Self-normalized Cramér-Type Large Deviations for Independent Random Variables

Qi-Man Shao

National University of Singapore
and
University of Oregon

qmshao@darkwing.uoregon.edu

1. Introduction

Let $X, X_1, X_2, \cdots, X_n$ be i.i.d. random variables and let

$$S_n = \sum_{i=1}^{n} X_i, \quad V_n^2 = \sum_{i=1}^{n} X_i^2.$$

- Chernoff's large deviation:

  If $Ee^{t_0 X} < \infty$ for some $t_0 > 0$, then $\forall\ x > EX$,

  $$P\left(\frac{S_n}{n} \geq x\right)^{1/n} \rightarrow \inf_{t \geq 0} e^{-tx} Ee^{tX}.$$

- Self-normalized large deviation (Shao, 1997):

  If $EX = 0$ or $EX^2 = \infty$, then $\forall\ x > 0$

  $$P\left(S_n/V_n \geq xn^{1/2}\right)^{1/n} \rightarrow \lambda(x)$$

  where $\lambda(x) = \sup_{c \geq 0} \inf_{t \geq 0} Ee^{t(cX - x(|X|^2 + c^2)/2)}$

- Cramér's moderate deviation:

  Assume $EX = 0$ and $\sigma^2 = EX^2 < \infty$.

  – If $Ee^{t_0|X|^{1/2}} < \infty$ for $t_0 > 0$, then

$$\frac{P\left(\frac{S_n}{\sigma\sqrt{n}} \geq x\right)}{1 - \Phi(x)} \to 1$$

  uniformly in $0 \leq x \leq o(n^{1/6})$.

  – If $Ee^{t_0|X|} < \infty$ for $t_0 > 0$, then for $x \geq 0$ and $x = o(n^{1/2})$

$$\frac{P\left(\frac{S_n}{\sigma\sqrt{n}} \geq x\right)}{1 - \Phi(x)} = \exp\left\{\frac{x^3}{\sqrt{n}}\lambda(\frac{x}{\sqrt{n}})\right\}\left(1 + O(\frac{1+x}{\sqrt{n}})\right),$$

  where $\lambda(t)$ is the Cramér's series. In particular,

$$\ln P\left(\frac{S_n}{\sigma\sqrt{n}} \geq x_n\right) \sim -\frac{1}{2}x_n^2$$

  for $x_n \to \infty$ and $x_n = o(\sqrt{n})$.

- Self-normalized moderate deviations (Shao, 1997, 1999):

  – If $EX = 0$ and $EX^2 I\{|X| \le x\}$ is slowly varying, then
  $$\ln P(S_n/V_n > x_n) \sim -x_n^2/2$$
  for $x_n \to \infty$ and $x_n = o(\sqrt{n})$.

  – If $EX = 0$ and $E|X|^3 < \infty$, then
  $$\frac{P(S_n/V_n \ge x)}{1 - \Phi(x)} \to 1$$
  uniformly in $0 \le x \le o(n^{1/6})$.

What if $\{X_n, n \geq 1\}$ are independent random variables?

2. Self-normalized Cramér type large deviation for independent random variables

Let $X_1, X_2, \cdots, X_n$ be independent random variables with $EX_i = 0$ and $EX_i^2 < \infty$. Put

$$S_n = \sum_{i=1}^n X_i, \quad V_n^2 = \sum_{i=1}^n X_i^2,$$

$$B_n^2 = \sum_{i=1}^n EX_i^2, \quad L_{n,p} = \sum_{i=1}^n E|X_i|^p.$$

- Petrov (1968):

  Suppose there exist positive constants t_0, c_1, $c_2, \cdots$ such that

  $$|\ln Ee^{tX_i}| \le c_i \text{ for } |t| \le t_0,$$

  $$\limsup_{n\to\infty} \frac{1}{n}\sum_{i=1}^n |c_i|^{3/2} < \infty,$$

  $$\liminf_{n\to\infty} B_n/\sqrt{n} > 0.$$

  If $x \ge 0$ and $x = o(\sqrt{n})$, then

  $$\frac{P(S_n/B_n \ge x)}{1-\Phi(x)} = \exp\left\{\frac{x^3}{\sqrt{n}}\lambda_n(\frac{x}{\sqrt{n}})\right\}\Big(1 + O(\frac{1+x}{\sqrt{n}})\Big)$$

  where λ_n is a power series.

