Best-Arm Identification

- Treat each batsmen as an arm a, such that their average performance (hit rate) is given by μ_a
- At every round t,
  - Choose an arm (batsman) a_t and play a ball
  - Observe their performance in terms of reward $r_t \sim D(\mu_{a_t})$
  - If you are confident at round τ_δ, announce the best batsman $\hat{a}_t$, else continue
- Objective: Minimise the number of samples (balls) $\mathbb{E}[\tau_\delta]$ such that the error is small i.e. $\mathbb{P}(\hat{a}_{\tau_\delta} \neq \arg\max_a \mu_a) \leq \delta$

Recall from Hoeffding's inequality

The empirical mean $\bar{X}(t) \in [\mu_1 \pm \alpha_t]$ **where** $\alpha_t = \sqrt{\frac{4\log(Kt/\delta)}{t}}$

with probability atleast $1 - \delta$