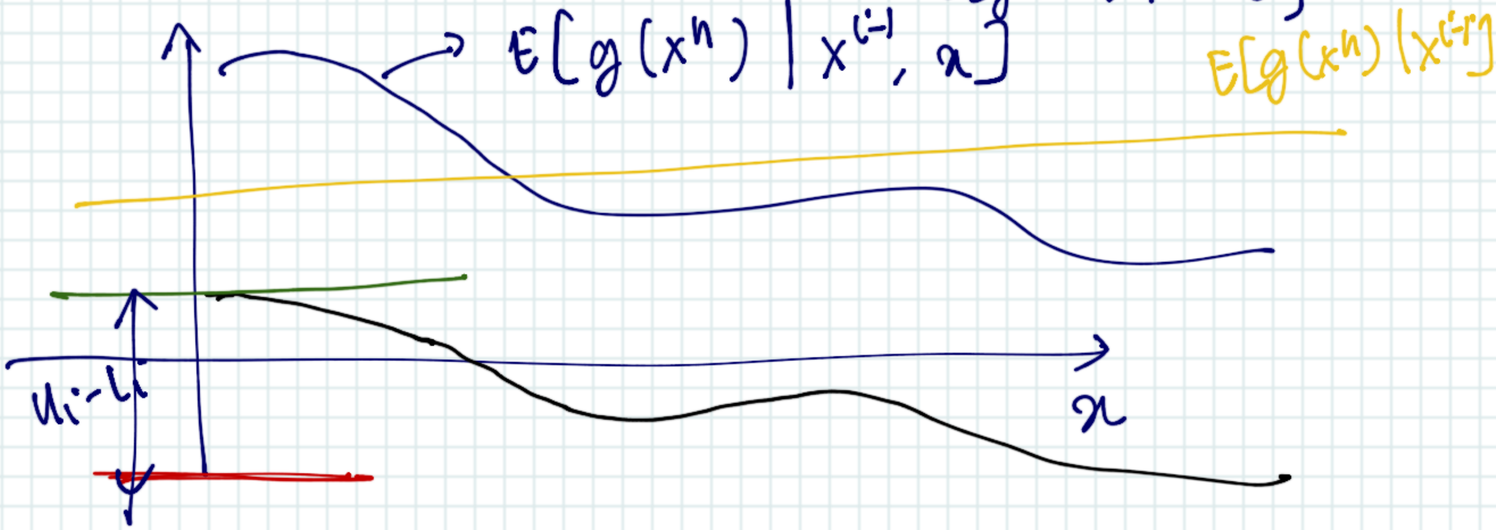
31/1/19 EE5603: Concentration Inequalities

Today: • Review McDiarmid's inequality

• Efron-Stein inequality

Review: Show that $U_i - L_i \leq c_i$, $U_i = \sup_{x' \in \mathcal{X}} \{ E[g(X^n) \mid X^{i-1}, x'] - E[g(X^n) \mid X^{i-1}] \}$

$$L_i = \inf_{x \in \mathcal{X}} \{ E[g(X^n) \mid X^{i-1}, x] - E[g(X^n) \mid X^{i-1}] \}$$



$$U_i - L_i = \sup_{x \in \mathcal{X}} \sup_{x' \in \mathcal{X}} \{ E[g(X^n) \mid X^{i-1}, x] - E[g(X^n) \mid X^{i-1}, x'] \}$$

Efron-Stein inequality: Let $X_1 \ldots X_n$ be independent RVs. Let $f: \mathcal{X}^n \to \mathbb{R}$ be a square integrable, $Z = f(X_1 \ldots X_n)$.

$$\mathrm{Var}(Z) \leq \sum_{i=1}^{n} E[(Z - E^{i}(Z))^2] \stackrel{\text{def}}{=} v$$

• $E_i(Z) = E[f(X_1 \ldots X_n) \mid X^i]$; $E_0 = E$

- $E^i(z) = \int_{x_i \in X} f(x_1 \ldots x_i, \ldots x_n)\, dP(x_i)$

- If $\Delta_i = E_i(z) - E_{i-1}(z)$, $\sum_{i=1}^{n} \Delta_i = Z - E(z)$ — ①

- $Var(z) = E[(z - E(z))^2]$ (from defn)

$$= E\left[\left(\sum_{i=1}^{n} \Delta_i\right)^2\right] \quad \text{(from ①)}$$

$$= E\left[\sum_{i=1}^{n} \Delta_i^2 + 2\sum_{j>i} \Delta_i \Delta_j\right]$$

$$= E\left[\sum_{i=1}^{n} \Delta_i^2\right] + \underbrace{2 \cdot \sum_{j>i} E[\Delta_j \Delta_i]}_{?}$$

Claim: If $j > i$, $E_i \Delta_j = 0$ (Fubini's theorem)

$$E_i \Delta_j = E_i[\underline{E_j(z)} - \underline{E_{j-1}(z)}]$$

$$= E[\underbrace{E_j(z)} - \underbrace{E_{j-1}(z)} \mid x^i]$$

$$= E_i(z) - E_i(z) \quad \text{(since } j > i\text{)}$$

$$= 0$$

If $E_i \Delta_j = 0$ $j > i$, then $2\sum_{j>i} E[\Delta_i \Delta_j] = ?$ $\overset{0}{\underset{(HW)}{=}}$

$$\Rightarrow \mathrm{Var}(Z) = E\left(\sum_{i=1}^{n} \Delta_i^2\right). \quad - (2)$$

• Recall. $E_i\left[E^i(Z)\right] = E_{i-1}(Z)$

$$E^i(Z) = \int_{x_i \in \mathcal{X}} f(x_1 \ldots x_{i-1}, x_i, x_{i+1} \ldots x_n) \cdot dP(x_i)$$

$$E_i(Z) = \int_{x_{i+1}^n \in \mathcal{X}^{n-i}} f(x_1 \ldots x_{i-1}, x_i, x_{i+1} \ldots x_n) \cdot dP(x_{i+1}^n)$$

$$\Rightarrow E_i\left[E^i(Z)\right] = \int_{x_{i+1}^n \in \mathcal{X}^{n-i}} \int_{x_i \in \mathcal{X}} f(x_1 \ldots x_{i-1}, x_i, x_{i+1} \ldots x_n) \cdot dP(x_i)\, dP(x_{i+1}^n)$$

$$= \int_{x_i^n \in \mathcal{X}^{n-(i-1)}} f(x_1 \ldots x_{i-1}, x_i, \ldots x_n) \cdot dP(x_i^n)$$

(due to independence of X_i)

$$= E_{i-1}(Z) \quad - (3)$$

• $\Delta_i = E_i(Z) - E_{i-1}(Z)$

(from (3))

$$\Rightarrow \Delta_i = E_i\left[Z - E^i(Z)\right].$$

$$\Delta_i^2 = \left[E_i\left[Z - E^i(Z)\right]\right]^2.$$

Jensen's inequality: If g is a convex function and X is a random variable then

$$E[f(X)] \geq f(E[X])$$

Apply this result to Δ_i

$$E[\Delta_i^2] \geq \left(E[\Delta_i]\right)^2 \quad \text{or}$$

$$\left(E[\Delta_i]\right)^2 \leq E[\Delta_i^2] \qquad - \quad (4)$$

Recall $\quad Var(Z) = E\left[\sum_{i=1}^{n} \Delta_i^2\right]$