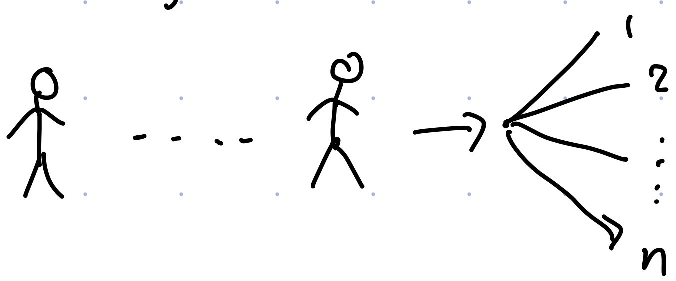
CSE 541 Interactive Learning

https://courses.cs.washington.edu/courses/cse541/24sp/

n drugs / treatments



At t^{th} arrival

I assign treatment

$I_t \in [n] = \{1, \dots, n\}$

The outcome is binary $X_{t,i} = \begin{cases} 1 & \text{if treatment } i \text{ assigned and patient survived} \\ 0 & \text{o.w.} \end{cases}$

We can model $X_{t,i}$ is an IID Bernoulli R.V.

s.t. $\mathbb{E}[X_{t,i}] = \theta_i \in [0,1] \quad \forall t.$

Objectives: 1) Identify best treatment

$\arg\max_i \theta_i$

w/ high prob as fast as possible.

2) Maximize expected reward $\mathbb{E}\left[\sum_{t=1}^{T} X_{t,I_t}\right]$

I_t is chosen given $\{(I_s, X_{I_s,s})\}_{s=1}^{t-1}$

$$\text{Regret } R_T := \max_{i=1,\dots,n} \mathbb{E}\left[\sum_{t=1}^{T} X_{i,t} - X_{I_t,t} \right]$$

$$= \max_{i=1,\dots,n} T\theta_i - \sum_{j=1}^{n} \theta_j \mathbb{E}[T_j]$$

$$T = \sum_{j=1}^{n} \mathbb{E}[T_j] \qquad = \max_i \sum_{j=1}^{n} \theta_i \mathbb{E}[T_j] - \theta_j \mathbb{E}[T_j]$$

$$= \mathbb{E}\left[\sum_{j=1}^{n} T_j\right] \qquad = \max_{i=1,\dots,n} \sum_{j=1}^{n} \Delta_j \mathbb{E}[T_j].$$

$$\Delta_j := \max_i \theta_i - \theta_j.$$

For any fixed i $\mathbb{E}[X_{i,t}] = \theta_i$. $\mathbb{E}\left[\sum_t X_{i,t}\right] = \sum_t \mathbb{E}[X_{i,t}] = T\theta_i$

$$\mathbb{E}\left[\sum_t X_{I_t,t}\right] = \mathbb{E}\left[\sum_t \underbrace{\mathbb{E}\left[X_{I_t,t} \mid \{(I_s, X_{I_s,s})\}_{s=1}^{t-1}\right]}_{\theta_{I_t}}\right]$$

$$= \mathbb{E}\left[\sum_{t=1}^{T} \theta_{I_t}\right]$$

$$= \mathbb{E}\left[\sum_{t=1}^{T}\sum_{i=1}^{n} \theta_i \mathbb{1}\{I_t = i\}\right]$$

$$= \mathbb{E}\left[\sum_{i=1}^{n} \theta_i \underbrace{\left(\sum_{t=1}^{T} \mathbb{1}\{I_t = i\}\right)}_{=:T_i}\right]$$

$$= \sum_{i=1}^{n} \theta_i \mathbb{E}[T_i]$$

$$R_T = \max_i \mathbb{E}\Big[\sum_t X_{i,t} - X_{I_t,t}\Big] = \sum_{j \neq i} \Delta_j \mathbb{E}[T_j]$$

"Good" $\stackrel{?}{=}$ $R_T = o(T)$.

Ex. If $R_T \leq \sqrt{T} \Rightarrow \frac{1}{T} R_T \to 0$

$$\max_i \mathbb{E}\Big[\frac{1}{T} \sum_{t=1}^{T} X_{i,t} - X_{I_t,t}\Big] \to 0$$

(A,B) R.V.s $\quad \mathbb{E}[f(A,B)] = \sum_{a,b} f(a,b)\, \mathbb{P}(A=a, B=b)$

$$= \sum_a \mathbb{P}(A=a) \underbrace{\sum_b f(a,b)\, \mathbb{P}(B=b \mid A=a)}_{\mathbb{E}[f(A,B) \mid A=a]}$$

$$= \mathbb{E}\big[\mathbb{E}[f(A,B) \mid A=a]\big]$$

Consider $n=2$ case.

Algorithm. Fix $\xi \in \mathbb{N}$.

