Optimism in the Face of Uncertainty.

Pull arm i T_i times we know $|\hat{\theta}_i - \theta_i| \le c\sqrt{\frac{1}{T_i}}$

Upper confidence Bound Algorithm (UCB) (2002)

Input: n arms, confidence δ

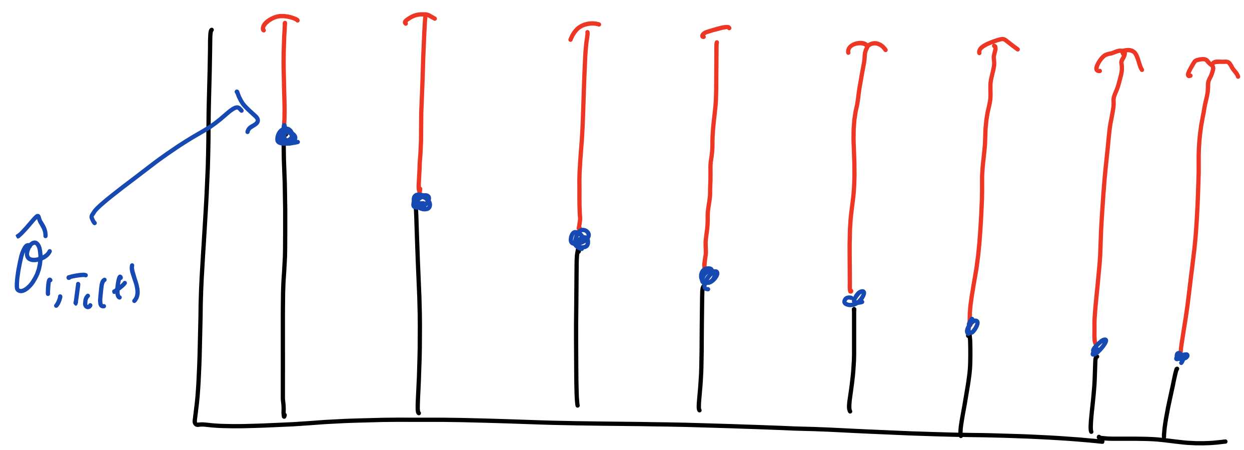
Initialize: Pull every arm once

Let $T_i(t)$ denote # times arm i has been pulled up to time t,

and let $\hat{\theta}_{i,T_i(t)}$ denote empirical mean

Pull arm $I_t = \underset{i \in [n]}{\arg\max}\ \hat{\theta}_{i,T_i(t)} + \sqrt{\frac{2\log(4nT_i(t)^2/\delta)}{T_i(t)}}$

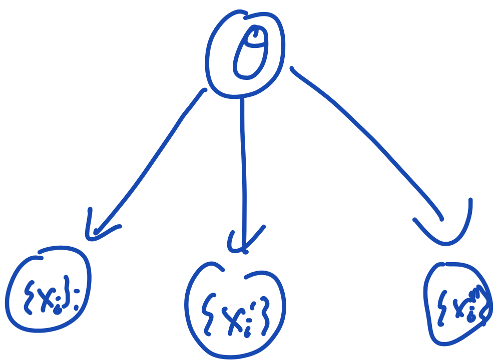
$\boxed{\text{Claim}}$ W.p. $\ge 1-\delta$ $\quad \theta_i \le \hat{\theta}_{i,T_i(t)} + \sqrt{\frac{2\log(4nT_i(t)^2/\delta)}{T_i(t)}} \quad \forall t, i$



Frequentist Statistics

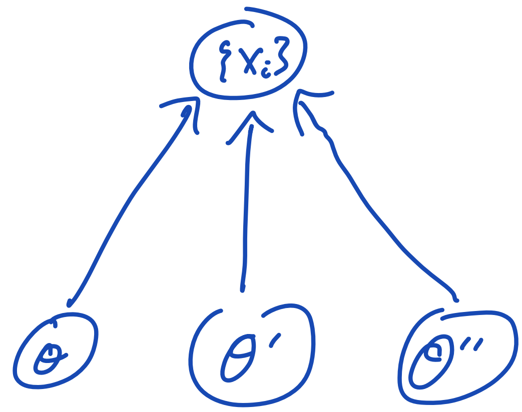
Exists a true parameter θ that is fixed and unknown.

