$

Hence (44) follows directly from (46).

Corollary 3. Under the conditions of the Theorem 1 and Theorem 2, if $\{X_n, n \geq 1\}$ are independent and identically distribution of Poisson random variables with parameters λ , then

$$\begin{aligned} & \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n (X_k - m_k) \\ & < mpe + \frac{r(\omega) - mp(1+t)}{t} \quad (45) \end{aligned}$$

$0 < t < 1 < t_0$ a.s.,

$$\begin{aligned} & \liminf_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n (X_k - m_k) > \frac{r(\omega)}{t} \quad (46) \\ & -t_0 < t < 0, \quad \text{a.s.} \end{aligned}$$

Proof. The moment generating function of distribution of Poisson is defined by $M_k(t) = e^{\lambda(e^t - 1)}$, Using inequality $e^t < et (t < 1)$, we have

$$\begin{aligned} & \frac{\ln M_k(t)}{t} - m_k + \frac{x}{t} = \frac{\lambda(e^{2t} - 1)}{t} - \lambda + \frac{x}{t} \quad (47) \\ & < e\lambda^2 - \lambda - \frac{\lambda - x}{t} \end{aligned}$$

Hence (45) follows directly from (47). At the same time, by using inequality $e^t - 1 < t (t < 0)$, we have

$$\frac{\ln M_k(t)}{t} - m_k + \frac{x}{t} = \frac{\lambda(e^{2t} - 1)}{t} - \lambda + \frac{x}{t} > \lambda^2 - \lambda + \frac{x}{t}.$$

3. Acknowledgements

The authors would like to thank the National Science Funding of China (Grant No. 41172299) and the Funding Project Education Department Science Project of Hebei Province (Grant No. Z2010297).

REFERENCES

- [1] W. Liu, "Relative Entropy Densities and a Class of Limit Theorems of Sequence of M-Valued Random Variables," *Annals of Probability*, Vol. 18, No. 2, 1990, pp. 829-839. [doi:10.1214/aop/1176990860](https://doi.org/10.1214/aop/1176990860)
- [2] W. Liu, "Strong Deviation Theorems and Analytic Method," Science Press, Beijing, 2003.
- [3] W. Liu, "A Kind of Strong Deviation Theorems for the Sequence of Nonnegative Integer-Valued Random Variables," *Statistics & Probability Letters*, Vol. 32, No. 4, 1997, pp. 269-276.
- [4] W. Liu, "Some Limit Properties of the Multivariate Function Sequences of Discrete Random Variables," *Statistics & Probability Letters*, Vol. 61, No. 1, 2003, pp. 41-50. [doi:10.1016/S0167-7152\(02\)00304-8](https://doi.org/10.1016/S0167-7152(02)00304-8)
- [5] W. G. Yang and X. Yang, "A Note on Strong Limit Theorems for Arbitrary Stochastic Sequences," *Statistics & Probability Letters*, Vol. 78, No. 14, 2008, pp. 2018-2023. [doi:10.1016/j.spl.2008.01.084](https://doi.org/10.1016/j.spl.2008.01.084)
- [6] G. R. Li, S. Chen and J. H. Zhang, "A Class of Random Deviation Theorems and the Approach of Laplace Transform," *Statistics & Probability Letters*, Vol. 79, No. 2, 2009, pp. 202-210. [doi:10.1016/j.spl.2008.07.048](https://doi.org/10.1016/j.spl.2008.07.048)
- [7] G. R. Li, S. Chen and S. Y. Feng, "A Strong Limit Theorem for Functions of Continuous Random Variables and an Extension of the Shannon-McMillan Theorem," *Journal of Applied Mathematics*, Vol. 2008, 2008, pp. 1-10. [doi:10.1155/2008/639145](https://doi.org/10.1155/2008/639145)
- [8] W. G. Yang, "Some Limit Properties for Markov Chains Indexed by a Homogeneous Tree," *Statistics & Probability Letters*, Vol. 65, No. 3, 2003, pp. 241-250. [doi:10.1016/j.spl.2003.04.001](https://doi.org/10.1016/j.spl.2003.04.001)
- [9] W. Liu, "Strong Deviation Theorems and Analytic Method (in Chinese)," Science Press, Beijing, 2003.
- [10] J. L. Doob, "Stochastic Processes," John Wiley & Sons, New York, 1953.