Lietuvos matematikos rinkinys ISSN 0132-2818
Proc. of the Lithuanian Mathematical Society, Ser. A Vol. **56**, 2015
DOI: 10.15388/LMR.A.2015.07 pages 36–41

On large deviations for random sums of the squares of weighted Gaussian random variables

Aurelija Kasparavičiūtė, Dovilė Deltuvienė

Faculty of Fundamental Sciences, Vilnius Gediminas Technical University
Saulėtekio 11, LT-10223, Vilnius
E-mail: aurelija@czv.lt, doviled@mail.lt

Abstract. The paper considers normal approximation to the distribution of random sums of the squares of independent weighted Gaussian random variables (r.vs.) taking into consideration large deviations in the Cramér zone.

Keywords: cumulant method, large deviations, Gaussian sequence.

Introduction

Assume that N denotes a non-negative integer-valued random variable (r.v.) with the distribution:

$$\mathbf{P}(N=m)=q_m, \quad 0<q_m<1,\ m\in\mathbb{N}_0,\ \mathbb{N}_0=\{0,1,2,\dots\}. \tag{1}$$

In addition, let $\{X, X_j,\ j=1,2,\dots\}$ be a family of independent standard normal r.vs. with the distribution function

$$\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{y^2}{2}}\,dy, \quad x\in\mathbb{R},$$

where $\mathbb{R}$ is the set of real numbers. Consider weighted random (compound) sum

$$Z_N=\sum_{j=1}^{N}\mu_j X_j^2, \tag{2}$$

where $0<\mu_j<\infty$. Throughout, we assume that N is independent of $\{X, X_j,\ j=1,2,\dots\}$, and for definiteness, we suppose that $Z_0=0$.

To define the mean and the variance of Z_N, we first introduce the following compound r.vs. $T_{N,r}$:

$$T_{N,r}=\sum_{j=1}^{N}\mu_j^r, \quad r\in\mathbb{N}, \tag{3}$$

where $0<\mu_j<\infty$, and $\mathbb{N}=\{1,2,\dots\}$. For definiteness, we assume $T_{0,r}=0$ for any fixed r. Clearly, $T_{N,0}=N$.

It is easy to verify that the probability characteristics of $T_{N,r}$ are expressed through the characteristics of non-random sum $T_{m,r}=\sum_{j=1}^{m}\mu_j^r$, $m\in\mathbb{N}$. For instance, the mean, second moment and variance are as follows

$$\mathbf{E}T_{N,r} = \sum_{m=1}^{\infty} T_{m,r} q_m, \qquad \mathbf{E}T_{N,r}^2 = \sum_{m=1}^{\infty} T_{m,r}^2 q_m, \qquad \mathbf{D}T_{N,r} = \mathbf{E}T_{N,r}^2 - (\mathbf{E}T_{N,r})^2. \quad (4)$$

It's well known, that the sum $\chi_m^2 = \sum_{j=1}^{m} X_j^2$ has a chi-square distribution with m degrees of freedom. In addition, the density and characteristic functions of χ_m^2 are

$$p_{\chi_m^2}(x) = \begin{cases} 2^{-m/2}((\frac{m}{2}-1)!)^{-1} x^{\frac{m}{2}-1} e^{-\frac{1}{2}x}, & x > 0, \\ 0, & x \leqslant 0, \end{cases}$$

$$f_{\chi_m^2}(u) = \mathbf{E}e^{iu\chi_m^2} = (1-2iu)^{-\frac{m}{2}}, \quad u \in \mathbb{R}, \quad (5)$$

where $\Gamma(m) = \int_0^{\infty} x^{m-1} e^{-x}\, dx$ is gamma function. Consequently,

$$\mathbf{E}\chi_m^2 = m, \qquad \mathbf{D}\chi_m^2 = 2m. \quad (6)$$

Application of (4), (6) together with (8) in [2, p. 257] leads to

$$\mathbf{E}Z_N = \mathbf{E}T_{N,1}, \qquad \mathbf{D}Z_N = 2\mathbf{E}T_{N,2} + \mathbf{D}T_{N,1}. \quad (7)$$

In this paper, we are interested in the normal approximation for the distribution of

$$\tilde{Z}_N = (Z_N - \mathbf{E}Z_N)/\sqrt{\mathbf{D}Z_N}, \quad \mathbf{D}Z_N > 0, \quad (8)$$

that takes into consideration large deviations in the Cramér zone in the case where cumulant method (see [6]) is used. In addition, this paper also considers the exponential inequalities for the probabilities $\mathbf{P}(\tilde{Z}_N \geqslant x)$, $\mathbf{P}(\tilde{Z}_N \leqslant -x)$.