Observe $X_i \overset{iid}{\sim} N(\theta, 1)$



Bayesian Stats

Observe $\{X_i\}_i$ and it is fixed.



Bayesian Statistics

Before seeing any data you have prior "beliefs" about what values θ could take. Prior dist. P_0

Ex. $\theta \sim N(0, \nu^2)$

When I observe $X_i \sim f_\theta(\cdot)$

Ex. $X_i \mid \theta \sim N(\theta, \sigma^2)$

Given prior distribution P_0 and $\{X_i\}_{i=1}^n \overset{iid}{\sim} f_\theta(\cdot)$

Bayes Rule: For any events A, B $\quad \mathbb{P}(B|A) = \frac{\mathbb{P}(A,B)}{\mathbb{P}(A)}$

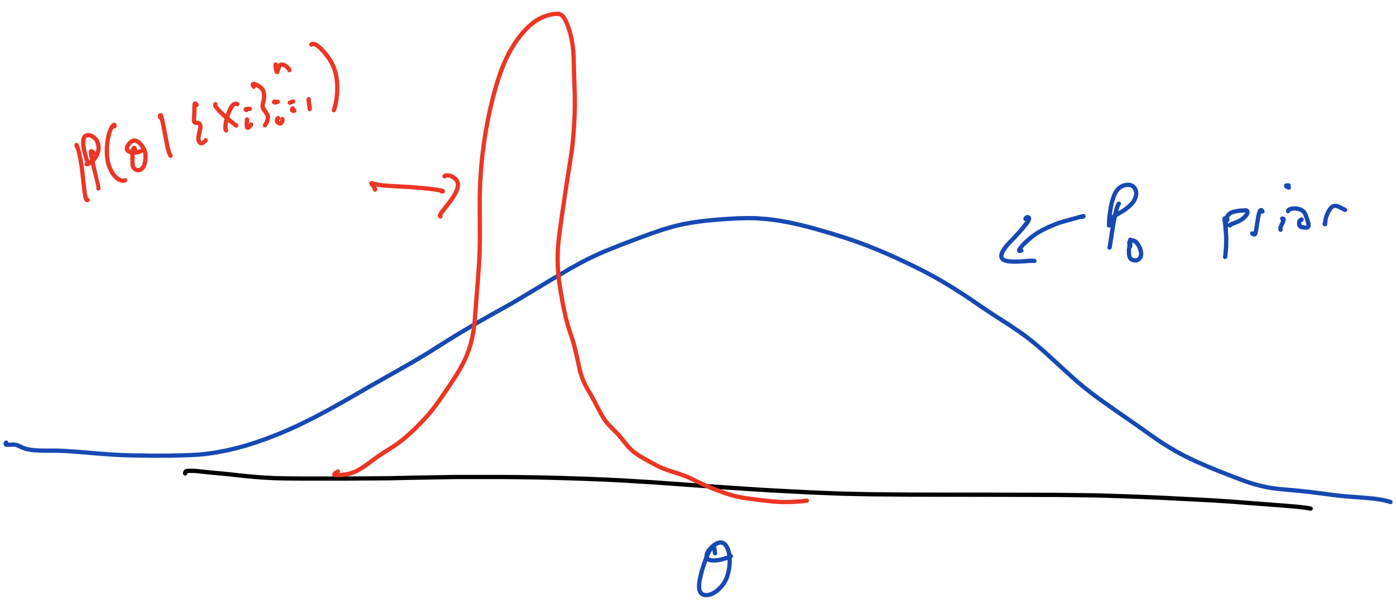
$$\mathbb{P}(A,B) = \mathbb{P}(B|A)\,\mathbb{P}(A)$$

