Announcements

- Office hours
  - Tue March 5, 4-5pm in room 3234
  - Thu March 7:
    - online, 5.30-7pm:
    - write me an email for appointment: gquer@ucsd.edu

- Thu March 7: NO LECTURE

- **Readings**: 09_ABtesting/readings.md **due on Friday 3/8 at 6PM**

A/B test: binary outcomes

- Email 1: $n_1 = 605$, clicks: $c_1 = 351$
- Email 2: $n_2 = 585$, clicks: $c_2 = 123$
- Click per email: $p_1 = 0.58$, $p_2 = 0.21$
- Is there **enough evidence** that Email 1 is better than email 2?
- Numbers are large (>100) so we can approximate with a Gaussian
- The null hypothesis is $p_1 = p_2$, we can calculate
- $p = (c_1 + c_2) / (n_1 + n_2)$: the mean click rate in the null hypothesis
- $\sigma^2 = p(1-p)$: the variance of the outcome
-
- If t>1.96, they are actually different (with 95% confidence)

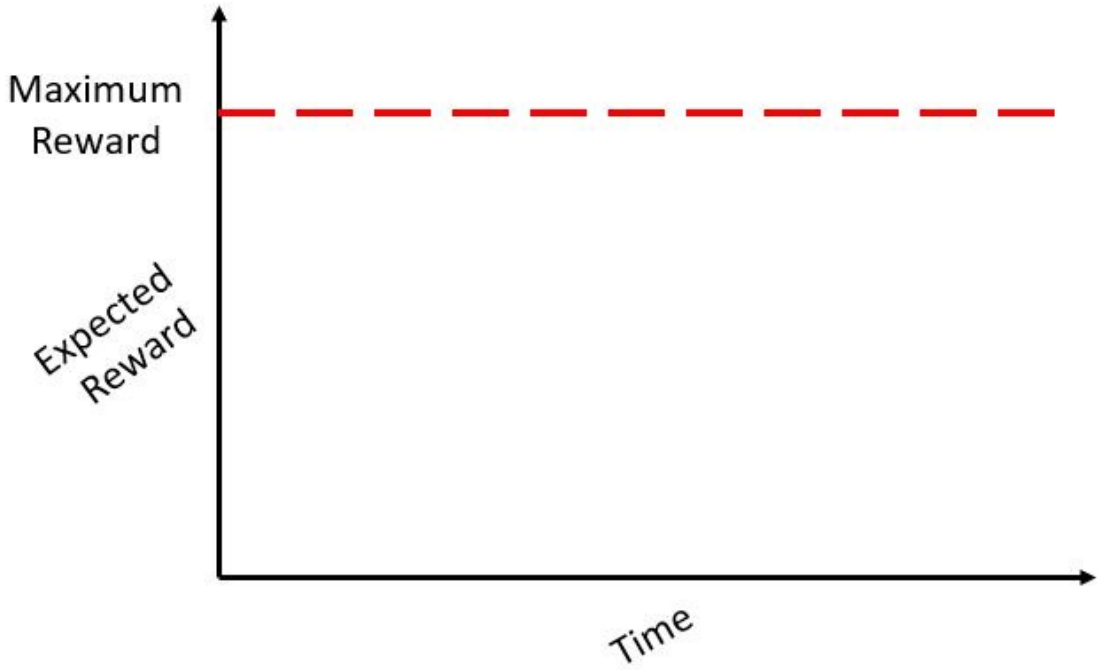
$$t = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\sigma^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