Since we are interested not only in the convergence to the normal distribution but also in a more accurate asymptotic analysis of the distribution function $F_{\tilde{Z}_N}(x)$, we must first find the suitable bound for the kth-order cumulants of (8). For that the combinatorial method is used. In order to obtain upper bounds for $\Gamma_k(\tilde{Z}_N)$, we must impose conditions for the kth-order cumulants of the compound r.v. $T_{N,1}$, which is defined by (3). Consequently, we assume that $T_{N,1}$ satisfies the condition (L): there exist constants $K > 0$, $\epsilon \geqslant 0$ such that

$$\left|\Gamma_k(T_{N,1})\right| \leqslant (1/2) k! K^{k-2} (\mathbf{D}T_{N,1})^{1+(k-2)\epsilon}, \quad k = 2, 3, \ldots. \quad (L)$$

Define the abbreviations $(a \vee b) = \max\{a, b\}$, $a, b \in \mathbb{R}$, $0 < \mu = \sup\{\mu_j,\ j = 1, 2, \ldots\} < \infty$, $0 < \bar{\mu} = \inf\{\mu_j,\ j = 1, 2, \ldots\} < \infty$.

Lemma 1. *Suppose that the r.v.* X *is distributed according to the standard normal law and that the r.v.* $T_{N,1}$ *defined by* (3) *satisfies condition* (L). *Then*

$$\left|\Gamma_k(\tilde{Z}_N)\right| \leqslant k!/\Delta_*^{k-2}, \qquad \Delta_* = \sqrt{\mathbf{D}Z_N}/M_*, \quad M_* = 2\big(K(\mathbf{D}T_{N,1})^{\epsilon} \vee 4\mu^2/\bar{\mu}\big), \quad (9)$$

$k = 3, 4, \ldots$. *Here* $\mathbf{D}Z_N$ *is defined by* (7). *In addition, the constants* K, ϵ *are defined by condition* (L), *and* $\mathbf{D}T_{N,1}$ *is defined by* (4).

Since the accurate upper bounds (9) for the kth-order cumulants of the standardized sum $\tilde{Z}_N$ have been derived, to prove theorems of large deviations and exponential inequalities we have to use general lemmas presented in [1, 4], respectively, about exponential inequalities and large deviations for an arbitrary r.v. with zero mean and unit variance.

We will use θ (with or without an index) to denote a value, not always the same, that does not exceed 1 in modulus.

Theorem 1. *Suppose that the r.v. X is distributed according to the standard normal law and that the r.v. $T_{N,1}$ defined by (3) satisfies condition (L). Then in the interval $0 \leqslant x < \Delta_*/24$, the ratios of large deviations*

$$\frac{1-F_{\tilde{Z}_N}(x)}{1-\Phi(x)} = \exp\{L_*(x)\}\bigl(1+24\theta_1 f(x)(x+1)/\Delta_*\bigr),$$
$$\frac{F_{\tilde{Z}_N}(-x)}{\Phi(-x)} = \exp\{L_*(-x)\}\bigl(1+24\theta_2 f(x)(x+1)/\Delta_*\bigr) \tag{10}$$

are valid, where

$$f(x) = \frac{60(1+0,02\Delta_*^2\exp\{-(1-24x/\Delta_*)\sqrt{\Delta_*/26}\,\})}{1-24x/\Delta_*},$$
$$L_*(x) = \sum_{k=3}^{\infty}\tilde{\lambda}_{*,k}x^k + \theta_3\left(\frac{24x}{\Delta_*}\right)^3. \tag{11}$$

