LIETUVOS MATEMATIKOS RINKINYS. LMD DARBAI
51 tomas, 2010, 459–464

ISSN 0132-2818
www.mii.lt/LMR/

Theorems on large deviations for the sum of random number of summands

Aurelija Kasparavičiūtė, Leonas Saulis

Vilnius Gediminas Technical University
Saulėtekio 11, LT-10223 Vilnius
E-mail: aurelija@czv.lt; isaulis@fm.vgtu.lt

Abstract. In this paper, we present the rate of convergence of normal approximation and the theorem on large deviations for a compound process $Z_t = \sum_{i=1}^{N_t} a_i X_i$, where $Z_0 = 0$ and $a_i \geqslant 0$, of weighted independent identically distributed random variables X_i, $i = 1, 2, \ldots$ with mean $\mathbf{E}X_i = \mu$ and variance $\mathbf{D}X_i = \sigma^2 > 0$. It is assumed that N_t is a non-negative integer-valued random variable, which depends on $t \geqslant 0$ and is independent of X_i, $i = 1, 2, \ldots$.

Keywords: cumulant, large deviations, compound Poisson process.

Introduction

Let's $\{X_i,\ i = 1, 2, \ldots\}$ consider a family of independent identically distributed (i.i.d.) random variables (r.v.'s) with mean $\mathbf{E}X_i = \mu$ and variance $\mathbf{D}X_i = \sigma^2 > 0$.

Denote a process

$$Z_t = \sum_{i=1}^{N_t} a_i X_i, \tag{1}$$

where N_t is a non-negative integer-valued r.v. depending on $t \geqslant 0$ and for each t the r.v.'s $N_t, X_1, X_2, \ldots$ are independent. It is assumed that $Z_0 = 0$, besides $a_i \geqslant 0$ and $a = \sup\{a_i,\ i = 1, 2, \ldots\}$. Let's say $\mathbf{E}N_t = \alpha_t$, $\mathbf{D}N_t = \beta_t^2$.

A compound process of a form $Z'_t = \sum_{i=1}^{N_t} X_i$ is a model for processes of insurance companies, where X_i is the i-th claim and N_t denotes the number of claims up to time t. Consequently, Z'_t defines the total claims in the time-interval $[0, t]$ (see [4]). If N_t is a Poisson process, then Z'_t is a compound Poisson process. Analysis of Z'_t is essential not only in insurance, but also in fields of mathematics such as finance mathematics.

Denote

$$T_t = \sum_{i=1}^{N_t} a_i, \qquad \tilde{T}_t = \sum_{i=1}^{N_t} a_i^2. \tag{2}$$

Futhermore, $S_l = \sum_{i=1}^{l} a_i X_i$, $l \geqslant 1$ with mean $\mathbf{E}S_l = \mu \sum_{i=1}^{l} a_i$ and variance $\mathbf{D}S_l = \sigma^2 \sum_{i=1}^{l} a_i^2$. At firts, we obtain

$$\mathbf{E}T_t = \sum_{l=1}^{\infty} \sum_{i=1}^{l} a_i \mathbf{P}(N_t = l), \qquad \mathbf{E}T_t^2 = \sum_{l=1}^{\infty} \left(\sum_{i=1}^{l} a_i \right)^2 \mathbf{P}(N_t = l), \tag{3}$$

$$\mathbf{E}\tilde{T}_t = \sum_{l=1}^{\infty}\sum_{i=1}^{l} a_i^2 \mathbf{P}(N_t = l), \qquad \sum_{l=1}^{\infty} \sum_{\substack{i \neq j \\ i,j=1}}^{l} a_i a_j \mathbf{P}(N_t = l) = \mathbf{E}T_t^2 - \mathbf{E}\tilde{T}_t. \tag{4}$$

