Stat 6390.001: High Dimensional Probability and Statistics, Fall 2026 **Topic 2**

Sub-Gaussian and Sub-Exponential Variables

Lecturer: Yun Wei *Scribe: Yizhe He*

1 Example of Sub-Gaussian Variables

Remark 1.1. If X is σ-sub-Gaussian for some $\sigma > 0$, then X is also σ'-sub-Gaussian for some $\sigma' \geq \sigma$.

Example 1.2 (Bounded random variables)**.** Let X be zero-mean, and supported on some interval $[a, b]$. We will show that X is sub-Gaussian with parameter at most $\sigma = b - a$.

Letting X' be an independent copy of X, so that $X \overset{d}{=} X'$ and $X \perp\!\!\!\perp X'$. For any $\lambda \in \mathbb{R}$, we have

$$\mathbb{E}_X\left[e^{\lambda X}\right] = \mathbb{E}_X\left[e^{\lambda(X - \mathbb{E}_{X'}[X'])}\right] \overset{\text{(i)}}{\leq} \mathbb{E}_X\left[\mathbb{E}_{X'} e^{\lambda(X-X')}\right] = \mathbb{E}_{X,X'}\left[e^{\lambda(X-X')}\right]$$

The first equality is because $\mathbb{E}[X'] = 0$, and (i) follows from the the convexity of the exponential, and Jensen's inequality. Letting ε be an independent Rademacher variable, note that since (X, X') are i.i.d, $X - X' \overset{d}{=} X' - X$, so $(X - X')$ is symmetric about 0. Thus, the distribution of $(X - X')$ is the same as that of $\varepsilon(X - X')$, so that we have

$$\mathbb{E}_{X,X'}\left[e^{\lambda(X-X')}\right] = \mathbb{E}_{X,X'}\left[\mathbb{E}_\varepsilon\left[e^{\lambda\varepsilon(X-X')}\right]\right] \overset{\text{(ii)}}{\leq} \mathbb{E}_{X,X'}\left[e^{\frac{\lambda^2}{2}(X-X')^2}\right] \leq e^{\frac{\lambda^2}{2}(b-a)^2}$$

The first equality follows from Tonelli's theorem and (ii) follows from the sub-Gaussianity of Radamacher variable and last inequality is because $|X - X'| \leq b - a$. Putting everything together, we have shown that X is sub-Gaussian with parameter at most $\sigma = b - a$.

Remark 1.3. This result is useful, but it can be sharpened. It can be shown that X is sub-Gaussian with parameter at most $\sigma = \frac{b-a}{2}$ (Exercise 2.4). The technique used in this example is a simple example of a *symmetrization argument*, in which we first introduce an independent copy X', and then symmetrize the problem with a Rademacher variable. Such symmetrization arguments are useful in a variety of contexts.

2 Hoeffding's Inequality

Another property of sub-Gaussianity is preserved by linear operations.

Lemma 2.1. *Let X_1 and X_2 be independent sub-Gaussian variables with parameters σ_1 and σ_2, respectively, then $X_1 + X_2$ is sub-Gaussian with parameter $\sqrt{\sigma_1^2 + \sigma_2^2}$.*

Proof. By independence we have

$$\mathbb{E}\left[e^{\lambda(X_1+X_2-\mathbb{E}X_1-\mathbb{E}X_2)}\right] = \mathbb{E}\left[e^{\lambda(X_1-\mathbb{E}X_1)}\right]\mathbb{E}\left[e^{\lambda(X_2-\mathbb{E}X_2)}\right] \leq e^{\frac{\lambda^2}{2}\sigma_1^2}e^{\frac{\lambda^2}{2}\sigma_2^2} = e^{\frac{\lambda^2}{2}(\sigma_1^2+\sigma_2^2)}$$

□

As a consequence of this fact, we obtain an important result known as the *Hoeffding's inequality*:

Proposition 2.2 (Hoeffding's inequality)**.** *Suppose that the variables X_i, $i = 1, \ldots, n$, are independent, and X_i has mean μ_i and sub-Gaussian parameter σ_i. Then for all $t \geq 0$, we have*

$$\mathbb{P}\left[\sum_{i=1}^{n}(X_i - \mu_i) \geq t\right] \leq \exp\left\{-\frac{t^2}{2\sum_{i=1}^{n}\sigma_i^2}\right\}$$

and

$$\mathbb{P}\left[\left|\sum_{i=1}^{n} X_i - \sum_{i=1}^{n}\mathbb{E}X_i\right| \geq t\right] \leq 2\exp\left\{-\frac{t^2}{2\sum_{i=1}^{n}\sigma_i^2}\right\}$$

Proof. By previous lemma, $\sum_{i=1}^{n} X_i$ is $\sqrt{\sigma_1^2 + \cdots + \sigma_n^2}$ sub-Gaussian. So, by basic sub-Gaussian tail bound we have the results. □

Corollary 2.3. *If X_i, $i = 1, \ldots, n$, are i.i.d, with mean μ and sub-Gaussian parameter σ. Then for all $t \geq 0$, we have*

$$\mathbb{P}\left[\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| \geq t\right] \leq 2\exp\left\{-\frac{nt^2}{2\sigma^2}\right\}$$

and

$$\mathbb{P}\left[\sqrt{n}\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| \geq t\right] \leq 2\exp\left\{-\frac{t^2}{2\sigma^2}\right\}$$

Remark 2.4. The above can be viewed as a finite sample Central Limit Theorem. CLT states that when $n \to \infty$,

$$\sqrt{n}\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| \to \mathcal{N}(0, \sigma^2).$$

The second statement of the above corollary is a finite sample version of CLT, by stating that the tail of $\sqrt{n}\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right|$ is upper bounded by the tail of $\mathcal{N}(0, \sigma^2)$.