- Wang and Jing (1999):
  - If X_i is symmetric with $E|X_i|^3 < \infty$, then

$$|P(S_n/V_n \le x) - \Phi(x)| \le A \min\left\{(1+|x|^3)\frac{L_{n,3}}{B_n^3}, 1\right\} e^{-x^2/2}.$$

  - If $X_1, \cdots, X_n$ i.i.d. with $\sigma^2 = EX_1^2$ and $E|X_1|^{10/3} < \infty$, then there exists an absolute constant $0 < \eta < 1$ such that

$$|P(S_n/V_n \le x) - \Phi(x)| \le \frac{AE|X_1|^{10/3}}{\sigma^{10/3}\sqrt{n}} e^{-\eta x^2/2}$$

- Chistyakov and Götze (1999):

  If $X_1, X_2, \cdots$ are symmetric with finite third moments, then

$$P(S_n/V_n \geq x) = (1 - \Phi(x))\Big(1 + O(1)(1+x)^3 B_n^{-3} L_{n,3}\Big)$$

  for $0 \leq x \leq B_n/L_{n,3}^{1/3}$, where $O(1)$ is bounded by an absolute constant.

Can the assumption of symmetry be removed?

Let

$$\Delta_{n,x} = \frac{(1+x)^2}{B_n^2} \sum_{i=1}^n EX_i^2 I_{\{|X_i|>B_n/(1+x)\}} + \frac{(1+x)^3}{B_n^3} \sum_{i=1}^n E|X_i|^3 I_{\{|X_i|\le B_n/(1+x)\}}$$

for $x > 0$.

Theorem 1 *[Jing-Shao-Wang (2003)] There is an absolute constant A such that*

$$\frac{P(S_n \ge xV_n)}{1-\Phi(x)} = e^{O(1)\Delta_{n,x}}$$

for all $x \ge 0$ satisfying

$(H1)$ $x^2 \max_{1\le i\le n} EX_i^2 \le B_n^2$.

$(H2)$ $\Delta_{n,x} \le (1+x)^2/A$,

where $|O(1)| \le A$.

Theorem 1 provides a very general framework. The following results are direct consequences of the above general theorem.

Theorem 2 *Let $\{a_n, n \geq 1\}$ be a sequence of positive numbers. Assume that*

$$a_n^2 \leq B_n^2 / \max_{1 \leq i \leq n} EX_i^2$$

and

$$\forall\, \varepsilon > 0, \quad B_n^{-2} \sum_{i=1}^{n} EX_i^2 I_{\{|X_i| > \varepsilon B_n/(1+a_n)\}} \longrightarrow 0 \text{ as } n \to \infty.$$

Then

$$\frac{\ln P(S_n/V_n \geq x)}{\ln(1 - \Phi(x))} \longrightarrow 1$$

holds uniformly for $x \in (0, a_n)$.

The next corollary is a special case of Theorem 2 and may be of independent interest.

Corollary 1 *Suppose that $B_n \geq c\sqrt{n}$ for some $c > 0$ and that $\{X_i^2, i \geq 1\}$ is uniformly integrable. Then, for any sequence of real numbers x_n satisfying $x_n \to \infty$ and $x_n = o(\sqrt{n})$,*

$$\ln P(S_n/V_n \geq x_n) \sim -x_n^2/2.$$

When the X_i's have a finite pth moment, $2 < p \leq 3$, we obtain Chistyakov and Götze' result without assuming any symmetric condition.

Theorem 3 *Let $2 < p \leq 3$ and set*

$$L_{n,p} = \sum_{i=1}^{n} E|X_i|^p, \ d_{n,p} = B_n / L_{n,p}^{1/p}.$$

Then,

$$\frac{P(S_n/V_n \geq x)}{1 - \Phi(x)} = 1 + O(1)\Big(\frac{1+x}{d_{n,p}}\Big)^p$$

for $0 \leq x \leq d_{n,p}$, where $O(1)$ is bounded by an absolute constant. In particular, if $d_{n,p} \to \infty$ as $n \to \infty$, we have

$$\frac{P(S_n/V_n \geq x)}{1 - \Phi(x)} \to 1$$

uniformly in $0 \leq x \leq o(d_{n,p})$.

By the fact that $1 - \Phi(x) \leq 2e^{-x^2/2}/(1+x)$ for $x \geq 0$, we have the following exponential non-uniform Berry-Esseen bound

$$|P(S_n/V_n \geq x) - (1 - \Phi(x))| \leq A(1+x)^{p-1} e^{-x^2/2} / d_{n,p}^p$$

holds for $0 \leq x \leq d_{n,p}$.

For i.i.d. random variables, Theorem 3 simply reduces to

Corollary 2 *Let $X_1, X_2, \cdots$ be i.i.d. with $EX_i = 0$, $\sigma^2 = EX_1^2$, $E|X_1|^p < \infty$ $(2 < p \le 3)$. Then, there exists an absolute constant A such that*

$$\frac{P(S_n/V_n \ge x)}{1 - \Phi(x)} = 1 + O(1)\Big(\frac{(1+x)^p E|X_1|^p}{n^{(p-2)/2}\sigma^p}\Big)$$

for $0 \le x \le n^{1/2-1/p}\sigma/(E|X_1|^p)^{1/p}$, where $|O(1)| \le A$.

