

Explore then commit

- Uniform sampling doesn't use the information collected from the observed samples, and just dumbly explores over all the arms !!
- Explore then commit:
 - Do uniform sampling over first m iterations ($m \leq T$)
 - Sample from the empirically best arm over the next $T - m$ iterations.

Explore then commit when gap is known

- If the minimum gap $\Delta_{\min} = \min_{a \neq a^*} \Delta_a$ is known, then regret suffered by explore then commit is

$$\text{Reg}_T \leq m + 2T \cdot \exp(-m\Delta_{\min}^2/K)$$

- Optimal amount of exploration minimizing the regret is

$$m^* = O\left(\frac{K \log(T)}{\Delta_{\min}^2}\right)$$

and the corresponding regret is $\text{Reg}_T = O\left(\frac{K \log(T)}{\Delta_{\min}^2}\right)$.

- To **separate** two coins with probability of heads 0.4 and 0.6 ($\Delta_{\min} = 0.2$) within 1000 tosses ($T = 1000$) efficiently, we need at least 576 tosses for exploration