STAT 620: Asymptotic Statistics **Spring 2022**

Lecture: Mar 10

Lecturer: Xianyang Zhang

1 Sub-Gaussianity

X is a σ^2-sub-Gaussian random variable if

$$E[e^{\lambda(X-EX)}] \leq e^{\frac{\lambda^2\sigma^2}{2}},\ \forall \lambda \in \mathbb{R}.$$

1.1 Example: Hoeffding's inequality

If $X \in [a, b]$, then X is $\frac{(b-a)^2}{4}$-sub-Gaussian, which means

$$E[e^{\lambda(X-EX)}] \leq e^{\frac{\lambda^2(a-b)^2}{8}}.$$

Proof. Assume X has zero-mean (otherwise we can center the random variable).
We have $\lambda x = \frac{x-a}{b-a}\lambda b + \frac{b-x}{b-a}\lambda a$.
Then

$$e^{\lambda x} \leq \frac{x-a}{b-a}e^{\lambda b} + \frac{b-x}{b-a}e^{\lambda a} \text{ (Jensen's inequality)}$$

and thus

$$E[e^{\lambda x}] \leq \frac{-a}{b-a}e^{\lambda b} + \frac{b}{b-a}e^{\lambda a} = e^{\lambda a}\left(\frac{b}{b-a} - \frac{a}{b-a}e^{\lambda(b-a)}\right).$$

Let $p = -\frac{a}{b-a}$, $u = \lambda(b-a)$ so that $pu = -\lambda a$. Then

$$E[e^{\lambda x}] \leq e^{-pu}(1-p+pe^u) = e^{-pu+\log(1-p+pe^u)} \triangleq e^{\varphi(u)}$$

For $\varphi(u)$, we have

$$\varphi(0) = 0,\ \varphi'(0) = 0,\ \varphi''(u) \leq \frac{1}{4}$$

which implies that

$$\varphi(u) = \varphi(0) + \varphi'(0)u + \varphi''(\tilde{u})\frac{u^2}{2} \leq \frac{u^2}{8} \text{ (Taylor's theorem)}$$

where $\tilde{u} \in [0, u]$.

1.2 Proposition 1

Let X_i's be independent σ_i^2-sub-Gaussian random variables, then $\sum_{i=1}^n X_i$ is $\sum_{i=1}^n \sigma_i^2$-sub-Gaussian.

Proof. Use definition and independence.

1.3 Proposition 2 (Cramér-Chernoff Method)

Let X be σ^2-sub-Gaussian, then

$$\max\{P(X - EX \geq t),\ P(X - EX \leq -t)\} \leq e^{-\frac{t^2}{2\sigma^2}},\ \forall t \geq 0.$$

Proof. Assume $EX = 0$,

$$P(X \geq t) = P(e^{\lambda X} \geq e^{\lambda t}) \leq \frac{1}{e^{\lambda t}} E[e^{\lambda X}] \leq e^{\frac{\lambda^2\sigma^2}{2} - \lambda t},\ \forall \lambda \geq 0.$$

Therefore,

$$P(X \geq t) \leq \inf_{\lambda \geq 0} e^{\frac{\lambda^2\sigma^2}{2} - \lambda t} = e^{-\frac{t^2}{2\sigma^2}}.$$

1.4 Proposition 3

Let X_i's be independent σ_i^2-sub-Gaussian random variables, then

$$P\left(\frac{1}{n}\sum_{i=1}^n (X_i - EX_i) \geq t\right) \leq e^{-\frac{n^2 t^2}{2\sum_{i=1}^n \sigma_i^2}}.$$

Proof. Note that $P\left(\frac{1}{n}\sum_{i=1}^n (X_i - EX_i) \geq t\right) = P\left(\sum_{i=1}^n (X_i - EX_i) \geq nt\right)$. Then use the results of Propositions 1 and 2.

1.5 Corollary 1

Let $\{X_i\}_{i=1}^n$ be i.i.d. σ^2-sub-Gaussian random variables with $E(X_i) = \mu$, then

$$P\left(\frac{1}{n}\sum_{i=1}^n X_i - \mu \geq t\right) \leq e^{-\frac{nt^2}{2\sigma^2}}.$$

1.6 Corollary 2

Let $\{X_i\}_{i=1}^n$ be i.i.d. random variables with $E(X_i) = \mu$ and $X_i \in [a, b]$, then

$$P\left(\frac{1}{n}\sum_{i=1}^n X_i - \mu \geq t\right) \leq e^{-\frac{2nt^2}{(a-b)^2}}.$$

1.7 Proposition 4

Let $\{X_i\}_{i=1}^n$ be zero-mean σ^2-sub-Gaussian random variables (possibly dependent), then

$$E[\max_{1\le i\le n} X_i] \le \sqrt{2\sigma^2 \log n}$$

Proof. Notice that

$$\begin{aligned} e^{\lambda E[\max_{1\le i\le n} X_i]} &\le E[e^{\lambda \max_{1\le i\le n} X_i}] \text{ (Jensen's inequality)} \\ &\le E[\sum_{i=1}^n e^{\lambda X_i}] \le n \cdot e^{\frac{\lambda^2\sigma^2}{2}}. \end{aligned}$$

Thus we have

$$\lambda E[\max_{1\le i\le n} X_i] \le \log n + \frac{\lambda^2\sigma^2}{2},$$

which implies that

$$E[\max_{1\le i\le n} X_i] \le \frac{\log n}{\lambda} + \frac{\lambda\sigma^2}{2},\ \forall \lambda \ge 0.$$

Therefore,

$$E[\max_{1\le i\le n} X_i] \le \inf_{\lambda\ge 0}\left\{\frac{\log n}{\lambda} + \frac{\lambda\sigma^2}{2}\right\} = \sqrt{2\sigma^2 \log n} \text{ (by setting the derivative equal to zero).}$$

2 Symmetrization

2.1 Motivation

To derive the uniform laws of large numbers (ULLN), an intuitive idea is just using the definition and Markov's inequality:

$$\begin{aligned} P\left(\sup_{f\in\mathcal{F}} |P_n f - Pf| \ge t\right) \le& t^{-1} E\left[\sup_{f\in\mathcal{F}} |P_n f - Pf|\right] = t^{-1}\cdot E\left[\sup_{f\in\mathcal{F}} |\frac{1}{n}\sum_{i=1}^n (f(X_i) - Ef(X_i))|\right] \\ =& \frac{1}{tn}\cdot E\left[\sup_{f\in\mathcal{F}} |\sum_{i=1}^n (f(X_i) - Ef(X_i))|\right]. \end{aligned}$$

If we can show

$$E\left[\sup_{f\in\mathcal{F}} |\sum_{i=1}^n (f(X_i) - Ef(X_i))|\right] = o(n),$$

then

$$P\left(\sup_{f\in\mathcal{F}} |P_n f - Pf| \ge t\right) \to 0.$$

2.2 Rademacher random variable

ϵ is a Rademacher random variable if

$$P[\epsilon = 1] = P[\epsilon = -1] = \frac{1}{2}.$$

2.3 Symmetrization

Let $\{X_i\}_{i=1}^n$ be independent random variables in a normed space with the norm $\|\cdot\|$ and let $\{\epsilon_i\}_{i=1}^n$ be i.i.d. Rademacher random variables which are independent of $\{X_i\}_{i=1}^n$. For $p \geq 1$, we have

$$E\left[\|\sum_{i=1}^n (X_i - EX_i)\|^p\right] \leq 2^p E\left[\|\sum_{i=1}^n \epsilon_i X_i\|^p\right].$$