# Chapter 13
# Concentration of sub-Gaussian random variables

**Definition 13.1 (Sub-gaussian random variables[1]).** A random variable $m$ is called sub-gaussian if its MGF is dominated[2] by a mean zero normal random variable with variance $\sigma^2$, i.e.,

$$\mathbb{E}\left[e^{\lambda m}\right] \leq e^{\frac{\lambda^2\sigma^2}{2}}. \tag{13.1}$$

$m$ is sometimes called a $\sigma$-sub-gaussian or a sub-gaussian with *proxy* $\sigma^2$. A direct consequence of the definition 1 is that a sub-gaussian random variable has zero mean and its variance is bounded above by $\sigma^2$.

**Proposition 13.1.** *If $m$ is a $\sigma$-sub-gaussian, then $\mathbb{E}[m] = 0$ and $\mathbb{V}ar[m] \leq \sigma^2$.*

*Proof.* For any $\lambda$, using Taylor expansion for both sides of (13.1) we have

*Why the first equality is true?*

$$\sum_{n=0}^{\infty} \frac{\lambda^n}{n!}\mathbb{E}[m^n] = \mathbb{E}\left[e^{\lambda m}\right] \leq e^{\frac{\lambda^2\sigma^2}{2}} = \sum_{n=0}^{\infty} \frac{\lambda^{2n}\sigma^{2n}}{2^n n!}.$$

Now dividing both sides by $\lambda > 0$ and taking the limit $\lambda \to 0$, we conclude that $\mathbb{E}[m] = 0$. Next, dividing both sides by $\lambda^2$ and again taking the limit $\lambda \to 0$ we can conclude that $\mathbb{V}ar[m] \leq \sigma^2$.

The following theorem characterizes sub-gaussian random variables.

**Theorem 13.1 (Sub-gaussian properties).** *Let $m$ be a random variable. Then the following are equivalent:*

*i) There exists a constant $c_1$ such that the tail of $m$ satisfies*

$$\mathbb{P}[|m| \geq t] \leq 2e^{-\frac{t^2}{2c_1^2}}, \quad \forall t \geq 0. \tag{13.2}$$

[1] A weaker definition is based on the tail bound, i.e., $m$ is a sub-gaussian random variable if

$$\mathbb{P}[|m| \geq t] \leq 2e^{-t^2/c^2},$$

for some constant $c$.

[2] That is, the Laplace transform of $m$ is dominated by the Laplace transform of $\mathcal{N}\left(0, \sigma^2\right)$.

*ii) There exists a constant $c_2$ such that the moments of $m$ satisfy*

$$\|m\|_p \stackrel{def}{=} \left(\mathbb{E}\left[|m|^p\right]\right)^{1/p} \leq c_2\sqrt{p}, \quad \forall p \geq 1.$$

*iii)There exists a constant $c_3$ such that the MGF of $m^2$ satisfies*

$$\mathbb{E}\left[e^{\lambda^2 m^2}\right] \leq e^{c_3^2\lambda^2}, \quad \forall |\lambda| \leq \frac{1}{c_3}.$$

*Moreover, if $\mathbb{E}[m] = 0$, then all the above are equivalent to*

*iv) There exists a constant $c_4$ such that the MGF of $m$ satisfies*

$$\mathbb{E}\left[e^{\lambda m}\right] \leq e^{\sigma^2\lambda^2/2}, \quad \forall \lambda \in \mathbb{R}.$$

*Proof.* We prove $1 \Rightarrow 2$, $2 \Rightarrow 3$, $3 \Rightarrow 1$, and $4 \Rightarrow 1$.