Proof. By Hoeffding,

$$\mathbb{P}\left[\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| \geq t\right] = \mathbb{P}\left[\left|\sum_{i=1}^{n} X_i - n\mu\right| \geq nt\right] \leq 2\exp\left\{-\frac{(nt)^2}{2n\sigma^2}\right\} = 2\exp\left\{-\frac{nt^2}{2\sigma^2}\right\}$$

Second statement follows analogously. □

One application of Hoeffding's bound is to derive finite sample guarantees for the error of the sample mean.

Example 2.5 (Finite-sample error guarantee)**.** Let $\overline{X}_n := \frac{1}{n}\sum_{i=1}^{n} X_i$ and fix a confidence level $\delta \in (0, 1)$. By the preceding corollary,

$$\mathbb{P}\left[\left|\overline{X}_n - \mu\right| \geq t\right] \leq 2\exp\left\{-\frac{nt^2}{2\sigma^2}\right\}$$

Choose $t = \sqrt{\frac{2\sigma^2}{n}\log\left(\frac{2}{\delta}\right)}$, so that $\delta = 2\exp\left\{-\frac{nt^2}{2\sigma^2}\right\}$. In other words, with probability at least $1 - \delta$,

$$\left|\overline{X}_n - \mu\right| \leq \sqrt{\frac{2\sigma^2}{n}\log\left(\frac{2}{\delta}\right)}$$

Example 2.6 (Expected absolute error)**.** We can also control the expected absolute error $\mathbb{E}\left|\overline{X}_n - \mu\right|$ directly from the tail bound. Recall that for any nonnegative random variable Y, Tonelli's theorem gives

$$\mathbb{E}[Y] = \int_0^\infty \mathbb{P}(Y \geq t)\, dt$$

Taking $Y = \left|\overline{X}_n - \mu\right|$, the preceding corollary and the trivial bound $\mathbb{P}(Y \geq t) \leq 1$ imply

$$\mathbb{P}\left[\left|\overline{X}_n - \mu\right| \geq t\right] \leq \min\left\{1,\ 2\exp\left\{-\frac{nt^2}{2\sigma^2}\right\}\right\}$$

Let t_0 be the point at which the two upper bounds agree. Thus $2e^{-\frac{nt_0^2}{2\sigma^2}} = 1$, which gives $t_0 = \sigma\sqrt{\frac{2\log 2}{n}}$. Hence

$$\begin{aligned}
\mathbb{E}\left|\overline{X}_n - \mu\right| &= \int_0^\infty \mathbb{P}\left[\left|\overline{X}_n - \mu\right| \geq t\right]\, dt \\
&\leq \int_0^{t_0} 1\, dt + \underbrace{\int_{t_0}^\infty 2\exp\left\{-\frac{nt^2}{2\sigma^2}\right\} dt}_{z = \frac{\sqrt{n}}{\sigma} t, \quad dt = \frac{\sigma}{\sqrt{n}}\, dz} \\
&= t_0 + \frac{2\sigma}{\sqrt{n}} \int_{\frac{\sqrt{n}}{\sigma} t_0}^\infty e^{-z^2/2}\, dz \\
&= \frac{\sigma}{\sqrt{n}}\left(\sqrt{2\log 2} + 2\int_{\sqrt{2\log 2}}^\infty e^{-z^2/2}\, dz\right) \\
&\leq 2\frac{\sigma}{\sqrt{n}}.
\end{aligned}$$

We conclude our discussion of sub-Gaussianity with a theorem that provides different characterizations of sub-Gaussian variables.

Theorem 2.7 (Equivalent characterizations of sub-Gaussian variables)**.** *Given any zero-mean random variable X, the following properties are equivalent:*

1. *There is a constant $\sigma \geq 0$ such that*

$$\mathbb{E}\left[e^{\lambda X}\right] \leq e^{\frac{\lambda^2\sigma^2}{2}} \qquad \textit{for all } \lambda \in \mathbb{R}.$$

2. *There is a constant $c \geq 0$ and Gaussian random variable $Z \sim \mathcal{N}(0, \tau^2)$ such that*

$$\mathbb{P}\left[|X| \geq s\right] \leq c\,\mathbb{P}\left[|Z| \geq s\right] \qquad \textit{for all } s \geq 0.$$

3. *There is a constant $\theta \geq 0$ such that*

$$\mathbb{E}\left[X^{2k}\right] \leq \frac{(2k)!}{2^k k!}\theta^{2k} \qquad \textit{for all } k = 1, 2, \ldots$$

4. *There is a constant $\sigma \geq 0$ such that*

$$\mathbb{E}\left[e^{\frac{\lambda X^2}{2\sigma^2}}\right] \leq \frac{1}{\sqrt{1-\lambda}} \qquad \textit{for all } \lambda \in [0, 1).$$

Proof. See Appendix A (Section 2.4) in [1]. □

3 Sub-Exponential Variables

We now turn to the class of sub-exponential variables, which are defined by a slightly milder condition on the moment generating function.

Definition 3.1. A random variable X with mean $\mu = \mathbb{E}[X]$ is *sub-exponential* if there are non-negative parameters (ν, α) such that

$$\mathbb{E}\left[e^{\lambda(X-\mu)}\right] \leq e^{\frac{\nu^2\lambda^2}{2}} \qquad \text{for all } |\lambda| < \frac{1}{\alpha}$$

References

[1] M. J. Wainwright, *High-Dimensional Statistics: A Non-Asymptotic Viewpoint*, Cambridge University Press, 2019.