Bayesian Inference: Towards MAB

$$\mathbb{P}(\theta \mid \text{data}) = \frac{\mathbb{P}(\text{data} \mid \theta) \times \mathbb{P}(\theta)}{\mathbb{P}(\text{data})}$$

- Suppose that you want to estimate the bias θ (= probability of getting heads) of a coin using coin flips.
- Let the prior = Uniform distribution([0,1]) = beta($\alpha = 1, \beta = 1$)
- We interpret α = #tails and β = #heads

Suppose at time n,

α_n = #tails and β_n = #heads such that $\alpha_n + \beta_n$ = n.

Then if at time n+1, you get a

head then posterior = beta($\alpha_n, \beta_n + 1$)

tail then posterior = beta($\alpha_n + 1, \beta_n$)

Thompson Sampling

- Maintain posteriors for different arms.

- Let $\alpha_a(n)$ and $\beta_a(n)$ denote the number of tails and heads of arm a resp. at time n. Then beta$(\alpha_a(n), \beta_a(n))$ represent "belief" about the true bias of arm a.

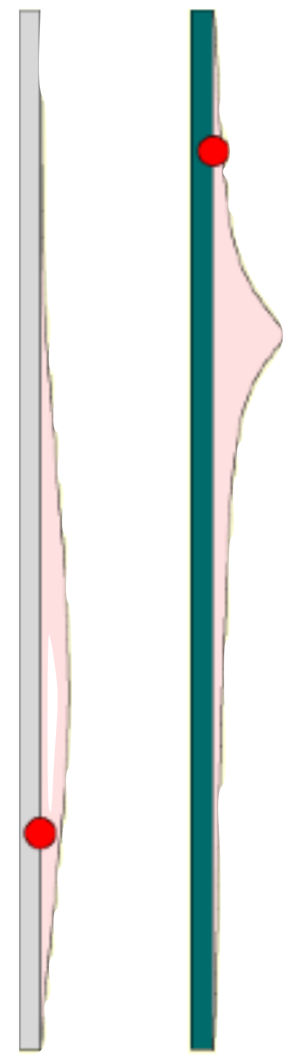
At time n,

Computational step:

Sample each arm $x_a(n) \sim$ beta$(\alpha_a(n), \beta_a(n))$

Sampling step:

Pull arm $a(n) = \arg\max_{a \in [K]} x_a(n)$



- Achieves optimal regret (Kaufmann et al., 2012); is excellent in practice (Chapelle and Li, 2011)