**Explore-Exploit**

1. **First N Rounds** try all doors, uniformly.
   - Switching to pick poor with highest avg reward.
   - "Explore - Then - Exploit".

2. **E-prob Explore** "Pick a Poor Random".
   - 1-E prob **Exploit** "Greedy".
   - (E-Greedy)

3. **Thompson Sampling**
   - Maintain a distribution $P(door)$ of how good door is.
   - Sample door from this dist.
   - Update posterior.

4. **UCB**: Choose door with highest upper confidence bound.