With reference, to (3)–(4), we get

$$\mathbf{E}Z_t = \mathbf{E}\sum_{i=1}^{N_t} a_i X_i = \sum_{l=1}^{\infty} \mathbf{E}S_l \mathbf{P}(N_t = l) = \mu \mathbf{E}T_t, \tag{5}$$

$$\mathbf{D}Z_t = \mathbf{E}Z_t^2 - (\mathbf{E}Z_t)^2 = \sigma^2 \mathbf{E}\tilde{T}_t + \mu^2\big(\mathbf{E}T_t^2 - (\mathbf{E}T_t)^2\big) = \sigma^2 \mathbf{E}\tilde{T}_t + \mu^2 \mathbf{D}T_t. \tag{6}$$

Remark 1. If $a_i \equiv 1$, then $\mathbf{E}T_t = \mathbf{E}\tilde{T}_t = \alpha_t$, $\mathbf{D}T_t = \beta_t^2$. And if N_t, $t \geqslant 0$ is a homogeneous Poisson process with intensity $\lambda > 0$, then $\mathbf{E}T_t = \mathbf{E}\tilde{T}_t = \mathbf{D}T_t = \lambda t$, consequently $\mathbf{D}Z_t = (\sigma^2 + \mu^2)\lambda t$.

Denote

$$\tilde{Z}_t = (Z_t - \mathbf{E}Z_t)(\mathbf{D}Z_t)^{-1/2}, \qquad F_{\tilde{Z}_t}(x) = \mathbf{P}(\tilde{Z}_t < x). \tag{7}$$

The characteristic function and the k-th order cumulant of a r.v. are denoted by

$$\varphi_X(u) = \mathbf{E}\exp\{iuX\}, \qquad \Gamma_k(X) = \frac{1}{i^k}\frac{d^k}{du^k}\ln\varphi_X(u)\Big|_{u=0}, \quad k = 1, 2, \ldots,$$

respectively.

To achieve the purpose of this paper, we have to use the cumulants method offered by V.Statulevičius in [8] and developed by R. Rudzkis, L. Saulis, V. Statulevičius in [5].

We assume that i.i.d. random variables X_i, $i = 1, 2, \ldots$ with mean $\mathbf{E}X_i = \mu$ and variance $\mathbf{D}X_i = \sigma^2 > 0$ satisfy the condition (B_γ), if exist constants $\gamma \geqslant 0$ and $K > 0$ such that

$$\big|\mathbf{E}X_i^k\big| \leqslant (k!)^{1+\gamma} K^{k-2} \mathbf{E}X_1^2, \quad k = 3, 4, \ldots. \tag{B_γ}$$

Also we assume, that the r.v. T_t defined by (2) satisfies the condition: exist constants $K_2 > 0$ and $0 \leqslant p < \frac{1}{2}$ such that

$$\big|\Gamma_m(T_t)\big| \leqslant (1/2)m! K_2^{m-2} (\mathbf{D}T_t)^{1+(m-2)p}, \quad m = 2, 3, \ldots. \tag{L}$$

On the grounds of condition (L) the cumulants $\Gamma_m(T_t/\sqrt{\mathbf{D}T_t}) \to 0$ if $\sqrt{\mathbf{D}T_t} \to \infty$. We estimated

$$\big|\Gamma_m(\tilde{T}_t)\big| \leqslant a^m \big|\Gamma_m(T_t)\big| \leqslant a^{2m}\big|\Gamma_m(N_t)\big|, \quad m = 1, 2, \ldots,$$

where $a = \sup\{a_i,\ i = 1, 2, \ldots\}$, $a_i \geqslant 0$. Besides, in the case of a homogeneous Poisson process with intensity $\lambda > 0$, $|\Gamma_m(\tilde{T}_t)| \leqslant a^{2m}\lambda t$, $m = 1, 2, \ldots$. And if $a_i \equiv 1$, then $\Gamma_m(\tilde{T}_t) = \Gamma_m(N_t) = \lambda t$, $m = 1, 2, \ldots$. So condition (L) holds with $p = 0$ and $K_2 = 1$.

