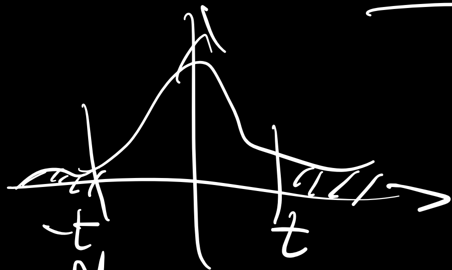
Concentration Ineq. . Let $X_1, \cdots, X_n$ be indep. r.v.

(Hoeffding's ineq). bounded a.s. in $[a,b]$.

Let $S_n = \sum_{i=1}^{n} X_i$. $\qquad Var\left(\frac{S_n}{n}\right) \sim \frac{1}{n}$.

Then $\forall t>0$, $\Pr\left[\left| \frac{S_n}{n} - \frac{\mathbb{E}[S_n]}{n} \right| \geq t \right] \leq 2e^{-2nt^2/(b-a)^2}$.



$\boxed{(b-a)\sqrt{\frac{1}{2n}\ln\frac{2}{\delta}}}$ $\qquad \boxed{\delta}$

Informally, w.p. $\geq 1-\delta$. $\left| \frac{S_n}{n} - \frac{\mathbb{E}[S_n]}{n} \right| \leq (b-a)\sqrt{\frac{1}{2n}\ln\frac{2}{\delta}}$.

Supervised Learning. $\{(X_i, Y_i)\}_{i=1}^{n} \overset{i.i.d.}{\sim} P$.

binary.

$$\mathcal{F} \subseteq \mathcal{X} \to \mathcal{Y}.$$

ERM: $\hat{f} = \underset{f \in \mathcal{F}}{\operatorname{argmin}} \ \hat{\mathbb{E}}\left[\mathbb{I}[f(X) \neq Y] \right]$

$$\frac{1}{n} \sum_{i=1}^{n} \mathbb{I}[f(X_i) \neq Y_i].$$

Define $f^* = \underset{f \in \mathcal{F}}{\operatorname{argmin}} \ \mathbb{E}_P\left[\mathbb{I}[f(X) \neq Y] \right]$.

Bound:

$$\mathbb{E}[\mathbb{I}[\hat{f}(X)\neq Y]] - \mathbb{E}[\mathbb{I}[f^*(X)\neq Y]].$$

Proof: Fix any $f \in \mathcal{F}$. w.p. $\geq 1-\delta$.

$$\left|\hat{\mathbb{E}}[\mathbb{I}[f(X)\neq Y]] - \mathbb{E}[\mathbb{I}[f(X)\neq Y]]\right| \leq \sqrt{\frac{1}{2n}\ln\frac{2}{\delta}}$$

$\forall f \in \mathcal{F}$. w.p. $\geq 1-\delta$.

w.p. $\geq 1-\delta$ $\forall f \in \mathcal{F}$

$$\mathbb{E}[\mathbb{I}[\hat{f}(X)\neq Y]] - \mathbb{E}[\mathbb{I}[f^*(X)\neq Y]]$$

$$\leq \cdots - \hat{\mathbb{E}}[\mathbb{I}[\hat{f}(X)\neq Y]] + \hat{\mathbb{E}}[\mathbb{I}[f^*(X)\neq Y]] - \cdots$$

$$\leq 2 \cdot \max_{f\in\mathcal{F}} \left|\mathbb{E}[\mathbb{I}[f(X)\neq Y]] - \hat{\mathbb{E}}[\mathbb{I}[f(X)\neq Y]]\right|$$

$$\leq 2 \cdot \sqrt{\frac{1}{2n}\ln\frac{2|\mathcal{F}|}{\delta}}$$

$f^1, f^2, \ldots, f^m \qquad m = |\mathcal{F}|.$

$$\forall f \in \mathcal{F} \; \Pr\left[\left| \mathbb{E} f - \hat{\mathbb{E}} f \right| > \cdots \right] \leq \delta.$$

$$\Pr\left[\exists f \in \mathcal{F}, \; \cdots \right] \leq |\mathcal{F}| \delta.$$

$\vdots$

$$\Pr[A \cup B] \leq \Pr[A] + \Pr[B].$$

Multi-armed Bandit (MAB)

K arms (actions). each arm i is associated

$\sim$ reward distribution $\underline{R_i} \in \Delta([0,1])$.

denote $\underline{\mu_i} = \mathbb{E}_{r \sim R_i}[r]$.

