Notes on the saddlepoint expansion

Giuseppe Durisi and Alejandro Lancho

I. Preliminaries

Let $Z_1, \dots, Z_n$ be independent and identically distributed, real-valued, zero-mean random variables. Let $m(\tau) = \mathbb{E}\left[e^{\tau Z_1}\right]$ be the moment-generating function of these random variables and $\psi(\tau) = \log m(\tau)$ be the cumulant-generating function. We use the notation m', m'' and m''' to denote the first three derivatives of $m(\tau)$. Similarly, ψ', ψ'', and ψ''' denote the first three derivatives of $\psi(\tau)$.

A random variable Z is said to be lattice if it is supported on the points $b, b \pm h, b \pm 2h, \dots$ for some b and h. A random variable that is not lattice will be referred to as nonlattice. Throughout, we will focus only on nonlattice random variables. Finally, we will denote by $Q(\cdot)$ the Gaussian Q function:

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_{x}^{\infty} \exp\left(-\frac{u^2}{2}\right) \mathrm{d}u.$$

II. The saddlepoint expansion

The goal is to estimate accurately the tail probability

$$\mathbb{P}\left[\frac{1}{n}\sum_{\ell=1}^{n} Z_\ell > \gamma\right]$$

for some $\gamma > 0$, in the regime in which both the central-limit theorem and the large-deviation bound provided by Chernoff inequality provide loose estimates.

The saddlepoint method [1] yields such an accurate estimate. The resulting expansion is given below. A self-contained proof (for a slightly more general setup) can be found in, e.g., [2, App. I.A].

Theorem 1 (saddlepoint approximation): Let the zero-mean i.i.d. random variables $\{Z_\ell\}_{\ell=1}^{n}$ be nonlattice. Suppose that there exists a $\overline{\tau} > 0$ and a $\underline{\tau} < 0$ such that

$$\sup_{\tau \in [\underline{\tau}, \overline{\tau}]} \left|m'''(\tau)\right| < \infty$$

and

$$\inf_{\tau \in [\underline{\tau}, \overline{\tau}]} \psi''(\tau) > 0.$$

Then, if $\gamma \geq 0$ and the solution to the stationary equation $\psi'(\tau) = \gamma$ gives a $\tau \in [0, \overline{\tau}]$, we have

$$\mathbb{P}\left[\sum_{\ell=1}^{n} Z_\ell \geq n\gamma\right] = e^{n[\psi(\tau) - \tau\psi'(\tau)]}\left[\Psi(\tau, n) + \frac{K(\tau, n)}{\sqrt{n}} + o\left(\frac{1}{\sqrt{n}}\right)\right]. \tag{1}$$

where

$$\begin{aligned}
\Psi(\tau, n) &= e^{n\frac{u^2}{2}\psi''(\tau)} Q\left(u\sqrt{n\psi''(\tau)}\right) \\
K(\tau, n) &= \frac{\psi'''(\tau)}{6\psi''(\tau)^{3/2}}\left(-\frac{1}{\sqrt{2\pi}} + \frac{u^2 n\psi''(\tau)}{\sqrt{2\pi}} - u^3\psi''(\tau)^{3/2} n^{3/2}\Psi(\tau, n)\right)
\end{aligned}$$

and $o(1/\sqrt{n})$ comprises terms that vanish faster than $1/\sqrt{n}$ and are uniform in τ, i.e.,

$$\lim_{n\to\infty} \sup_{\tau \in [0, \tau_0)} \frac{o(1/\sqrt{n})}{1/\sqrt{n}} = 0.$$

A. Remarks

Here is the intuition behind the saddlepoint approximation. One performs the usual exponential tilting on the distribution of Z that is needed to prove the achievability of large-deviation exponent $[\psi(\tau) - \tau\psi'(\tau)]$ (see e.g., [3, Ch. 5.11]). This allows one to pull out of the integral the exponential term. The pre-exponential factor is computed by approximating the distribution of the average of the tilted random variables by a Gaussian distribution using the central-limit theorem.