- $1 \Rightarrow 2)$: Using Proposition 12.1 we have

$$\mathbb{E}\left[|X|^p\right] = \int_0^\infty \mathbb{P}\left[|m|^p > t\right] dt = \int_0^\infty \mathbb{P}\left[|m| > u\right] pu^{p-1}\, du \leq 2\int_0^\infty pu^{p-1} e^{-\frac{u^2}{2c_1^2}}\, du$$
$$= \left(\sqrt{2}c_1\right)^p 2\int_0^\infty pt^{p-1}e^{-t^2}\, dt \leq \left(\sqrt{2}c_1\right)^p p\Gamma(p/2) \leq \left(\sqrt{2}c_1\right)^p p\,(p/2)^{p/2},$$

  where we have used a change of variable $t = u^p$ in the second equality, i) in the first inequality, definition of the Gamma function in the third equality, and a Stirling's approximation in the second inequality. Taking the $pth$-root both side and using the fact that $p^{1/p} < 2$ ends the proof with $c_2 = 2c_1$.
- $2 \Rightarrow 3)$: By Taylor expansion and the monotone convergence theorem we have

$$\mathbb{E}\left[e^{\lambda^2 m^2}\right] = 1 + \sum_{k=1}^\infty \frac{\lambda^{2k}}{k!}\mathbb{E}\left[m^{2k}\right] \leq 1 + \sum_{k=1}^\infty \frac{\lambda^{2k}}{(k/e)^k}\left(8c_1^2\right)^k k^k = \frac{1}{1 - 8c_1^2 e\lambda^2},$$

  where we have used $ii)$ and a Stirling approximation $k! \geq (k/e)^k$ in the first inequality. The last equality holds provided that $8c_1^2 e\lambda^2 \leq 1$. We can further bound the right hand side if we take $\lambda$ such that $8c_1^2 e\lambda^2 \leq 1/2$ and use the inequality $(1-x)^{-1} \leq e^{2x}$ for $x \in [0, 1/2]$, i.e.,

$$\mathbb{E}\left[e^{\lambda^2 m^2}\right] \leq e^{16c_1^2 e\lambda^2},$$

  which ends the proof by taking $c_3^2 = 16c_1^2 e$.
- $3 \Rightarrow 1)$: we have

$$\mathbb{P}\left[|m| \geq t\right] = \mathbb{P}\left[e^{m^2} \geq e^{t^2}\right] \stackrel{Markov}{\leq} \frac{\mathbb{E}\left[e^{m^2}\right]}{e^{t^2}},$$

which ends the proof by applying *iii*) with $\lambda = c_3 = 1$.

- $4 \Rightarrow 1$): Using Chernoff inequality (12.1) we have

$$\mathbb{P}[m \geq t] \leq \min_{\lambda>0} \frac{\mathbb{E}\left[e^{\lambda m}\right]}{e^{\lambda t}} \leq \min_{\lambda>0} e^{\sigma^2\lambda^2/2-\lambda t} = e^{-\frac{t^2}{2\sigma^2}},$$

where we have used *iv*) in the second inequality.

*Remark 13.1.* Note that the first three assertions are equivalent for any random variable. The equivalence says that if a random variable has a exponential decaying tail bound of the form (13.2), not only its expectation (see Proposition 12.1) is bounded, but its $L^p$-norm grows like $\mathscr{O}(\sqrt{p})$. It also says that the exponential decaying tail bound (13.2) is necessary and sufficient for the integrability of fast growing function $e^{\lambda^2 m^2}$. Assertion *iv*) tells us that sub-gaussian random variables have all these properties, which is not surprising since a Gaussian random varialbe is also a sub-gaussian random variable.

Similar to Gaussian distributions, a finite sum of sub-gaussian random variables is sub-gaussian.

**Proposition 13.2 (Sum of independent sub-gaussians).** *Assume $m_1, \ldots, m_N$ are independent, sub-gaussian random variables with proxy $c_i^2$, then $\sum_{i=1}^{N} \mathbf{a}_i m_i$ is also a sub-gaussian random variable with proxy $\sum_{i=1}^{N} c_i^2 \mathbf{a}_i^2$.*

*Proof.* The proof is straightforward:

$$\mathbb{E}\left[e^{\lambda \sum_{i=1}^{N} \mathbf{a}_i m_i}\right] = \Pi_{i=1}^{N} \mathbb{E}\left[e^{\lambda \mathbf{a}_i m_i}\right] \leq e^{\lambda^2 \sum_{i=1}^{N} c_i^2 \mathbf{a}_i^2}.$$