The coefficients $\tilde{\lambda}_{*,k}$ (expressed by cumulants of (8)) coincide with the coefficients of the Cramér–Petrov series (see, e.g. [3]) given by the formula $\tilde{\lambda}_{*,k} = -b_{*,k-1}/k$, where the $b_{*,k}$ are determined successively from the equations

$$\sum_{r=1}^{j}\frac{1}{r!}\Gamma_{r+1}(\tilde{Z}_N)\sum_{j_1+\cdots+j_r=j,\, j_i\geqslant 1}\prod_{i=1}^{r}b_{*,j_i} = \begin{cases}1, & j=1,\\ 0, & j=2,3,\ldots.\end{cases}$$

Observe, that for $k=2,3,\ldots$, estimates are valid

$$|\tilde{\lambda}_{*,k}| \leqslant \frac{2}{k}\left(\frac{16}{\Delta_*}\right)^{k-2}, \qquad L_*(x) \leqslant \frac{x^3}{2(x+\Delta_*/3)}, \qquad L_*(-x) \geqslant -\frac{8x^3}{\Delta_*}.$$

Theorem 2. *Under the conditions of Theorem 1, the ratios*

$$\frac{1-F_{\tilde{Z}_N}(x)}{1-\Phi(x)} \to 1, \qquad \frac{F_{\tilde{Z}_N}(-x)}{\Phi(-x)} \to 1 \tag{12}$$

hold for $x \geqslant 0$, $x = o((\mathbf{D}T_{N,1})^{((1/2)-\epsilon)/3})$ if $\mathbf{D}T_{N,1} \to \infty$ when $0 \leqslant \epsilon < 1/2$.

Theorem 3. *Suppose that the r.v. X is distributed according to the standard normal law and that the r.v. $T_{N,1}$ defined by (3) satisfies condition (L). Then for all $x \geqslant 0$,*

$$\mathbf{P}(\pm\tilde{Z}_N \geqslant x) \leqslant \begin{cases}\exp\{-x^2/4\}, & 0 \leqslant x \leqslant \Delta_*,\\ \exp\{-x\Delta_*/4\}, & x \geqslant \Delta_*.\end{cases}$$

Here $\mathbf{P}(\pm\tilde{Z}_N \geqslant x)$ denotes $\mathbf{P}(\tilde{Z}_N \geqslant x)$ or $\mathbf{P}(\tilde{Z}_N \leqslant -x)$.

It should be noted that the sum Z_N defined by (2) is a partial sum in which the deterministic index $n \in \mathbb{N}$ of the partial sum $Z_n = \sum_{j=1}^{n} \mu_j X_j^2$ is replaced by the r.v. N. Let us note, that the paper [5] considers the sum $\zeta_n = \sum_{s,t=1}^{n} a_{s,t} Y_s Y_t$ of a real stationary Gaussian sequence $\{Y_t,\ t = 1,2,\dots\}$ with the mean $\mathbf{E}Y_t = 0$ and the covariance matrix $R = [\mathbf{E}Y_s Y_t]_{s=\overline{1,n}}^{t=\overline{1,n}}$, $\det R \neq 0$. If μ_j, $j = 1,2,\dots$, is a spectrum of eigenvalues of matrix RA obtained in the solution of the nth degree algebraic equation $\det(A - \mu R^{-1}) = 0$, where $A = [a_{s,t}]_{s=\overline{1,n}}^{t=\overline{1,n}}$ is a symmetric matrix, then the distribution of Z_n is the same as that of the r.v. ζ_n. Aforementioned paper is addressed for asymptotic expansions in the large deviation Cramér zone for the distribution and it's density functions of the quadratic form of a stationary Gaussian sequence ζ_n.

Remark 1. Assume N is non-random: $N := n \in \mathbb{N}$. Then $T_{N,r} = T_{n,r} = \sum_{j=1}^{n} \mu_j^r$, $r \in \mathbb{N}$, where $T_{N,r}$ is defined by (3). Thus in accordance with (4), we have $\mathbf{E}T_{N,r} = T_{n,r}$, $\Gamma_k(T_{n,r}) = 0$, $k = 2,3,\dots$. Consequently, taking into account (7), we get $\mathbf{E}Z_n = T_{n,1}$, $\mathbf{D}Z_n = 2T_{n,2}$. Equality (16) in Section 1 yields

$$\left|\Gamma_k(\tilde{Z}_n)\right| \leqslant k!/\tilde{\Delta}_n^{k-2}, \qquad \tilde{\Delta}_n = \sqrt{\mathbf{D}Z_n}/(2\mu), \quad k = 3,4,\dots. \tag{13}$$

