How to get sub-linear regret ?

- **Explore sub-optimal arms:** We need to sample every sub-optimal arms infinitely often as budget $T \to \infty$, to be certain about their sub-optimality.

- What if we sample some sub-optimal arm finite no. of times? The collected samples can be large with positive probability and the arm may appear good !!

- **Exploit the best arm:** Our algorithm must satisfy: $\lim_{T \to \infty} \frac{\mathbb{E}[N_{a^\star}(T)]}{T} = 1.$

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.