In this paper, we have evaluated the accuracy of approximation of the distribution function $F_{\tilde{Z}_t}(x)$ by the Normal Law $\Phi(x) \sim N(0, 1)$, where $N(0, 1)$ is the standard Normal distribution. Moreover, a large deviation theorem was proofed and exponential inequalities were obtained. Note that there is a large amount of literature on normal approximation for random sums of independent sequences of r.v.'s see, for instance, [1, 2, 3]. If we are interested in the rate of convergence, then we have to estimate $\Gamma_k(\tilde{Z}_t)$, to $t \geqslant 0$, $k \geqslant 3$. Such type of research in the case of the sum $Z_t' = \sum_{i=1}^{N_t} X_i$, can be found in the paper [6].

1 The estimate of the k-th order cumulants

The estimate of the k-th order cumulants $\Gamma_k(\tilde{Z}_t)$ of the sum $\tilde{Z}_t$ which is defined by (7), is developed by using the method showed in [6].

Lemma 1. *If* B_γ *is satisfied, then the* k*-th order cumulant of the process* $\tilde{Z}_t$ *defined by* (7) *is*

$$\left|\Gamma_k(\tilde{Z}_t)\right| \leqslant \frac{(3/2)k!}{\Delta_t^{k-2}}, \quad k = 3, 4, \ldots, \tag{8}$$

where

$$\Delta_t = L_t\sqrt{\mathbf{D}Z_t}, \qquad L_t = \left(3K_2|\mu|(\mathbf{D}T_t)^p \vee \left(1 \vee \sigma|\mu|^{-1}\right)4aM\right)^{-1}, \tag{9}$$

where $M = 2(\sigma \vee K)$*,* $\mu \neq 0$*,* $K > 0$*,* $K_2 > 0$*,* $a = \sup\{a_i,\ i = 1, 2, \ldots\}$*, and* $\mathbf{D}Z_t$ *is defined by* (6)*.*

Proof. At first we obtain the characteristic function of the process Z_t which is defined by (1). Since i.i.d. random variables X_i, $i = 1, 2, \ldots$ and N_t are independent, then the characteristic function of Z_t is

$$\begin{aligned}
\varphi_{Z_t}(u) = \mathbf{E}e^{iuZ_t} &= \sum_{l=1}^{\infty}\mathbf{E}e^{uS_l}\mathbf{P}(N_t = l) = \sum_{l=1}^{\infty}\varphi_{S_l}(u)\mathbf{P}(N_t = l) \\
&= \sum_{l=1}^{\infty}\exp\left\{\ln\varphi_{S_l}(u)\right\}\mathbf{P}(N_t = l) \\
&= \sum_{l=1}^{\infty}\exp\left\{\sum_{k=1}^{\infty}(1/k!)\Gamma_k(S_l)(iu)^k\right\}\mathbf{P}(N_t = l) \\
&= \sum_{l=1}^{\infty}\exp\left\{\sum_{k=1}^{\infty}(1/k!)\Gamma_k\left(\sum_{i=1}^{l}a_iX_i\right)(iu)^k\right\}\mathbf{P}(N_t = l) \\
&= \sum_{l=1}^{\infty}\exp\left\{\sum_{k=1}^{\infty}\frac{1}{k!}\sum_{i=1}^{l}a_i^k\Gamma_k(X_1)(iu)^k\right\}\mathbf{P}(N_t = l) \\
&= \sum_{l=1}^{\infty}\exp\left\{\sum_{k=1}^{\infty}\frac{1}{k!}\sum_{i=1}^{l}a_i^2a_i^{k-2}\Gamma_k(X_1)(iu)^k\right\}\mathbf{P}(N_t = l) \\
&= \sum_{l=1}^{\infty}\exp\left\{\theta a^{-2}\left|\sum_{k=1}^{\infty}\frac{1}{k!}\Gamma_k(X_1)(iau)^k\right|\sum_{i=1}^{l}a_i^2\right\}\mathbf{P}(N_t = l) \\
&= \varphi_{\tilde{T}_t}\left(\theta a^{-2}\left|\ln\varphi_{X_1}(au)\right|\right),
\end{aligned} \tag{10}$$

where $a = \sup\{a_i,\ i = 1, 2, \ldots\}$ and θ is a quantity such that $|\theta| < 1$.

