15/1/19 EE5203: Concentration Inequalities

Today:
- Review
  - Hoeffding's lemma
  - Hoeffding's inequality
- Sub-Gaussian RV
- Bennett's inequality

Review:
- Hoeffding's lemma: For a random variable X with $E[X]=0$ and $a \le X \le b$,

$$E[e^{sX}] \le e^{s^2\frac{(b-a)^2}{8}}.$$

Recall $L(h) = -hp + \log[(1-p) + p \cdot e^h]$ where

$$h = s.(b-a), \quad p = \frac{-a}{(b-a)}$$

$$L''(h) \le \frac{1}{4} \text{ for any } h.$$

- Hoeffding's inequality: If $S_n = \sum_{i=1}^{n} X_i$ where X_i's are independent RVs with $a_i \le X_i \le b_i$ then

$$P\{S_n - ES_n \ge t\} \le \exp\left[\frac{-2t^2}{\sum_{i=1}^{n}(b_i-a_i)^2}\right],$$

$$P\{S_n - ES_n \le -t\} \le \exp\left[-\frac{2t^2}{\sum_{i=1}^{n}(b_i-a_i)^2}\right]$$

- Taylor's theorem: If $f(x)$ is a continuous function in the bounded interval $[a,b]$ and has $f'(x)$, and $f''(x)$ defined in this interval then

$$f(h) = f(0).h^0 + f'(0).\frac{h^1}{1!} + f''(v)\frac{h^2}{2!}$$ where $v \in (0, h)$.

- **Sub-Gaussian RV:** A real valued RV X is said to be σ^2-sub-Gaussian if there exists a σ such that
$$E[e^{\lambda X}] \leq \exp\left(\frac{\lambda^2\sigma^2}{2}\right) \quad \text{for any } \lambda \in \mathbb{R}.$$

- observation about the Hoeffding's inequality: the bound does not involve the variance of the RV.
- Question: Can we find a tighter bound when the RV has low variance? Yes, via the Bennett's inequality.

Bennett's inequality: Let $X_1, \ldots X_n$ be independent RVs with finite variance and $X_i \leq b$ for $b > 0$ almost surely for $i \leq n$. Let $v = \sum_{i=1}^{n} E[X_i^2]$. If $S = \sum_{i=1}^{n}(X_i - EX_i)$, then
$$\log\left[Ee^{\lambda S}\right] \leq n \cdot \log\left[1 + \frac{v}{nb^2}\phi(\lambda b)\right] \leq \frac{v}{b^2}\phi(\lambda b),$$
where $\phi(u) = e^u - u - 1 \quad u \in \mathbb{R}$.

Proof: • Let us assume $b = 1$.

- Show that $u^{-2}\cdot\phi(u)$ is a non decreasing function.
$$g(u) = \frac{\phi(u)}{u^2} \Rightarrow g'(u) = \frac{e^u(u-2) + (u+2)}{u^3} \overset{?}{\geq} 0$$
$$(\lambda X_i)^{-2}\phi(\lambda X_i) \leq \lambda^{-2}\phi(\lambda) \text{ (due } X_i \leq 1,\ \lambda > 0$$
to $g(u)$ being non-decreasing)
$$\Rightarrow \phi(\lambda X_i) \leq X_i^2\,\phi(\lambda)$$
i.e. $e^{\lambda X_i} - \lambda X_i - 1 \leq X_i^2(e^\lambda - \lambda - 1)$.

Apply the expectation operator.

$$E[e^{\lambda X_i} - \lambda X_i - 1] \leq \phi(\lambda) \cdot E[X_i^2]$$

$$\underbrace{E[e^{\lambda X_i}]} \leq (E[\lambda X_i] + 1 + E[X_i^2] \cdot \phi(\lambda))$$

Apply log.

$$\log E[e^{\lambda X_i}] \leq \log (E[\lambda X_i] + 1 + E[X_i^2] \cdot \phi(\lambda)]$$

$$\sum_{i=1}^{n} \log E[e^{\lambda X_i}] \leq \sum_{i=1}^{n} \log (E[\lambda X_i] + 1 + E[X_i^2] \cdot \phi(\lambda))$$

$$\log \left[\prod_{i=1}^{n} E[e^{\lambda X_i}]\right] \leq \text{---}\,\text{''}\,\text{---}$$

$$\underbrace{\log \left[\prod_{i=1}^{n} E[e^{\lambda \cdot X_i - E[X_i]}]\right]} + \sum_{i=1}^{n} \lambda \cdot E[X_i] \leq \sum_{i=1}^{n} \log (E[\lambda X_i] + 1 + E[X_i^2] \phi(\lambda)]$$

$\psi_S(\lambda)$ where $\psi_S(\lambda) = \log [E[e^{\lambda S}]]$.

$$\Rightarrow \psi_S(\lambda) \leq \sum_{i=1}^{n} \log [E[\lambda X_i] + 1 + E[X_i^2] \phi(\lambda)]$$

$$- \sum_{i=1}^{n} \lambda \cdot E[X_i].$$

$$\leq n \sum_{i=1}^{n} \frac{\log}{n} [\qquad\qquad] - \sum_{i=1}^{n} \lambda E[X_i]$$