Pull each arm ξ times and compute $\hat{\theta}_i = \frac{1}{\xi} \sum_{t=1}^{\xi} X_{i,\xi} \quad i \in \{1,2\}$

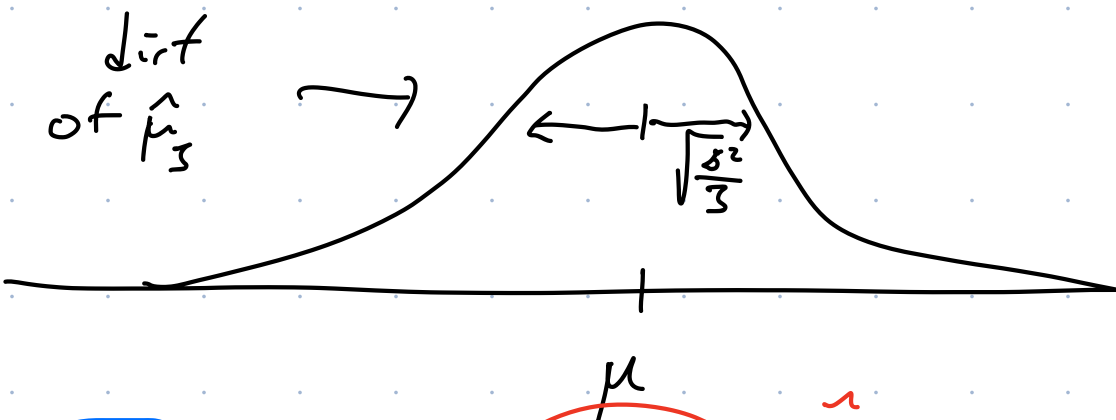
for $t > 2T$ play $\arg\max_i \hat{\theta}_i$

$\underline{\text{CLT}}$ Given IID $Z_1, \dots, Z_T$ w/ $\mathbb{E}[Z_i] = \mu$

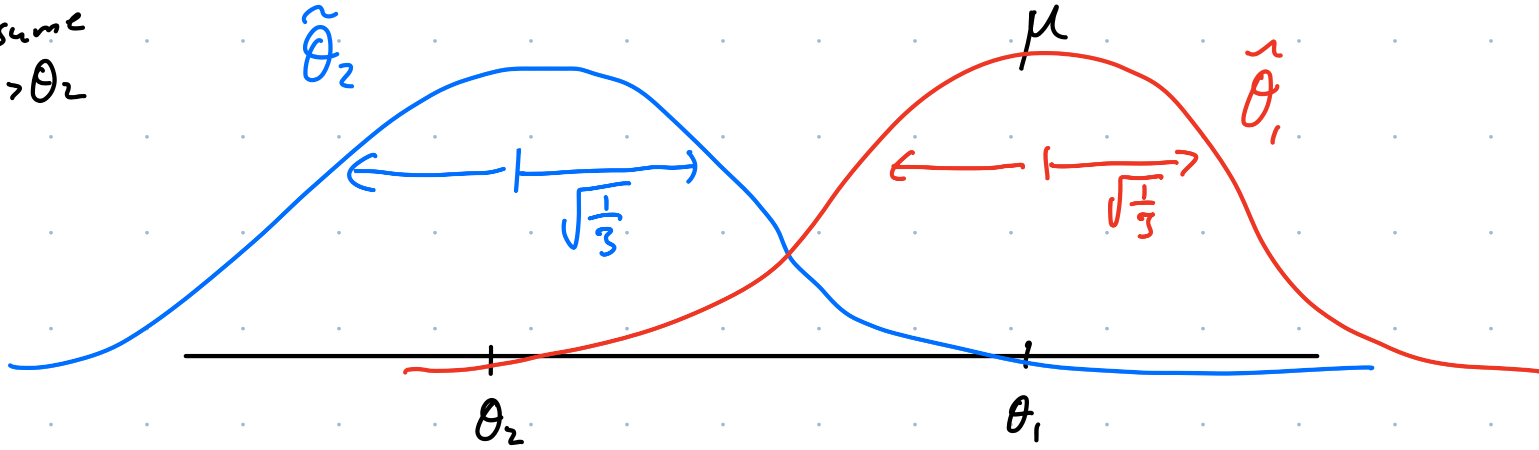
$\mathbb{E}[(Z_i - \mathbb{E}[Z_i])^2] = \sigma^2$, $\quad \hat{\mu}_T = \frac{1}{T} \sum_{t=1}^{T} Z_t$

Then $\quad \sqrt{T}(\hat{\mu}_T - \mu) \xrightarrow[T \to \infty]{\text{dist}} \mathcal{N}(0, \sigma^2)$

$$\hat{\mu}_T \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{T}\right)$$



Assume $\theta_1 > \theta_2$



Suffices to have $\hat{\theta}_2 < \theta_2 + \frac{1}{\sqrt{T}}$ and $\hat{\theta}_1 > \theta_1 - \frac{1}{T}$

$\Rightarrow \hat{\theta}_1 > \hat{\theta}_2$ if $\theta_1 - \theta_2 > \frac{2}{\sqrt{T}} \Rightarrow T \gtrsim \frac{1}{\Delta^2}$, $\Delta = \theta_1 - \theta_2$