On large deviations for sums of independent random variables

Valentin V. Petrov[12] and John Robinson[1*]

[1]School of Mathematics and Statistics, University of Sydney,
NSW 2006, Australia,

[2]Faculty of Mathematics and Mechanics, St. Petersburg University,
Stary Peterhof, St. Petersburg 198504, Russia

Abstract

This paper considers large deviation results for sums of independent random variables, generalizing the result of Petrov (1968) by using a weaker and more natural condition on bounds of the cumulant generating functions of the sequence of random variables.

MSC 60F10, 60G50, 62E20.

Keywords: Limit theorems; sums of independent random variables; large deviation; Cramér's condition; asymptotic expansions.

1 Introduction

The classical Cramér limit theorem on large deviations has the following formulation: Suppose that $X_1, X_2, \cdots$ is a sequence of independent random variables having a common distribution such that

$$Ee^{hX_1} < \infty \text{ in the interval } -H < h < H \text{ for some } H > 0 \tag{1}$$

*Address correspondence to John Robinson, School of Mathematics and Statistics, University of Sydney, NSW 2006, Australia; E-mail: johnr@maths.usyd.edu.au. This research was supported by ARC Discovery Grant DP0451722.

(the Cramér condition). Suppose $EX_1 = 0$, $EX_1^2 = \sigma^2 > 0$. Put $S_n = \sum_{j=1}^n X_j$, $Z_n = S_n/(\sigma\sqrt{n})$, $F_n(x) = P(Z_n \le x)$. If $x > 1$, $x = o(\sqrt{n})$ as $n \to \infty$, then

$$\frac{1 - F_n(x)}{1 - \Phi(x)} = \exp\Big\{\frac{x^3}{\sqrt{n}}\lambda\Big(\frac{x}{\sqrt{n}}\Big)\Big\}\Big[1 + O\Big(\frac{x}{\sqrt{n}}\Big)\Big] \tag{2}$$

$$\frac{F_n(-x)}{\Phi(-x)} = \exp\Big\{-\frac{x^3}{\sqrt{n}}\lambda\Big(-\frac{x}{\sqrt{n}}\Big)\Big\}\Big[1 + O\Big(\frac{x}{\sqrt{n}}\Big)\Big] \tag{3}$$

where $\Phi(x)$ is the standard normal distribution function, $\lambda(t) = \sum_{k=0}^\infty a_k t^k$ is a power series with coefficients depending only on the cumulants of X_1, which converges for all sufficiently small $|t|$. This is a strengthening of Cramér (1938), first given in Petrov (1954) together with a generalization to the case of non-identically distributed random variables. A detailed proof of Cramér's theorem can be found in Petrov (1995).

Cramér's theorem was extended by Feller (1943) to sequences of not necessarily identically distributed random variables under restrictive conditions (Feller considered only random variables taking values in finite intervals), thus Cramér's theorem does not follow from Feller's theorem.

The following result from Petrov (1954) (see also Petrov (1961) for some minor improvement of formulation) is a generalization of Cramér's theorem.

Theorem 1 *Suppose that $X_1, X_2, \cdots$ is a sequence of independent random variables with zero means satisfying the following condition: there exist positive numbers H, G and g such that*

$$g \le |Ee^{hX_n}| \le G \text{ in the circle } |h| < H,\ n = 1, 2, \cdots. \tag{4}$$

Also suppose that

$$\liminf B_n/n > 0, \tag{5}$$

where $B_n = \sum_{j=1}^n EX_j^2$. Put $S_n = \sum_{j=1}^n EX_j$, $Z_n = S_n/\sqrt{B_n}$ and $F_n(x) = P(Z_n \le x)$. If $x > 1$, $x = o(\sqrt{n})$ as $n \to \infty$, then

$$\frac{1 - F_n(x)}{1 - \Phi(x)} = \exp\Big\{\frac{x^3}{\sqrt{n}}\lambda_n\Big(\frac{x}{\sqrt{n}}\Big)\Big\}\Big[1 + O\Big(\frac{x}{\sqrt{n}}\Big)\Big], \tag{6}$$

$$\frac{F_n(-x)}{\Phi(-x)} = \exp\Big\{-\frac{x^3}{\sqrt{n}}\lambda_n\Big(-\frac{x}{\sqrt{n}}\Big)\Big\}\Big[1 + O\Big(\frac{x}{\sqrt{n}}\Big)\Big], \tag{7}$$

where $\lambda_n(t) = \sum_{k=0}^\infty a_{kn} t^k$ is a power series convergent for all sufficiently small t, uniformly in n.