$i^* = \arg\max_{i \in \underline{[K]}} \underline{\underline{\mu_i}}$. $\{1, 2, \cdots, K\}$.

Interaction Protocol.

For round $t = 1, 2, \cdots, T$.

1. choose arm $i_t \in [K]$.
2. receive $r_t \sim R_{i_t}$.

Goal: after T rounds. recommend $\hat{i}$.

w.t. bound. $\boxed{\mu_{i^*} - \mu_{\hat{i}}}$.

Δ

Alg: 1. sample each arm $\underline{T/K}$ times.

$\hookrightarrow$ int.

2. form emp. estimation $\hat{\mu}_i$.

3. $\hat{i} = \arg\max_i \hat{\mu}_i$.

$\mu_{i^*} - \mu_{\hat{i}} \leq |\mu_i - \hat{\mu}_i|$.

$\leq \boxed{\mu_{i^*} - \hat{\mu}_{i^*}} + \hat{\mu}_{\hat{i}} - \mu_{\hat{i}}$.

$\leq 2 \cdot \max_{i \in [K]} |\mu_i - \hat{\mu}_i| \overset{\text{w.p.} \geq 1-\delta}{\leq} 2 \cdot \sqrt{\frac{1}{2n} \ln \frac{2K}{\delta}}$.

Fix i. w.p. $\geq 1-\delta$. $|\hat{\mu}_i - \mu_i| \leq \sqrt{\frac{1}{2n} \ln \frac{2}{\delta}}$.

$n = T/K$.

w.p. $\geq 1-\delta$. $\forall i$, $|\hat{\mu}_i - \mu_i| \leq \sqrt{\frac{K}{2T} \ln \frac{2K}{\delta}}$.

"Proof": $\Pr[\, |\hat{\mu}_i - \mu_i| > \underbrace{\sqrt{\frac{1}{2n} \ln \frac{2}{\delta}}}_{A_i} \,] \leq \underbrace{\delta}_{\delta/K}$.

$\forall i, \delta$.

$$\Pr\Big[\exists i, \big\{|\hat{\mu}_i - \mu_i| > \boxed{?}\big\}\Big] \le \delta.$$

$|\hat{\mu}_i - \mu_i| > \boxed{?}$: event A_i

$= \Pr\left[\bigcup_i A_i\right] \le \underline{\sum_i \Pr[A_i]} \overset{\text{need}}{\le} \delta.$

suffices if $\Pr[A_i] \le \delta / K$.

$\forall i$, w.p. $\ge 1-\delta$. $|\mu_i - \hat{\mu}_i| \le \ldots$ &

Prove.

w.p. $1-\delta$ $\quad \forall i \quad |\mu_i - \hat{\mu}_i| \le \ldots$

$\mu_{i^*} - \mu_{\hat{i}} \le \ldots$

$\vdots$

deterministic.

$\Pr[\text{Conclusion (event)}] \ge 1-\delta.$

each i is sampled. $n = T/K$.

sub-opt $\propto \sqrt{\frac{1}{n}}$.

① each i is sampled n_i times.

$$\sum_i n_i = T.$$

sub-opt $\propto \boxed{\sqrt{\frac{1}{\min_i n_i}}}$.

②. $\sum_i n_i = T$.

change alg.

sub-opt $\propto \sqrt{\frac{1}{n_{i^*}}}$.