Question: Can condition $(H1)$ be removed?

- Shao (2003): Theorem 2 remains valid under $(H2)$. That is, if $a_n \to \infty$ and

$$\forall\ \varepsilon > 0, \quad B_n^{-2} \sum_{i=1}^{n} E X_i^2 I_{\{|X_i| > \varepsilon B_n/(1+a_n)\}} \to 0 \ \text{ as } n \to \infty,$$

then

$$\frac{\ln P(S_n/V_n \geq x)}{\ln(1-\Phi(x))} \to 1$$

holds uniformly for $x \in (0, a_n)$.

3. Self-normalized law of the iterated logarithm

Theorem 4 (Shao (2003)) *Let $X_1, X_2, \cdots$ be independent random variables with $EX_i = 0$ and $0 < EX_i^2 < \infty$. Assume that $B_n \to \infty$ and that*

$$\forall\ \varepsilon > 0,\ \ B_n^{-2} \sum_{i=1}^{n} EX_i^2 I_{\{|X_i| > \varepsilon B_n/(\log\log B_n)^{1/2}\}} \longrightarrow 0 \ \ \textit{as} \ \ n \to \infty.$$

Then,

$$\limsup_{n\to\infty} \frac{S_n}{V_n(2\log\log B_n)^{1/2}} = 1 \ \ a.s.$$

4. Self-normalized central limit theorem

- Gine-Götze-Mason (1995):
  Let $\{X_n, n \geq 1\}$ be i.i.d. Then
  $EX = 0$ and $\max_{1 \leq i \leq n} |X_i|/V_n \to 0$ in probability

$$\iff S_n/V_n \xrightarrow{d.} N(0,1)$$

- Egorov (1996):
  If X_i are independent and symmetric, then

$$S_n/V_n \xrightarrow{d.} N(0,1) \iff \max_{1 \leq i \leq n} |X_i|/V_n \to 0 \text{ in probability}$$

- Shao and Zhou (2003):

  Let $\{X_n, n \geq 1\}$ be independent. Suppose the following conditions are satisfied

$$\max_{1 \leq i \leq n} |X_i|/V_n \to 0 \text{ in probability},$$

$$\sum_{i=1}^{n} \{E(X_i/V_n)\}^2 \to 0,$$

$$E\Big(\frac{S_n}{\max(V_n, a_n)}\Big) \to 0$$

where a_n satisfies

$$\sum_{i=1}^{n} E\frac{X_i^2}{a_n^2 + X_i^2} = 1.$$

Then

$$S_n/V_n \xrightarrow{d.} N(0,1)$$

5. An application to the Student-t statistic

Let

$$T_n = \frac{S_n}{s\sqrt{n}},$$

where $s^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}(X_i - S_n/n)^2$.

T_n and S_n/V_n are closely related via the following identity:

$$T_n = \frac{S_n}{V_n}\Big(\frac{n-1}{n-(S_n/V_n)^2}\Big)^{1/2}.$$

Hence

$$\{T_n \geq x\} = \Big\{\frac{S_n}{V_n} \geq x\Big(\frac{n}{n+x^2-1}\Big)^{1/2}\Big\}$$

and the results for S_n/V_n remain valid for T_n.

6. The main idea of proof of Theorem 1

It suffices to show that

$$P(S_n \geq xV_n) \geq (1-\Phi(x))e^{-A\Delta_{n,x}}$$

and

$$P(S_n \geq xV_n) \leq (1-\Phi(x))e^{A\Delta_{n,x}}$$

for all $x > 0$ satisfying $(H1)$ and $(H2)$.

Let

$$b := b_x = x/B_n.$$

Observe that, by the Cauchy inequality

$$xV_n \leq (x^2 + b^2V_n^2)/(2b).$$

Thus, we have

$$\begin{aligned} P(S_n \geq xV_n) &\geq P(S_n \geq (x^2+b^2V_n^2)/(2b)) \\ &= P(2bS_n - b^2V_n^2 \geq x^2). \end{aligned}$$

Therefore, the lower bound follows from the following proposition immediately.

Proposition 1 *There exists an absolute constant $A > 1$ such that*

$$P(2bS_n - b^2V_n^2 \geq x^2) = (1-\Phi(x))e^{O(1)\Delta_{n,x}}$$

for all $x > 0$ satisfying $(H1)$ and $(H2)$, where $|O(1)| \leq A$.