The coefficient a_{kn} of this series is expressed in terms of the cumulants of the random variables $X_1, \cdots, X_n$ of order up to and including $k+3$. In particular, if γ_{kj} is the cumulant of order k of X_j and

$$\Gamma_{kn} = \sum_{j=1}^{n} \gamma_{kj}/n, \tag{8}$$

then

$$a_{0n} = \frac{\Gamma_{3n}}{6\Gamma_{2n}^{3/2}},\ a_{1n} = \frac{\Gamma_{4n}\Gamma_{2n} - 3\Gamma_{3n}^2}{24\Gamma_{2n}^3},\ a_{2n} = \frac{\Gamma_{5n}\Gamma_{2n}^2 - 10\Gamma_{4n}\Gamma_{3n}\Gamma_{2n} + 15\Gamma_{3n}^3}{120\Gamma_{2n}^{9/2}}.$$

If the variables $X_1, X_2, \cdots$ have identical distributions then $\Gamma_{kn} = \gamma_k$ and $\lambda_n(t)$ coincides with $\lambda(t)$ in Cramér's theorem which has coefficients independent of n. Therefore, Cramér's theorem follows from Theorem 1.

It is possible to replace condition (4) by a weaker condition. Let us formulate a result from Petrov (1968) (see also Theorem 2 of Chapter 8 of Petrov (1975)).

Use the notation of Theorem 1. Assume that there exists a circle centered at $z = 0$ within which the cumulant generating functions $L_j(z) = \log E \exp(zX_j)$ are analytic for all j (here log denotes the principal value of the logarithm, so that $L_j(0) = 0$ for every j). Within this circle $L_j(z)$ can be expanded in a convergent power series

$$L_j(z) = \sum_{k=1}^{\infty} \gamma_{kj} z^k / k!, \tag{9}$$

where γ_{kj} is the cumulant of X_j of order k. We have $\gamma_{1j} = EX_j = 0$ and $\gamma_{2j} = EX_j^2$ for every j.

Theorem 2 *Suppose there exist positive constants $c_1, c_2, \cdots$ such that*

$$|L_j(z)| \leq c_j \text{ in the circle } |z| < H,\ n = 1, 2, \cdots, \tag{10}$$

$$\limsup \sum_{j=1}^{n} c_j^{3/2}/n < \infty. \tag{11}$$

If condition (5) is satisfied and if $x \geq 0$, $x = o(\sqrt{n})$ as $n \to \infty$, then

$$\frac{1 - F_n(x)}{1 - \Phi(x)} = \exp\Big\{\frac{x^3}{\sqrt{n}}\lambda_n\Big(\frac{x}{\sqrt{n}}\Big)\Big\}\Big[1 + O\Big(\frac{x+1}{\sqrt{n}}\Big)\Big], \tag{12}$$

$$\frac{F_n(-x)}{\Phi(-x)} = \exp\Big\{-\frac{x^3}{\sqrt{n}}\lambda_n\Big(-\frac{x}{\sqrt{n}}\Big)\Big\}\Big[1 + O\Big(\frac{x+1}{\sqrt{n}}\Big)\Big], \tag{13}$$

where $\lambda_n(t) = \sum_{k=0}^{\infty} a_{kn} t^k$ is the same power series (convergent for all sufficiently small t, uniformly in n) as in Theorem 1.

We shall prove a generalization of all the above mentioned results.

2 Results

Theorem 3 *Let $X_1, X_2, \cdots$ be a a sequence of independent random variables with zero means satisfying the conditions of Theorem 2 with condition (11) replaced by the condition*

$$\limsup \sum_{j=1}^{n} c_j/n < \infty. \tag{14}$$

Then the assertion of Theorem 2 holds for $x \geq 0$, $x = o(\sqrt{n})$ as $n \to \infty$.

Condition (14) is obviously weaker than (11) so Theorem 3 generalizes Theorem 1 of Petrov (1968). We note that the results of Sections 3 and 4 of Petrov (1968) hold with condition (11) replaced by (14).

