On the Maxima of Sub-Gaussian Random Variables

Behrad Moniri
University of Pennsylvania
bemoniri@seas.upenn.edu

Abstract

In this note, we will first prove a bound on the expected maxima of a sequence of weighted sub-gaussian random variables. Next, we show an upper bound for the expected value of the maximum of a finite number of sub-gaussian random variables. Finally, we prove a high probability version of these results.

The following theorem on the expected maxima of an infinite sequence of weighted sub-gaussian random variables is stated in Exercise (2.5.10) of [Ver18].

Theorem 1. *Let $X_1, X_2, \dots$ be a sequence of independent σ-sub Gaussian random variables, then*

$$\mathbb{E}\left[\max_i \frac{|X_i|}{\sqrt{1+\log(i)}}\right] \leq 2\sigma + \frac{\pi^2}{3}\sqrt{2\pi}. \tag{1}$$

Proof. The expected value can be written in terms of an integral of tail probabilities:

$$\mathbb{E}\left[\frac{|X_i|}{\sqrt{1+\log(i)}}\right] = \int_0^\infty \mathbb{P}\Big(\max_i \frac{|X_i|}{\sqrt{1+\log(i)}} \geq t\Big) dt$$

If X is σ-sub gaussian, we have $\mathbb{P}[|X| \geq t] \leq 2e^{-\frac{t^2}{2}}$. Let $a = 2\sigma$, we can divide the integral into two parts and write the following chain of inequalities

$$\begin{aligned}
\mathbb{E}\left[\frac{|X_i|}{\sqrt{1+\log(i)}}\right] &\leq \int_0^{2\sigma} dt + \int_{2\sigma}^\infty \mathbb{P}\left(\max_i \frac{|X_i|}{\sqrt{1+\log(i)}} \geq t\right) dt \\
&\leq 2\sigma + \int_{2\sigma}^\infty \sum_{i=1}^\infty \mathbb{P}\left(\frac{|X_i|}{\sqrt{1+\log(i)}}\right) dt && \text{(union bound)} \\
&\leq 2\sigma + \sum_{i=1}^\infty \int_{2\sigma}^\infty 2\exp\left(\frac{-t^2}{2\sigma^2}\big(1+\log(i)\big)\right) dt && \text{(sub-gaussian tails)} \\
&\leq 2\sigma + 2\sum_{i=1}^\infty \int_2^\infty e^{-u^2/2}.\ i^{-u^2/2}\, du \\
&\leq 2\sigma + 2\sum_{i=1}^\infty \int_2^\infty e^{-u^2/2}.\ i^{-2}\, du \\
&\leq 2\sigma + 2\sum_{i=1}^\infty i^{-2} \int_2^\infty e^{-u^2/2}\, du \\
&\leq 2\sigma + \frac{\pi^2}{3}\sqrt{2\pi},
\end{aligned}$$

which finished the proof. □

We will now prove a similar bound, for the maximum of finite number of sub-gaussian random variables.

Theorem 2. *Let $X_1, \dots, X_n$ be independent σ-sub gaussian random variables. We have*

$$\mathbb{E}\big[\max_{i\in\{1,\dots,n\}} X_i\big] \leq \sigma\sqrt{2\log n}. \tag{2}$$

Proof. Let $Y = \max_{i \in \{1,\dots,n\}} X_i$.

$$\begin{aligned} e^{\lambda \mathbb{E}[\max X_i]} &\le \mathbb{E}\Big[e^{\lambda \max X_i}\Big] \\ &= \mathbb{E}\Big[\max e^{\lambda X_i}\Big] \\ &\le \mathbb{E}\left[\sum_{i=1}^{n} e^{\lambda X_i}\right] \\ &\le n e^{\frac{\lambda^2 \sigma^2}{2}}. \end{aligned}$$

Where the first inequality is a consequence of Jensen's inequality. Hence,

$$\mathbb{E}[\max X_i] \le \frac{\log(n)}{\lambda} + \frac{\lambda \sigma^2}{2}.$$

Optimizing the RHS, yields $\lambda^* = \sqrt{\frac{2 \log n}{\sigma^2}}$. Thus

$$\mathbb{E}[\max X_i] \le \sigma \sqrt{2 \log(n)},$$

which proves the theorem. □

Note 3. *Let $X_1, \dots, X_n \sim \mathcal{N}(0, \sigma^2)$ be i.i.d. random variables. In [Kam], the following upper and lower bounds are proved:*

$$\frac{1}{\sqrt{\pi \log 2}} \sigma \sqrt{\log n} \le \mathbb{E}\big[\max_{i \in \{1,\dots,n\}} X_i\big] \le \sigma \sqrt{2 \log n}. \tag{3}$$

Hence, the bound (2) *is sharp.*

We will now show that the maximum, is less than $\sqrt{2\sigma^2 \log n}$ with high probability.

Theorem 4. *Let $X_1, \dots, X_n$ be independent σ-sub gaussian random variables:*

$$\mathbb{P}\Big(\max X_i - \sqrt{2\sigma^2 \log n} \ge t\Big) \le \exp\left(\frac{-t^2}{2\sigma^2}\right). \tag{4}$$

Proof.

$$\begin{aligned} \mathbb{P}\Big(\max X_i - \sqrt{2\sigma^2 \log n} \ge t\Big) &= \mathbb{P}\Big(\exists i \in [n] : \quad X_i \ge t + \sqrt{2\sigma^2 \log n}\Big) \\ &\le n\, \mathbb{P}\Big(X_i \ge t + \sqrt{2\sigma^2 \log n}\Big) && \text{(union bound)} \\ &\le n \ \exp\left(\frac{-\big(t + \sqrt{2\sigma^2 \log n}\big)^2}{2\sigma^2}\right) && \text{(sub-gaussian tail)} \\ &= \exp\left(\frac{-t^2}{2\sigma^2}\right) \exp\left(\frac{-2t\sqrt{2\sigma^2 \log n}}{2\sigma^2}\right) \le \exp\left(\frac{-t^2}{2\sigma^2}\right). \end{aligned}$$

This proves the theorem. □

References

[Kam] Gautam Kamath. Bounds on the expectation of the maximum of samples from a gaussian. `http://www.gautamkamath.com/writings/gaussian_max.pdf`.

[Ver18] Roman Vershynin. *High-Dimensional Probability: An Introduction with Applications in Data Science*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2018.