Common Pitfalls of A/B Tests

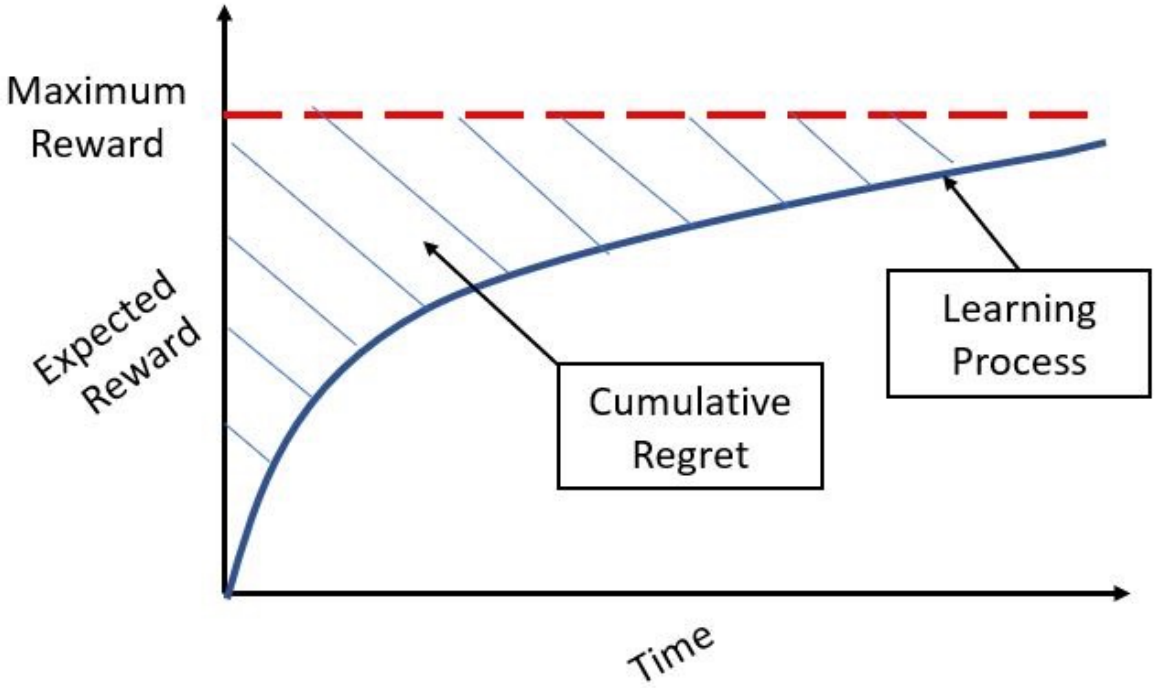
1. Optimizing for the wrong metric
2. Failing to correctly randomize
3. Stopping early
4. Unethical testing
5. Non-stationary problem

Multi-Armed Bandits

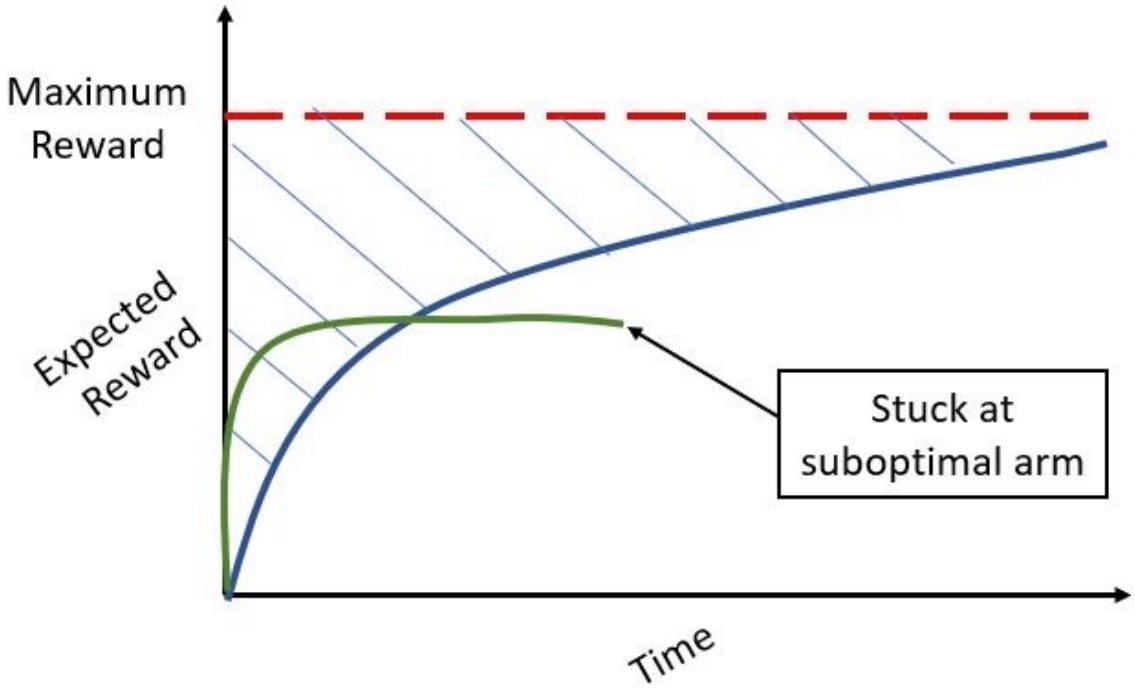
Perfect Knowledge



Exploring causes regret



Not exploring enough causes more regret



Explore then greedy

Hypothesis testing approach:

1. Run each experiment *N* times.
2. Choose a winner, and use that forever.

How to determine *N*?

What might go wrong with this approach?

- *N* too small and you might choose the wrong winner
- *N* too large and you spend too long experimenting

Upper Confidence Bound

j=1,...,K possible actions

Each action has a stationary random reward between 0 and 1

Choose the action that maximizes:

$$m_j + \sqrt{\frac{2 \ln N}{N_j}}$$

- m_j = the current average reward of arm *j*
- N = the number of pulls, and
- N_j = the number of pulls on lever *j*

Other Approaches

- Epsilon Greedy
- Softmax Exploration
- Decayed Epsilon Greedy
- Thompson Sampling