III. Numerical example

Assume that the $\{Z_\ell\}_{\ell=1}^n$ are independent $\mathrm{Gamma}(k,\theta)$-distributed random variables, i.e., their probability density function is given by:

$$f_Z(x) = \frac{1}{\Gamma(k)\theta^k} x^{k-1} e^{-x/\theta}.$$

Then their sum is $\mathrm{Gamma}(nk,\theta)$-distributed. Hence, the tail probability

$$\mathbb{P}\left[\sum_{\ell=1}^n Z_\ell > n\gamma\right]$$

can be easily evaluated numerically. We want to compare the approximation obtained using the saddlepoint method, as well as the normal approximation resulting from the central limit-theorem

$$\mathbb{P}\left[\sum_{\ell=1}^n Z_\ell \geq n\gamma\right] \approx Q\left(\frac{n(\mu-\gamma)}{\sqrt{n\sigma^2}}\right) \tag{2}$$

where $\mu = \mathbb{E}[Z_1]$ and $\sigma_Z^2 = \mathrm{Var}[Z_1]$, and the Chernoff bound

$$\mathbb{P}\left[\sum_{\ell=1}^n Z_\ell \geq n\gamma\right] \leq e^{n[\psi(\tau)-\tau\psi'(\tau)]} \tag{3}$$

where τ is the solution of $\psi'(\tau) = \gamma$.

Throughout, we set $k = 4$, $\theta = 1$, and $n = 100$. It then follows that $\mu = \sigma^2 = 4$. Furthermore, $\psi(\tau) = -4\log(1-\tau)$. Hence, for all $\gamma > 4$, $\psi'(\tau) = \gamma$ if $\tau = 1 - 4/\gamma$.

In the figure we compare the exact tail probability with the normal approximation (2), the Chernoff bound (3) and the saddle-point approximation in Theorem 1, obtained by neglecting the $o(1/\sqrt{n})$ term in (1). As expected, the normal approximation is accurate for values of γ close to the mean, and the Chernoff bound captures the correct slope of decay of the error probability. The saddlepoint expansion is accurate over the entire range of γ values considered in the figure.

References

[1] J. L. Jensen, *Saddlepoint approximations*. Oxford, U.K.: Oxford Univ. Press, 1995.

[2] A. Lancho, J. Ostman, G. Durisi, T. Koch, and G. Vazquez-Vilar, “Saddlepoint Approximations for Short-Packet Wireless Communications,” *IEEE Trans. Wireless Commun.*, 2020.

[3] G. R. Grimmett and D. R. Stirzaker, *Probability and Random Processes*, 3rd ed. Oxford, U.K.: Oxford Univ. Press, 2001.

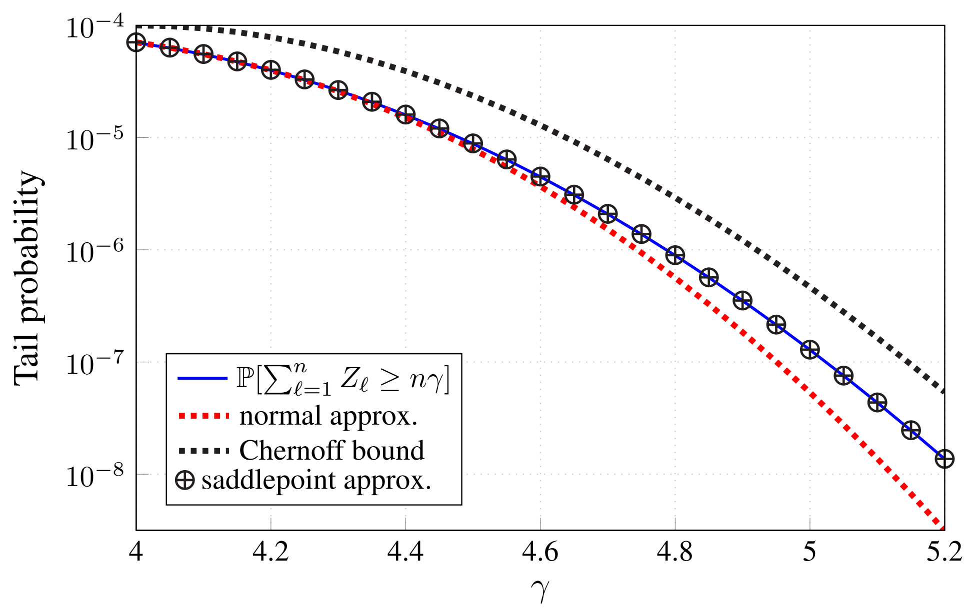

Fig. 1. The probability $\mathbb{P}[\sum_{\ell=1}^{n} Z_\ell \geq n\gamma]$ as a function of γ. The Z_ℓ are independent $\mathrm{Gamma}(4,1)$ random variables. The exact tail proability is compared with the normal approximation (2), the Chernoff bound (3), and the saddle-point approximation in Theorem 1.