The upper estimate (13) coincides with the estimate (1.12) presented in [5, p. 89]. In our considered instance, estimate (1.12) holds with the parameters $\Delta_n := \tilde{\Delta}_n$, $\bar{B}_n^2 := \mathbf{D}Z_n$. Note that $\tilde{\Delta}_n = C\sqrt{T_{n,2}}$, where $C = \sqrt{2}/(2\mu) > 0$. In consideration of the proof of Theorem 2, 2 the ratios (12) are valid for $x \geqslant 0$ such that $x = o(T_{n,2}^{1/6})$, if $T_{n,2} \to \infty$.

1 Proofs of Lemma 1 and Theorems 1, 2, 3.

Proof of Lemma 1. First, note that the kth-order cumulants of $\chi_m^2 = \sum_{j=1}^{m} X_j^2$ are

$$\Gamma_k(\chi_m^2) = \frac{1}{i^k}\frac{d^k}{du^k} \ln f_{\chi_m^2}(u)\bigg|_{u=0} = 2^{k-1}(k-1)!m, \quad k = 1,2,\dots. \tag{14}$$

Aforementioned equality is obtained due to the characteristic function (5) and definition of the kth-order cumulants (see, e.g. (1.31) in [6, p. 8]).

Recall that $T_{m,k} = \sum_{j=1}^{m} \mu_j^k$, $0 < \mu_j < \infty$, and q_m is defined by (1). Since N is independent of the i.i.d. r.vs. $\{X, X_j,\ j = 1,2,\dots\}$, given (5) and (14), we derive that the characteristic function

$$f_{Z_N}(u) = \mathbf{E}e^{iuZ_N} = \sum_{m=0}^{\infty} e^{\sum_{j=1}^{m} \ln f_{\chi_1^2}(\mu_j u)} q_m = \sum_{m=0}^{\infty} e^{\sum_{k=1}^{\infty} \frac{1}{2k} T_{m,k}(2u)^k} q_m \tag{15}$$

of (2) exists if the kth-order cumulants (14) exist. For a detailed calculations see, e.g. [2, p. 258].

Observe that $f_{Z_N}(u)|_{u=0} = 1$ and $(d^m/dy^m)\ln y|_{y=1} = (-1)^{m-1}(m-1)!$, $m = 1,2,\dots$. Thus, according to (15) and Lemma 5.6 in [3], together with definitions of the kth-order moments and cumulants, we can assert that for all $k = 1, 2, \dots,$

$$\Gamma_k(Z_N) = \frac{d^k}{i^k du^k} \ln f_{Z_N}(u)\Big|_{u=0} = k!2^k \sum\nolimits_1^* \frac{(-1)^{m-1}(m-1)!}{m_1! \cdot \dots \cdot m_k!} \cdot \prod_{j=1}^{k} \Biggl(\sum\nolimits_2^* \frac{\mathbf{E}(T_{N,1}^{\eta_1} \cdot \dots \cdot T_{N,j}^{\eta_j})}{2^\eta \eta_1! \cdot \eta_2! \cdot \dots \cdot \eta_j!} \prod_{n=1}^{j} \left(\frac{1}{n}\right)^{\eta_n} \Biggr)^{m_j}. \tag{16}$$

where $\mathbf{E}(T_{N,1}^{\eta_1} \cdot \dots \cdot T_{N,j}^{\eta_j}) = \sum_{s=0}^{\infty} (T_{s,1}^{\eta_1} \cdot \dots \cdot T_{s,j}^{\eta_j}) q_s$, $T_{s,j} = \sum_{r=1}^{s} \mu_r^j$, $j = 1, 2, \dots$. Here summation $\sum_1^*$ is carried out over all non-negative integer solutions $(m_1, m_2, \dots, m_k)$ of the equation $m_1 + 2m_2 + \dots + km_k = k$, $m_1 + m_2 + \dots + m_k = m$, where $0 \leqslant m_1, \dots, m_k \leqslant k$, and $1 \leqslant m \leqslant k$. In addition, $\sum_2^*$ is taken over all non-negative integer solutions $(\eta_1, \dots, \eta_j)$ of the equation $\eta_1 + 2\eta_2 + \dots + j\eta_j = j$, $\eta_1 + \eta_2 + \dots + \eta_j = \eta$, where $0 \leqslant \eta_1, \dots, \eta_j \leqslant j$, and $1 \leqslant \eta \leqslant j$.