Let $\bar{X}_1 = X_1 - \mathbf{E}X_1 = X_1 - \mu$ and $\bar{Z}_t = Z_t - \mathbf{E}Z_t = Z_t - \mu\mathbf{E}T_t$. Using condition (3.1) in [7] with $\gamma = 0$, we get

$$\ln\varphi_{\bar{X}_1}(au) = \sum_{k=1}^{\infty}\frac{\Gamma_k(\bar{X}_1)}{k!}(iau)^k = (3/2)\theta(\sigma au)^2, \quad |u| \leqslant (2aM)^{-1}. \tag{11}$$

We assume that $\mu \neq 0$. Using condition (L) and inequalities (10), (11), we evaluate

$$\begin{aligned}\ln\varphi_{\bar{Z}_t}(u) &= -\mu\mathbf{E}T_t iu + \ln\varphi_{\tilde{T}_t}\big(\theta a^{-2}\big(|a\mu u| + \big|\ln\varphi_{\bar{X}_1}(au)\big|\big)\big)\\ &= -\mu\mathbf{E}T_t iu + \sum_{m=1}^{\infty}\frac{\mathit{\Gamma}_m(\tilde{T}_t)}{m!}\big(\theta a^{-2}\big(|a\mu u| + \big|\ln\varphi_{\bar{X}_1}(au)\big|\big)\big)^m\\ &= -\mu\mathbf{E}T_t iu + \mathit{\Gamma}_1(\tilde{T}_t)\theta a^{-2}\big(|a\mu u| + \big|\ln\varphi_{\bar{X}_1}(au)\big|\big)\\ &\quad + (1/2)\theta\mathbf{D}T_t a^2\big(a^{-2}\big(|a\mu u| + \big|\ln\varphi_{\bar{X}_1}(au)\big|\big)\big)^2\\ &\quad \times\sum_{m=2}^{\infty}\big(a^{-1}K_2\big(|a\mu u| + \big|\ln\varphi_{\bar{X}_1}(au)\big|\big)(\mathbf{D}T_t)^p\big)^{m-2}\\ &= \theta a^{-2}\mathbf{E}\tilde{T}_t\big|\ln\varphi_{\bar{X}_1}(au)\big| + \theta\mathbf{D}T_t\big(a^{-1}\big(|a\mu u| + \big|\ln\varphi_{\bar{X}_1}(au)\big|\big)\big)^2,\end{aligned}$$

for $a^{-1}K_2(|a\mu u| + |\ln\varphi_{\bar{X}_1}(au)|)(\mathbf{D}T_t)^p \leqslant 1/2$.

Furthermore,

$$\ln\varphi_{\bar{Z}_t}(u) = (3/2)\theta\big(\sigma^2\mathbf{E}\tilde{T}_t + \mu^2\mathbf{D}T_t\big)u^2 = (3/2)\theta\mathbf{D}Z_t u^2, \quad |u| \leqslant L_t, \tag{12}$$

where L_t is defined by (9). It is obvious that

$$\begin{aligned}\big|\ln\varphi_{\bar{X}_1}(au)\big|\big(a|\mu u|\big)^{-1} \leqslant 1/5, &\qquad |u| \leqslant \big(\big(1\vee\sigma|\mu|^{-1}\big)4aM\big)^{-1},\\ a^{-1}K_2\big(|a\mu u| + \big|\ln\varphi_{\bar{X}_1}(au)\big|\big)(\mathbf{D}T_t)^p \leqslant 1/2, &\quad |u| \leqslant L_t.\end{aligned}$$