3 Proof

In the proof of Theorem 2 given in Petrov (1968), (11) is used only to prove his inequality (30); that is that

$$\sup_x |\bar{F}_n(x) - \Phi(x)| \leq c/\sqrt{n}, \tag{15}$$

for all sufficiently large n, where $\bar{F}_n(x)$ is the distribution function of $\bar{Z}_n$ defined in (19) below. We give an alternative proof of (15) assuming (14) in place of (11). The method used is related to methods used in Saulis and Statulevičius (1991).

We have

$$\gamma_{kj} = [d^k L_j(z)/dz^k]_{z=0}.$$

By the condition (10) and the Cauchy inequality for the derivatives of analytic functions we obtain the inequalities

$$|\gamma_{kj}| \leq k! c_j H^{-k}$$

for every k and j. Therefore

$$|\sum_{j=1}^{n} \gamma_{kj}| \leq k! C_0 n H^{-k}, \tag{16}$$

where C_0 is a positive constant satisfying

$$\sum_{j=1}^{n} c_j \leq C_0 n \tag{17}$$

which follows from (14). Taking into account (5), we obtain

$$b_1 n \leq B_n \leq b_2 n, \tag{18}$$

for all sufficiently large n, where b_1 and b_2 are positive constants.

Let $\bar{X}_1, \bar{X}_2, \cdots$ be a sequence of conjugate random variables with

$$P(\bar{X}_j \leq x) = e^{-L_j(h)} \int_{-\infty}^{x} e^{hy} dP(X_j \leq y), \quad j = 1, 2, \cdots,$$

let $L(h) = \sum_{j=1}^{n} L_j(h)$, $\bar{S}_n = \sum_{j=1}^{n} \bar{X}_j$, $\bar{A}_n = E\bar{S}_n,$, $\bar{B}_n = E(\bar{S}_n - E\bar{S}_n)^2$, and let

$$\bar{Z}_n = (\bar{S}_n - \bar{A}_n)/\sqrt{\bar{B}_n}. \tag{19}$$

Choose h as the unique solution of the equation

$$x = \bar{A}_n/\sqrt{B_n}$$

where

$$\bar{A}_n = \frac{dL(h)}{dh}.$$

The cumulant generating function of $\bar{X}_j$ is $\bar{L}_j(z) = L_j(h+z) - L_j(h)$. Further let

$$\bar{\gamma}_{kj} = [d^k L_j(z)/dz^k]_{z=h}.$$

Then, for $|h| < H/2$, $\bar{L}_j(z)$ is an analytic function for $|z| < H/2$, and using the Cauchy inequality for derivatives of analytic functions we have

$$|\bar{\gamma}_{kj}| \leq k! c_j (H/2)^{-k}. \tag{20}$$

Then letting

$$\bar{\Gamma}_{kn} = \sum_{j=1}^{n} \bar{\gamma}_{kj}/n \tag{21}$$

and using (20) and (21), we have

$$|\bar{\Gamma}_{kn}| \leq k! C (H/2)^{-k} \text{ for } |h| < H/2. \tag{22}$$

Also let $\bar{L}(z) = \sum_{j=1}^{n} \bar{L}_j(z)$.

If $\bar{F}_n(u)$ and $\bar{f}_n(t)$ are the distribution function and the characteristic function, respectively, of $\bar{Z}_n$ then

$$\begin{aligned}
\bar{f}_n(t) &= \int_{-\infty}^{\infty} e^{itu} d\bar{F}_n(u) \\
&= \int_{-\infty}^{\infty} e^{itu} dP(\bar{S}_n \leq u\sqrt{\bar{B}_n} + \bar{A}_n) \\
&= \int_{-\infty}^{\infty} e^{it(y-\bar{A}_n)/\sqrt{\bar{B}_n}} dP(\bar{S}_n \leq y) \\
&= e^{-it\bar{A}_n/\sqrt{\bar{B}_n} - L(h)} \int_{-\infty}^{\infty} e^{(h+it/\sqrt{\bar{B}_n})y} dP(S_n \leq y) \\
&= e^{-it\bar{A}_n/\sqrt{\bar{B}_n} - L(h) + L(h+it/\sqrt{\bar{B}_n})}.
\end{aligned}$$