Because of the equality

$$k!2^k \sum\nolimits_1^* \frac{(-1)^{m-1}(m-1)!}{m_1! \cdot \dots \cdot m_k!} \prod_{j=1}^{k} \Biggl(\sum\nolimits_2^* \frac{\mathbf{E}T_{N,2}^{\eta}}{2^\eta \eta_1! \cdot \dots \cdot \eta_j!} \prod_{n=1}^{j} \left(\frac{1}{n}\right)^{\eta_n} \Biggr)^{m_j} = k!2^k \sum\nolimits_1^* \frac{\Gamma_m(T_{N,2})}{2^m m_1! \cdot \dots \cdot m_k!} \prod_{j=1}^{k} \left(\frac{1}{j}\right)^{m_j},$$

and inequalities $T_{N,1}^{\eta_1} \leqslant \bar{\mu}^{-\eta_1} T_{N,2}^{\eta_1}$, $T_{N,j}^{\eta_j} \leqslant \mu^{(j-2)\eta_j} T_{N,2}^{\eta_j}$ as $j \geqslant 2$, where $0 < \bar{\mu} = \inf\{\mu_j,\ j = 1, 2, \dots\} < \infty$, $0 < \mu = \sup\{\mu_j,\ j = 1, 2, \dots\} < \infty$, we have that the inequality

$$\bigl|\Gamma_k(Z_N)\bigr| \leqslant k!(2\mu)^k \sum\nolimits_1^* \frac{|\Gamma_m(T_{N,2})|}{m_1! \cdot \dots \cdot m_k!} (2\mu)^{m_1 - 2m} \bar{\mu}^{-m_1}, \quad k = 1, 2, \dots, \tag{17}$$

is valid. Consequently, from (17), (L) together with equality $\sum_3^* (m_1 + \dots + m_{k-1})!/(m_1! \cdot \dots \cdot m_{k-1}!) = 2^{k-1} - 1$, $k \geqslant 2$, and $|\Gamma_m(T_{N,2})| \leqslant \mu^m |\Gamma_m(T_{N,1})|$, $m = 1, 2, \dots,$ follows that

$$\bigl|\Gamma_k(Z_N)\bigr| \leqslant k!(2\mu)^{k-2}\mathbf{E}T_{N,2} + k!(2\mu)^k \mathbf{D}T_{N,1} \sum\nolimits_3^* \frac{\tilde{m}!(K_1(\mathbf{D}T_{N,1})^{\epsilon})^{\tilde{m}-2}(2\mu)^{m_1}}{(m_1! \cdot \dots \cdot m_{k-1}!)2(4\mu)^{\tilde{m}}\bar{\mu}^{m_1}}, \leqslant k!\mathbf{D}Z_N M_*, \tag{18}$$

where $\mathbf{D}Z_N$ and M_* are defined, accordingly, by (7) and (9). Here $\sum_3^*$ is taken over all the non-negative integer solutions $(m_1, m_2, \dots, m_{k-1})$ of the equation $m_1 + 2m_2 + \dots + (k-1)m_{k-1} = k$, $m_1 + m_2 + \dots + m_{k-1} = \tilde{m}$, where $0 \leqslant m_1, \dots, m_{k-1} \leqslant k$, $2 \leqslant \tilde{m} \leqslant k$.

