Module 2:
A/B Testing I: Measurement

DAV-6300-1: Experimental Optimization

David Sweet // 20240905

Coin Flipping I

Coin Flipping Ia

Key Terms

- Expectation
- Law of Large Numbers
- Central Limit Theorem
- Mean
- Standard Deviation
- Standard Error

Expectation

- Expectation is asymptotic value of running mean
- Expectation is 0
  - No flip is ever 0
  - Only observe 1, -1
- Expectation is unobservable

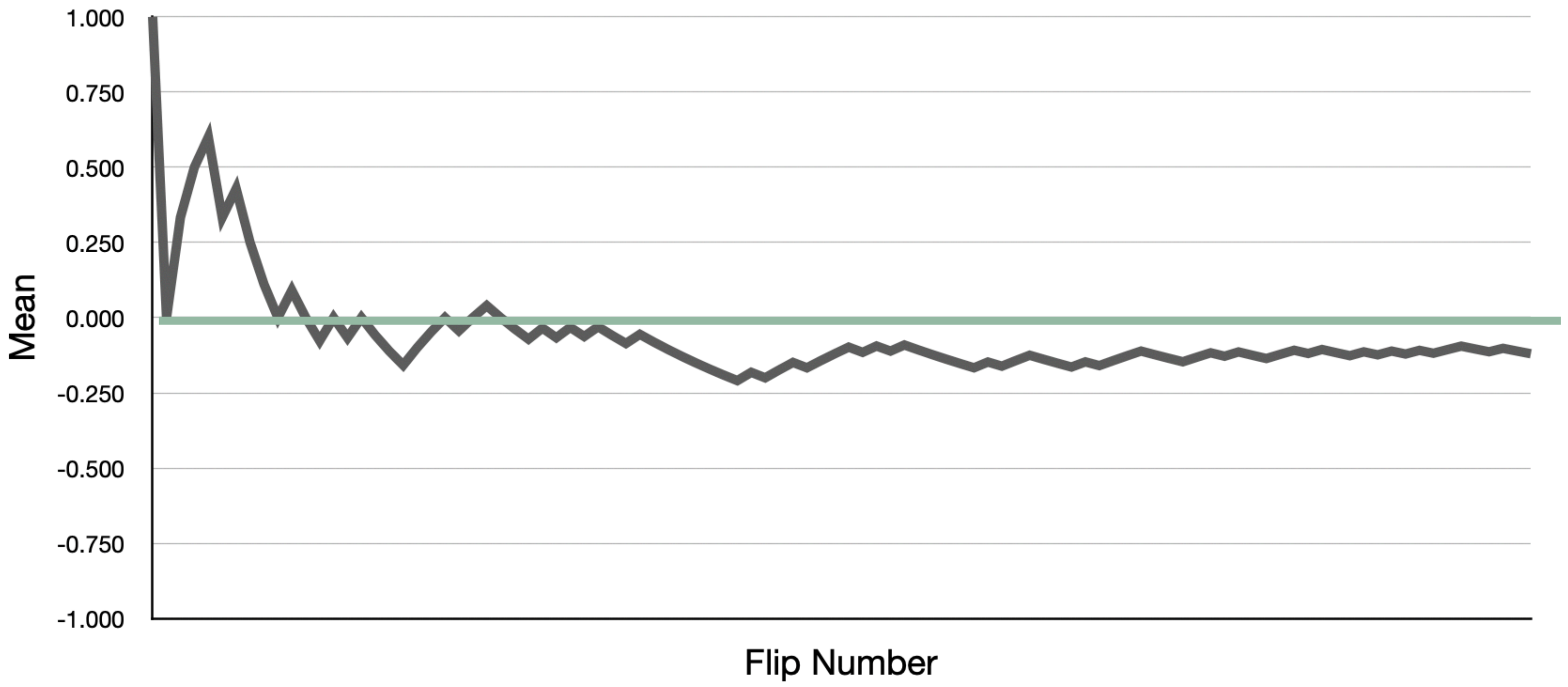
Notation

- observation: y
- observations, indexed: $y_1, y_2, y_3, \ldots$
- mean: $\mu = \frac{\sum_i^N y_i}{N}$
- expectation: $E[y]$
- Coin flipping: $y \in \{-1,1\}$, $E[y] = 0$

Law of Large Numbers

- The more observations we include in the mean, the closer the mean gets to the expectation.