So, for $h < H/2$ and $|t| < H\sqrt{\bar{B}_n}/2$

$$\begin{aligned}
\log \bar{f}_n(t) &= -it\bar{A}_n/\sqrt{\bar{B}_n} - \bar{L}(it/\sqrt{\bar{B}_n}) \\
&= -\frac{1}{2}t^2 + \frac{1}{6}(it/\sqrt{\bar{B}_n})^3 \Big[\frac{d^3\bar{L}(z)}{dz^3}\Big]_{z=\theta it/\sqrt{\bar{B}_n}},
\end{aligned}$$

where $0 \le |\theta| \le 1$. Now

$$\Big[\frac{d^3\bar{L}(z)}{dz^3}\Big]_{z=\theta it/\sqrt{\bar{B}_n}} = \sum_{k=3}^{\infty} \frac{n\bar{\Gamma}_{kn}}{(k-3)!}(\theta it/\sqrt{\bar{B}_n})^{k-3}.$$

So, noting that as for (18), $\bar{b}_1 n < \bar{B}_n < \bar{b}_2 n$, for all sufficiently large n, where $\bar{b}_1$ and $\bar{b}_2$ are positive constants and using (22) we have

$$\begin{aligned}\Big|\Big[\frac{d^3\bar{L}(z)}{dz^3}\Big]_{z=\theta it/\sqrt{\bar{B}_n}}\Big| &\le \frac{C\bar{B}_n}{(H/2)^3}\sum_{k=3}^{\infty} k(k-1)(k-2)\Big[\frac{|t|/\sqrt{\bar{B}_n}}{H/2}\Big]^{k-3} \\ &= \frac{6C\bar{B}_n}{(H/2)^3}\Big(1-\Big[\frac{|t|/\sqrt{\bar{B}_n}}{H/2}\Big]\Big)^{-4}.\end{aligned}$$

So if $0 \le h < H/2$ and $|t| < \delta H\sqrt{\bar{B}_n}/2$ with $0 < \delta < 1$, then

$$\Big|\Big[\frac{d^3\bar{L}(z)}{dz^3}\Big]_{z=\theta it/\sqrt{\bar{B}_n}}\Big| < \frac{6C\bar{B}_n}{(H/2)^3}(1-\delta)^{-4}.$$

Thus

$$|\log \bar{f}_n(t) + t^2/2| < \frac{|t|^3 C(1-\delta)^{-4}}{(H/2)^3\sqrt{\bar{B}_n}}. \tag{23}$$

Recall the following theorem of Esseen quoted from Petrov (1995):

Theorem 4 *Let $F(x)$ and $G(x)$ be distribution functions with characteristic functions $f(t)$ and $g(t)$. Suppose that $G(x)$ has a bounded derivative on the real line, so that $\sup_x G'(x) \le K$. Then for every $T > 0$ and every $b > 1/(2\pi)$ we have*

$$\sup_x |F(x) - G(x)| \le b\int_{-T}^{T} \frac{|f(t)-g(t)|}{|t|}dt + c(b)\frac{K}{T}, \tag{24}$$

where $c(b)$ is a positive constant depending only on b.

We can take $b = 1/\pi$ and $T = \delta H\sqrt{\bar{B}_n}/2$, and note that for $|t| < T$ we have from (23),

$$|\bar{f}_n(t) - e^{-t^2/2}| < C'|t|^3 e^{-t^2/4}/\sqrt{n}, \tag{25}$$

so (15) holds. The remainder of the proof of the theorem is exactly as in the proof of Theorem 1 in Petrov (1968).

References

Cramér, H. (1938). Sur un nouveau théorème-limite de la théorie des probabilités. *Actual. Sci. Indust. (Paris)*, 736: 5-23.

Feller, W. (1943). Generalization of a probability limit theorem of Cramér. *Trans. Amer. Math. Soc.* 54, 361–372.

Petrov, V. V. (1954). A generalization of Cramér's limit theorem. *Uspehi Mat. Nauk.* 9 (1954) No. 4 (62), 195-202. (In Selected Translations in Mathematical Statistics and Probability, 1966, 6, 1-8, American Mathematical Society, Providence, RI.)

Petrov, V. V. (1961) On large deviations of sums of random variables. (Russian) *Vestnik Leningrad. Univ.* 16 no. 1, 25–37.

Petrov, V. V. (1968). Asymptotic behavior of probabilities of large deviations. *Theory Probab. Appl.* 13: 408-420.

Petrov, V. V. (1975). *Sums of independent random variables.* Springer, New York.

Petrov, V. V. (1995). *Limit theorems of probability theory.* Oxford University Press, New York.

Saulis, L. and Statulevičius, V.A. (1991). *Limit Theorems for Large Deviations.* Kluwer Academic Publishers, Dordrecht, The Netherlands.