Explore then commit when gap is known

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

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

- **Optimal amount of exploration** minimizing the regret is

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

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

- 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

Limitations of explore then commit (contd.)

- **Question:** What if we don't know the minimum gap $\Delta_{\min}$?
- If $\Delta_{\min} \to 0$, $\quad m^\star = O\left(\frac{K\log(T)}{\Delta_{\min}}\right) \to \infty$!!!
- We need infinitely many samples to seperate arbitrarily close arms !!!
- Trick: For arms which are arbitarily close ($\Delta_a = O(1/T)$), we will suffer constant regret for pulling them. We can ignore such arms !!

Theorem: Upon having $m = O(N^{2/3}\log(N)^{1/3})$ explore then commit suffers a regret of at most $O(N^{2/3}(K\log(N))^{1/3})$ over all instances !!!

Question: Can there be an algorithm that learns the gap $\Delta_{\min}$ and **performs the optimal amount of exploration.**