Finally, usage of (12) and Cauchy formula give

$$\big|\mathit{\Gamma}_k(\bar{Z}_t)\big| \leqslant (3/2)k!\mathbf{D}Z_t\, L_t^{2-k}, \quad k = 3, 4, \ldots. \tag{13}$$

Whereas $\mathit{\Gamma}_k(\tilde{Z}_t) = (\mathbf{D}Z_t)^{-k/2}\mathit{\Gamma}_k(\bar{Z}_t)$, then from (13) we get the estimate (8).

It's clear that $\mathit{\Delta}_t \geqslant C(\mu, \sigma, a, K, K_2)(\mathbf{D}T_t)^{\frac{1}{2}-p}$. Thus, $0 \leqslant p < \frac{1}{2}$, $\mathit{\Delta}_t \to \infty$, if $\mathbf{D}T_t \to \infty$.

Remark 2. If $a_i \equiv 1$, then the characteristic function of the process Z_t defined by (1) is $\varphi_{Z_t}(u) = \varphi_{N_t}(\ln\varphi_{X_1}(u))$, and if N_t is a homogeneous Poisson process with intensity $\lambda > 0$, then $\varphi_{Z_t}(u) = \exp\{\lambda t(\varphi_{X_1}(u) - 1)\}$.

Remark 3. If $a_i \equiv 1$, then (8), holds with $\mathit{\Delta}_t = L_t\sqrt{\sigma^2\alpha_t + \mu^2\beta_t^2}$ and $L_t = (3K_2|\mu|\beta_t^{2p} \vee (1\vee\sigma|\mu|^{-1})4M)^{-1}$, $\mu \neq 0$, and in the case of homogeneous Poisson process with intensity $\lambda > 0$

$$\big|\mathit{\Gamma}_k(\tilde{Z}_t)\big| \leqslant \frac{(3/2)k!}{\mathit{\Delta}_t^{*k-2}}, \quad k = 3, 4, \ldots, \tag{14}$$

where

$$\mathit{\Delta}_t^* = L_t^*\sqrt{\lambda t\big(\sigma^2 + \mu^2\big)}, \qquad L_t^* = \big(3|\mu| \vee \big(1\vee\sigma|\mu|^{-1}\big)4M\big)^{-1}.$$

From the second remark we have $\mathit{\Gamma}_k(Z_t) = \lambda t\mathbf{E}X_1^k$, $k = 1, 2, \ldots$. According to the condition (B_γ) when $\gamma = 0$, we get $|\mathit{\Gamma}_k(Z_t)| \leqslant k!K^{k-2}\lambda t(\sigma^2 + \mu^2)$, $k = 3, 4, \ldots$. Thus, the estimate (14) with $\mathit{\Delta}_t^* = K^{-1}\sqrt{\lambda t(\sigma^2 + \mu^2)}$ holds for cumulants $\mathit{\Gamma}_k(\tilde{Z}_t)$. Comparing this expression with (14), we see that $K^{-1} < L_t^*$.

2 Large deviations theorems

In use of the estimate (8) of the k-th order cumulant $\Gamma_k(\tilde{Z}_t)$ of the sum $\tilde{Z}_t$ defined by (7), we obtain the results of the distribution function $F_{\tilde{Z}_t}(x)$.