As $N \to \infty$, $\mu \to E[y]$



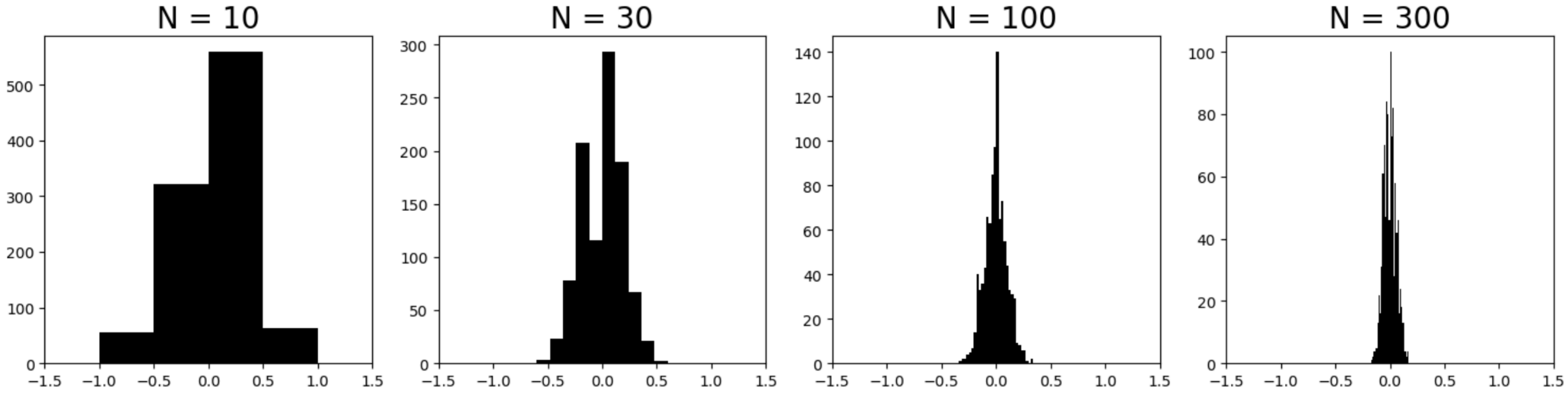
Expectation

- Measurement estimates expectation by mean (μ)
- Can't know expectation b/c N finite
- Aside: Experimentation cost ==> "Smaller N is better"
  - But LLN ==> "Larger N is better"
    - How much better?

Coin Flipping II

Central Limit Theorem

- Mean is normally distributed
- Variance of mean decreases with N



Notation

- variance of y: $VAR[y]$ ⟵ Also unobservable
- sample variance of y: $\sigma^2 = \frac{\sum_i^N (y_i - \mu)^2}{N}$ ⟵ ...but we estimate
- normal distribution: $\mathcal{N}(E[\cdot], VAR[\cdot])$
- standard deviation: $STD[y] = \sqrt{VAR[y]}$
- sample standard deviation: $\sigma = \sqrt{\sigma^2}$

Central Limit Theorem

- CLT: $\mu \sim \mathcal{N}(E[y], VAR[y]/N)$
- standard deviation of μ: $STD[\mu] = STD[y]/\sqrt{N}$
- or: $\dfrac{\mu - E[y]}{STD[\mu]} \sim \mathcal{N}(0,1)$ $\longleftarrow$ $E[y]$ and $STD[\mu]$ unobservable

Central Limit Theorem

- Can't observe, so estimate:
    - Estimate $E[y]$ by μ
    - Estimate $VAR[y]$ by σ^2
    - Estimate $SE[\mu]$ by $se = \sigma/\sqrt{N}$

Measurement

- Want to know $E[y]$ (but can't)
- Observe: $y_1, y_2, y_3, \ldots$
- Estimate

"Your estimate is this close to $E[y]$."

$$E[y] \quad \text{by} \quad \mu = \frac{\sum_i^N y_i}{N}$$

$$SE[\mu] \quad \text{by} \quad se = \left\{ \sqrt{\frac{\sum_i^N (y_i - \mu)^2}{N}} \right\} / \sqrt{N}$$

Measurement

$$VAR[aX] = a^2 VAR[X]$$

$$VAR[\mu] = VAR[\frac{\sum_i^N y_i}{N}] = \frac{1}{N^2}\sum_i^N VAR[y]$$
$$= \frac{1}{N^2} N \times VAR[y] = VAR[y]/N$$
$$SD[\mu] = \sqrt{VAR[\mu]} = \sqrt{VAR[y]/N} = SD[y]/\sqrt{N}$$

- Estimate with samples:

$$se = \sigma/\sqrt{N}$$

Measurement

- $\dfrac{\mu - E[y]}{STD[\mu]} \sim \mathcal{N}(0,1)$
- Bias / “inaccuracy”: $\mu - E[y]$
  - Can’t estimate
- Variance / “imprecision”: $STD[u]$
  - Can estimate: se

Experimental methods control
bias & variance,
accuracy & precision.

Readings for Week 3

- Chapter 2 of Experimentation for Engineers
  A/B testing: Evaluating a modification to your system
- A Refresher on A/B Testing
  https://hbr.org/2017/06/a-refresher-on-ab-testing
- Catalog of Biases
  https://catalogofbias.org/biases/
- Accuracy vs Precision: Differences & Examples
  https://statisticsbyjim.com/basics/accuracy-vs-precision/