As for the upper bound: when $0 < x \le 2$, this bound is a direct consequence of the Berry-Esseen bound. For $x > 2$, let

$$\tau := \tau_{n,x} = B_n/(1+x)$$

and define

$$\bar{X}_i = X_i I_{\{|X_i| \le \tau\}}, \ \ \bar{S}_n = \sum_{i=1}^n \bar{X}_i, \ \ \bar{V}_n^2 = \sum_{i=1}^n \bar{X}_i^2,$$

$$S_n^{(i)} = S_n - X_i, \ \ V_n^{(i)} = (V_n^2 - X_i^2)^{1/2}, \ \ \bar{B}_n^2 = \sum_{i=1}^n E\bar{X}_i^2.$$

Noting that for any $s, t \in R^1$, $c \ge 0$ and $x \ge 1$,

$$x\sqrt{c+t^2} \ge t + \sqrt{(x^2-1)c},$$

we have

$$\{s + t \ge x\sqrt{c+t^2}\} \subset \{s \ge (x^2-1)^{1/2}\sqrt{c}\}.$$

Hence,

$$\begin{aligned}
&P(S_n \ge xV_n) \\
&\le \ P(\bar{S}_n \ge x\bar{V}_n) + P(S_n \ge xV_n, \max_{1\le i\le n} |X_i| > \tau) \\
&\le \ P(\bar{S}_n \ge x\bar{V}_n) + \sum_{i=1}^n P(S_n \ge xV_n, |X_i| > \tau) \\
&\le \ P(\bar{S}_n \ge x\bar{V}_n) + \sum_{i=1}^n P(S_n^{(i)} \ge (x^2-1)^{1/2} V_n^{(i)}, |X_i| > \tau) \\
&\le \ P(\bar{S}_n \ge x\bar{V}_n) + \sum_{i=1}^n P(S_n^{(i)} \ge (x^2-1)^{1/2} V_n^{(i)}) P(|X_i| > \tau).
\end{aligned}$$

By the inequality $(1+y)^{1/2} \geq 1 + y/2 - y^2$ for any $y \geq -1$, we have

$$\begin{aligned}
& P(\bar{S}_n \geq x\bar{V}_n) \\
= \ & P(\bar{S}_n \geq x(\bar{B}_n^2 + \sum_{i=1}^{n}(\bar{X}_i^2 - E\bar{X}_i^2))^{1/2}) \\
\leq \ & P\Big(\bar{S}_n \geq x\bar{B}_n\Big\{1 + \frac{1}{2\bar{B}_n^2}\sum_{i=1}^{n}(\bar{X}_i^2 - E\bar{X}_i^2) \\
& \qquad - \frac{1}{\bar{B}_n^4}(\sum_{i=1}^{n}(\bar{X}_i^2 - E\bar{X}_i^2))^2\Big\}\Big) \\
:= \ & K_n.
\end{aligned}$$

Thus, the upper bound follows from the next two propositions:

Proposition 2 *There is an absolute constant A such that*

$$P(S_n^{(i)} \geq xV_n^{(i)}) \leq (1 + x^{-1})\frac{1}{\sqrt{2\pi}x}\exp(-x^2/2 + A\Delta_{n,x})$$

Proposition 3 *There exists an absolute constant A such that*

$$K_n \leq (1 - \Phi(x))e^{A\Delta_{n,x}}$$

for all $x > 2$ satisfying conditions $(H1)$ and $(H2)$.

[1] Shao, Q.M. (1997). Self-normalized large deviations. *Ann. Probab.* **25**, 285-328.

[2] Shao, Q.M. (1999). A Cramér type large deviation result for Student's t-statistic. *J. Theoret. Probab.* **12**, 387-398.

[3] Jing, B.Y., Shao, Q.M. and Wang, Q.Y. (2003). Self-normalized Cramér type large deviations for independent random variables. *Ann. Probab.* **31** , 2167-2215.

[4] Chernoff, H. (1952). A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Ann. Math. Statist.* **23**, 493–507.

[5] Griffin, P. and Kuelbs, J. (1989). Self-normalized laws of the iterated logarithm. *Ann. Probab.* 17, 1571–1601.

[6] Logan, B. F., Mallows, C. L., Rice, S. O. and Shepp, L. A. (1973). Limit distributions of self-normalized sums. *Ann. Probab.* 1, 788–809.

[7] Wang, Q. Y. and Jing, B.-Y. (1999). An exponential nonuniform Berry-Esseen bound for self-normalized sums. *Ann. Probab.* **27**, 2068-2088.

[8] Chistyakov, G. P. and Götze, F. (1999). Moderate deviations for self-normalized sums. Preprint.

[9] Gine, E., Götze, F. and Mason, D. M. (1997). When is the Student t-statistic asymptotically standard normal? *Ann. Probab.* **25**, 1514-1531.

[10] Hahn, M.G., Mason, D.M, Weiner, D.C. (eds) (1991). *Sums, Trimmed Sumsand Extremes*. Birkhauser, Boston.