

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 $\text{beta}(\alpha_a(n), \beta_a(n))$ represent “belief” about the true bias of arm a .

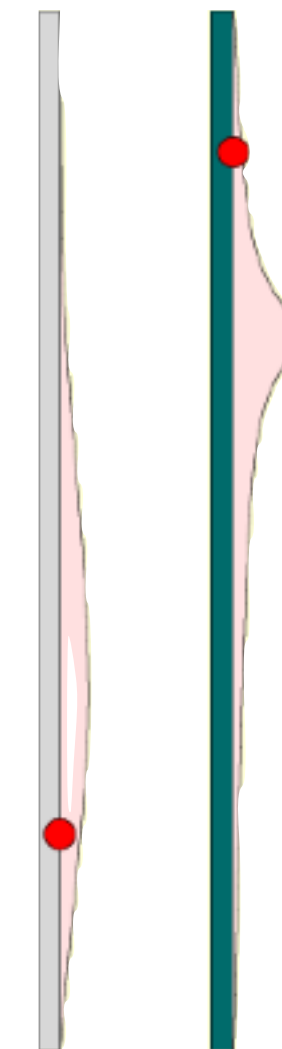
At time n ,

Computational step:

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

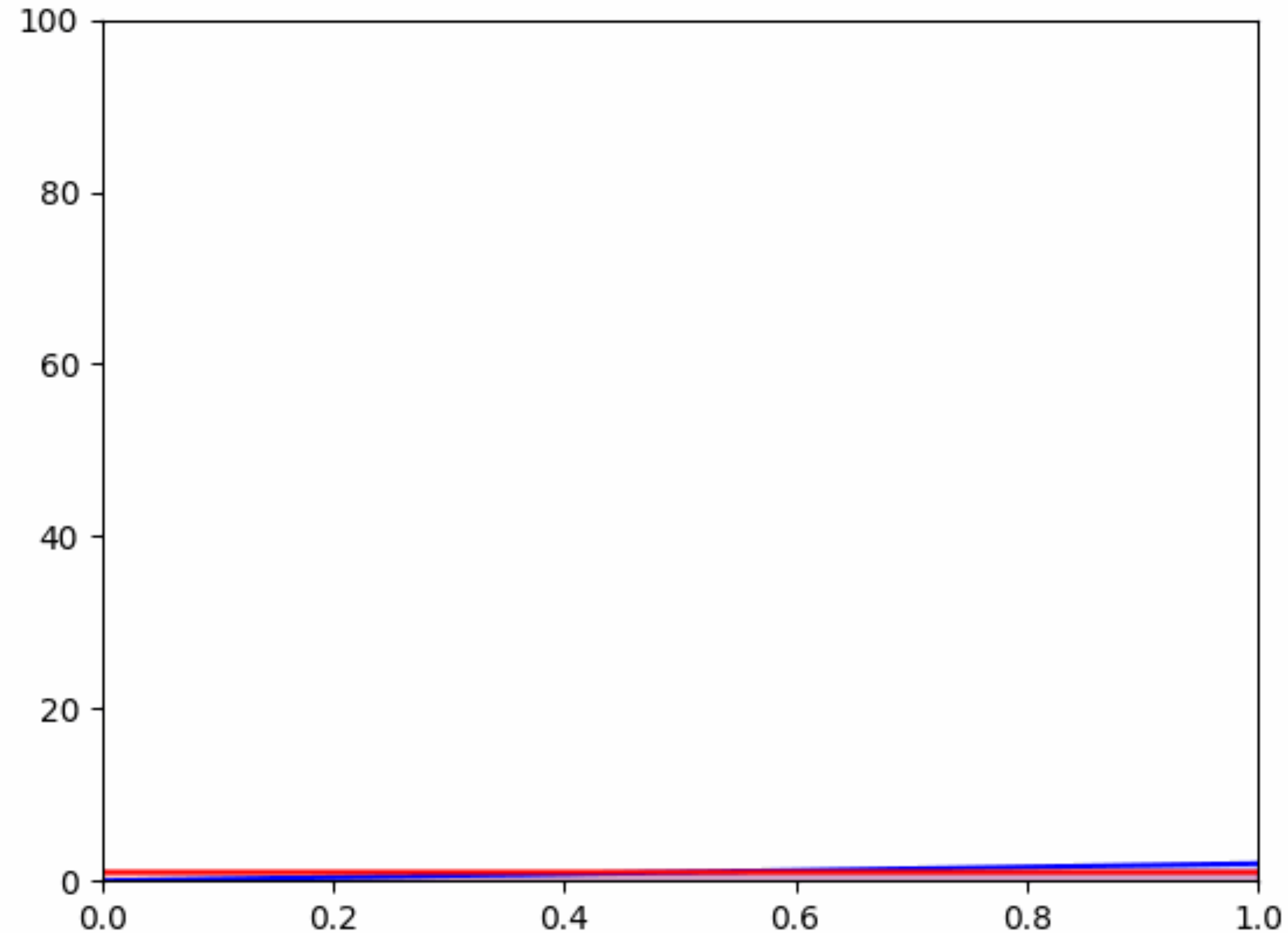
Sampling step:

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)

Thomson Sampling Working



Thompson sampling working for two coins
with probability of head 0.9 (red) and 0.5
(blue)