Theorem 1. *Let i.i.d. random variables X_i, $i = 1, 2, \ldots$ with mean $\mathbf{E}X_i = \mu$ and variance $\mathbf{D}X_i = \sigma^2 > 0$ satisfy condition* (B_γ)*, and r.v. T_t defined by* (2) *satisfies condition* (L)*. Non-negative integer-valued r.v. N_t is independent of $X_1, X_2, \ldots$, then*

$$\sup_x \left| F_{\tilde{Z}_t}(x) - \Phi(x) \right| \leqslant \frac{7}{\Delta_t},$$

where Δ_t defined by (9)*.*

Theorem 2. *Let i.i.d. random variables X_i, $i = 1, 2, \ldots$ satisfy the condition* (B_γ)*, and r.v. T_t satisfies condition* (L)*, then for $x \geqslant 0$, $x = o(\mathbf{D}T_t^{(\frac{1}{2}-p)/3})$*

$$\lim_{\mathbf{D}T_t \to \infty} \frac{1 - F_{\tilde{Z}_t}(x)}{1 - \Phi(x)} = 1, \qquad \lim_{\mathbf{D}T_t \to \infty} \frac{F_{\tilde{Z}_t}(-x)}{\Phi(-x)} = 1.$$

Theorem 3. *If random variables X_i and T_t satisfy condition of Theorem* 1*, then*

$$\mathbf{P}(\pm \tilde{Z}_t \geqslant x) \leqslant \begin{cases} \exp\{-x^2/12\}, & 0 \leqslant x \leqslant 3\Delta_t, \\ \exp\{-x\Delta_t/4\}, & x \geqslant 3\Delta_t, \end{cases}$$

where $\tilde{Z}_t$ and Δ_t are defined by (7) *and* (9)*, respectively.*

Proof. The proof of Theorems follows from (8) with Δ_t defined by (9) and Corollary 2.1, Lemmas 2.3 and 2.4, respectively in [7]

References

[1] V.E. Bening and V.Yu. Korolev. On approximations to generalized cox processes. *Probab. and Math. Statis.*, **18**(2):247–270, 1998.

[2] V.Yu. Korolev. O tochnosti normaln'oj aproksimacii dlja raspredelenij summ sluchajnovo chisla nezavisimyh sluchajnyh velichin. *Teor. Veroyatn. Primen.*, **33**(3):577–581, 1988.

[3] V.Yu. Korolev and S.Ya. Shorgin. On the absolute constant in the remainder term estimate in the central limit theorem for poisson random sums. In *Proc. 4th Intern. Petrozovodsk Conference*, pp. 305–308, VSP, Utrecht, 1997. Publisher.

[4] H. Pragarauskas. *Draudos matematika.* TEV, Vilnius, 2007.

[5] R. Rudzkis, L. Saulis and V. Statulevičius. A general lemma of large deviations. *Lithuanian Math. J.*, **18**(2):169–179, 1978.

[6] L. Saulis and D. Deltuvienė. Normal approximation for sum of random number of summands. *Liet. matem. rink.*, **47**:531–536, 2007.

[7] L. Saulis and V. Statulevičius. *Limit Theorems for Large Deviations.* Kluwer Academic Publisher, Dordrecht, Boston, London, 1991.

[8] V. Statulevičius. On large deviations. *Z. Wahrsch. verw. Geb.*, **6**:133–144, 1966.

REZIUMĖ

Atsitiktinio dėmenų skaičiaus sumoms didžiųjų nuokrypių teoremos

A. Kasparavičiūtė, L. Saulis

Šiame darbe, nagrinėjami sudėtinio proceso $Z_t = \sum_{i=1}^{N_t} a_i X_i$, kai $Z_0 = 0$ ir $a_i \geqslant 0$, aproksimacijos tikslumo įvertis ir didžiųjų nuokrypių teorema. Sumuojami nepriklausomi vienodai pasiskirstę atsitiktiniai dydžiai X_i, $i = 1, 2, \ldots$ su svoriniais koeficientais, turintys vidurkį $\mathbf{E}X_i = \mu$ ir dispersiją $\mathbf{D}X_i = \sigma^2 > 0$. Laikoma, kad neneigiamas sveikas reikšmes įgyjantis atsitiktinis dydis N_t, $t \geqslant 0$, nepriklauso nuo a.d. X_i, $i = 1, 2, \ldots$.

Raktiniai žodžiai: kumuliantai, didieji nuokrypiai, sudėtinis Puasono procesas.