Posterior distribution

$$\mathbb{P}(\theta \mid X_1, \dots, X_n) = \frac{\mathbb{P}(\theta, \{X_i\}_{i=1}^n)}{\mathbb{P}(\{X_i\}_{i=1}^n)}$$

$$= \frac{\mathbb{P}(\{X\}_{i=1}^n \mid \theta)\,\mathbb{P}(\theta)}{\mathbb{P}(\{X_i\}_{i=1}^n)}$$

$$= \frac{\prod_{i=1}^n f_\theta(X_i) \quad P_0(\theta)}{\mathbb{P}(\{X_i\}_{i=1}^n)}$$



$\theta \sim P_0$

$X_i \sim f_\theta \quad i=1,\dots,n$

Thompson Sampling (1930)

$\theta \in \mathbb{R}^n$, n arms

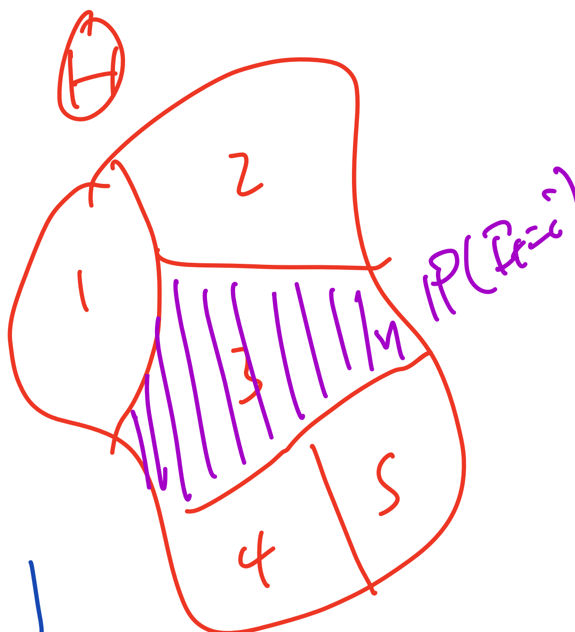
Initialize: $\theta \sim P_0$

for $t=1,2,\dots$

$$P_{t,i} = \mathbb{P}\left(i = \arg\max_j \theta_j \,\middle|\, \{X_{i,T_i(t)}\}_{i,t}\right)$$

$$\left(\sum_{i=1}^{n} P_{t,i} = 1\right)$$

$$I_t \sim P_{t,i}$$



$$\mathbb{P}\left(i = \arg\max_j \theta_j \,\middle|\, \{X_{i,T_i(t)}\}_{i,t}\right)$$

$$= \int_\theta \mathbb{1}\{\theta_i \geq \max_j \theta_j\}\, \mathbb{P}\left(\theta \,\middle|\, \{X_{i,T_i(t)}\}_{i,t}\right) d\theta =$$

for $t = 1, 2, \ldots$

$$\theta_t \sim \mathbb{P}(\theta \mid \{X_{i, T_i(n)}\}_{i,t})$$

Play $I_t = \arg\max_i \theta_{t,i}$

$$\mathbb{P}(I_t = i \mid \{X_{i,t}\}) = \int_\theta \mathbb{P}(I_t = i, \theta \mid \{X_{i,t}\}) d\theta$$

$$= \int_\theta \mathbb{1}\{I_t = i\} \mathbb{P}(\theta \mid \{X_{i,t}\}) d\theta$$

Frequentist regret: Think of TS as an algorithm where the choice of prior P_0, likelihood function f_θ are design choices. There exists a fixed θ^* and regret is

$$\sum_{i>1}^{n} \underbrace{(\theta_1^* - \theta_i^*)}_{\Delta_i} \mathbb{E}[T_i] \lesssim \min\left\{\sqrt{nT}, \sum_{i>1} \Delta_i^{-1} \log(T)\right\}$$

Bayesian regret: Assume "true" $\theta^* \sim P_0$.

And $f_\theta(x)$ is the "true" likelihood of observing $x \in \mathbb{R}$ given fixed θ. Then regret is

$$\mathbb{E}_{\theta^* \sim P_0}\left[\sum_{i \neq \arg\max_j \theta_j^*} \left(\max_{i'} \theta_{i'}^* - \theta_i\right) \mathbb{E}[T_i]\right].$$

$$\leq \sqrt{nT \cdot \underbrace{H(P_0)}}$$

entropy of the prior distribution assigning each arm to be the best arm.