To complete the proof of Lemma 1, it is sufficient to use (18), and then by noting that $\Gamma_k(\tilde{Z}_N) = (\mathbf{D}Z_N)^{-k/2}\Gamma_k(Z_N)$, $k = 2, 3, \dots,$ we arrive at (9). Here $\tilde{Z}_N$ is defined by (8). □

Proof of Theorem 1. Theorem 1 is proved using Lemma 1 and follows directly from the general Lemma 2.3 (Rudzkis, Saulis, Statulevičius, 1978) on large deviations (see, e.g. in [6, p. 18]). Clearly, $\tilde{Z}_N$ satisfies Statulevičius' condition (see condition (S_γ), e.g. in [6, p. 16]) with the parameters, $\gamma = 0$, $\Delta := \Delta_N$. Accordingly, Lemma 2.3 yields the assertion of Theorem 1. $\square$

Proof of Theorem 2. The proof of Theorem 2 follows immediately if we use the definition of $L_*(x)$ by relation (11). We shall prove that $L_*(x) \to 0$ and $x/\Delta_* \to 0$ as $\Delta_* \to \infty$, where Δ_* is defined by (9). It follows that

$$\Delta_* \geqslant C_1(\mathbf{D}T_{N,1})^{(1/2)-\epsilon} \quad \text{or} \quad \Delta_* \geqslant C_2(\mathbf{D}T_{N,1})^{1/2},$$

accordingly, if $M_* \leqslant 2K(\mathbf{D}T_{N,1})^{\epsilon}$ or $M_* \leqslant 8\mu^2/\bar{\mu}$. Here $C_1 = (2K)^{-1} > 0$, $C_2 = \bar{\mu}/(8\mu^2) > 0$, $0 < \mu = \sup\{\mu_j,\ j = 1, 2, \ldots\} < \infty$, $0 < \bar{\mu} = \inf\{\mu_j,\ j = 1, 2, \ldots\} < \infty$, and M_* is defined by (9). Thus $\Delta_* \to \infty$ as $\mathbf{D}T_{N,1} \to \infty$ when $0 \leqslant \epsilon < 1/2$.
Further, taking into account estimate (9), we obtain that

$$\lambda_{*,3}x^3 = \Gamma_3(\tilde{Z}_N)x^3/6 = o(1), \qquad x/\Delta_* = o((\mathbf{D}T_{N,1})^{-2((1/2)-\epsilon)/3}),$$

for all $x = o((\mathbf{D}T_{N,1})^{((1/2)-\epsilon)/3})$ with $0 \leqslant \epsilon < 1/2$, as $\mathbf{D}T_{N,1} \to \infty$. Thus, $L_*(x) \to 0$ as $\mathbf{D}T_{N,1} \to \infty$ when $0 \leqslant \epsilon < 1/2$. $\square$

Proof of Theorem 3. The proof of Theorem 3 is obtained by virtue of general Lemma 2.4 (Bentkus, Rudzkis, 1980) in [6, p. 19], where the inequality (2.12) holds with $H = 2$, $\Delta := \Delta_N$, $\gamma = 0$. $\square$

References

[1] R. Bentkus and R. Rudzkis. On exponential estimates of the distribution of random variables. *Lith. Math. J.*, **20**:15–30, 1980.

[2] A. Kasparavičiūtė and L. Saulis. Theorems on large deviations for randomly indexed sum of weighted r.vs. *Acta Appl. Math.*, **116**(3):255–267, 2011.

[3] V.V. Petrov. *Limit Theorems of Probability Theory.* Springer-Verlag, New York, 1955. Translated by A.Brown.

[4] R. Rudzkis, L. Saulis and V. Statulevičius. A general lemma on probabilities of large deviations. *Lith. Math. J.*, **18**(2):99–116, 1978.

[5] L. Saulis. Asymptotic expansions for the distribution and density functions of the quadratic form of a stationary gaussian process in the large deviation cramer zone. *Nonlinear Anal. Model. Control*, **6**(2):87–101, 2001.

[6] L. Saulis and V. Statulevičius. *Limit Theorems for Large Deviations.* Kluwer Academic Publishers, London, 1991.

REZIUMĖ

Atsitiktinio dėmenų skaičiaus Gauso atsitiktinių dydžių kvadratų sumos didžiųjų nuokrypių teoremos

A. Kasparavičiūtė, D. Deltuvienė

Šiame darbe yra nagrinėjama atsitiktinio dėmenų skaičiaus nepriklausomų Gauso atsitiktinių dydžių kvadratų su svoriniais koeficientais sumos pasiskirstymo funkcijos aproksimacija normaliuju dėsniu, didžiųjų nuokrypių Kramero zonoje.

Raktiniai žodžiai: kumuliantų metodas, didieji nuokrypiai, Gauso sekos.