FIN 5330 02.09. TU

Today's Agenda

- Thinking Statistically { 2 kinds of Prob. ① classical / frequentist ② subjectivist / Bayesian
- Review Bayes Rule (in brief)
- Review sampling distribution ✓ (LLN & CLT)
- Review some Julia bits (if time)

Monte Carlo Simulation

Key methodology

Bayes' Rule

- A probability statement a/b conditional prob's

- We're interest in some event A
- We observe the outcome of some event B

* Glossary
- Random variable
- observations (data)

Posterior

$$P(A \mid B) = \frac{P(B \mid A)\,[P(A)]}{P(B)}$$

- $P(A \mid B)$: conditional prob.
  - A: random variable of interest
  - B: observation
- $P(B \mid A)$: likelihood function
- $P(A)$: Prior (1)
- $P(B)$: marginal of B

Bayes Rule = updating beliefs in the light of evidence

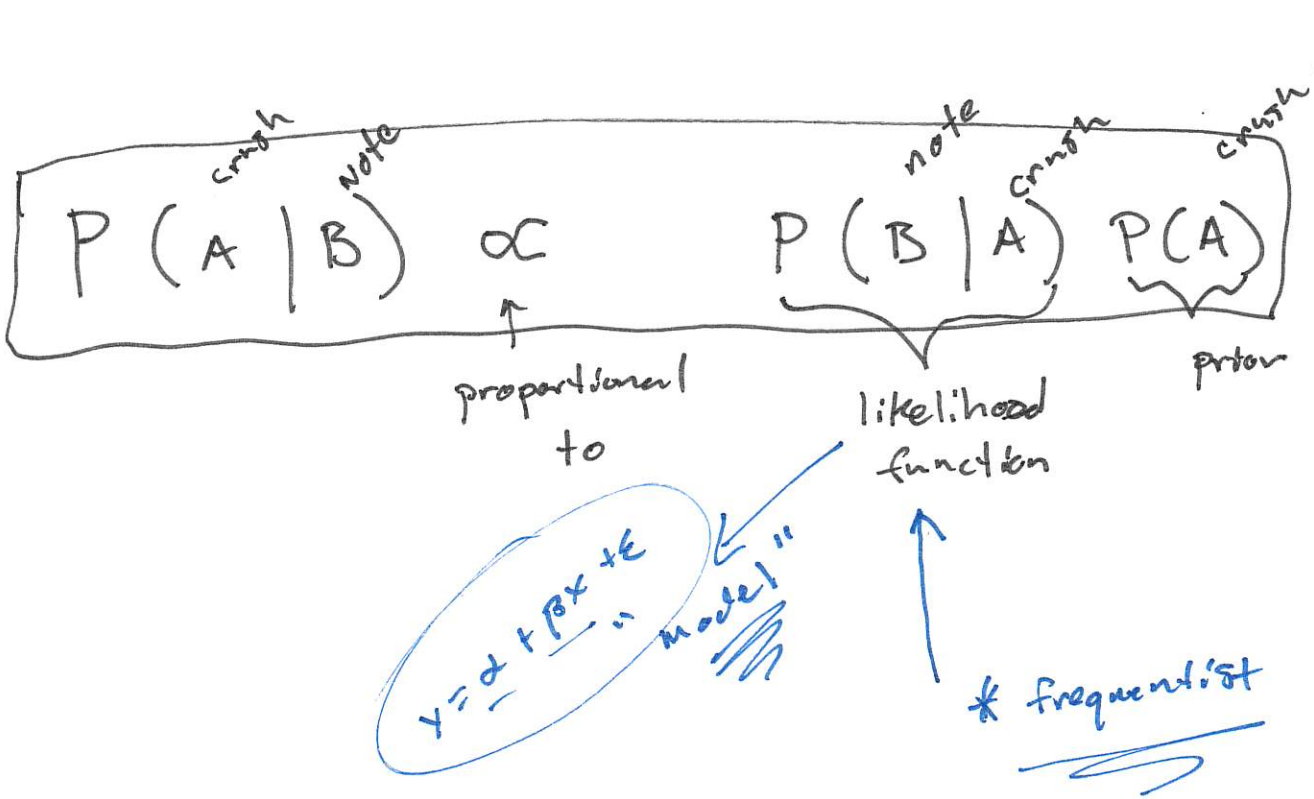
Prior: a statement a/b our beliefs before we see the data

Posterior: updated belief after we see the data

- online algorithm : optimal for sequential data process

- Information processing / updating beliefs

- Understanding Bayes' rule is what it means to "think statistically"



- The Frequentist

  physical property of the coin

  - $P(\theta)$ = prob. heads

$$\hat{\theta} = \frac{\#\text{ heads}}{\#\text{ flips}}$$ MLE

- Population: object of interest: PDF $f(x; \mu, \sigma)$

  prob. dist.

  $N(\mu, \sigma)$

  "Truth"

  Platonic

- Sample: data that results from an experiment

- Sampling Distribution

- Law of Large Numbers (LLN)
- Central Limit ~~Theorem~~ theorem (CLT)

  } Intro stats classes

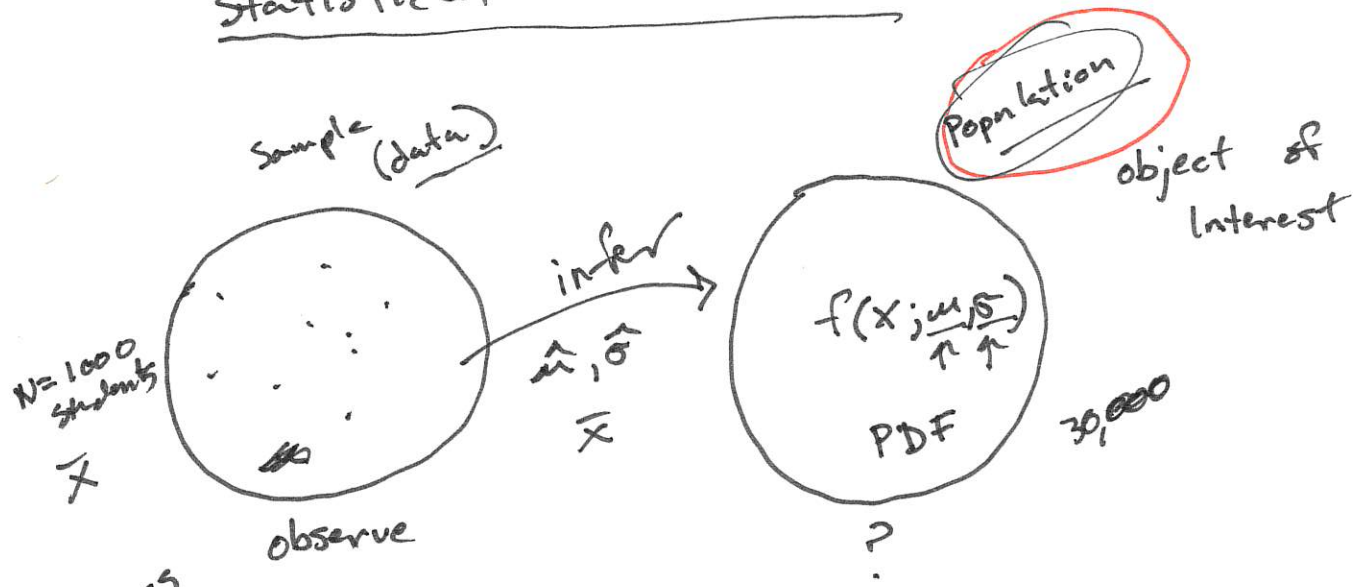
  ↳ ?

- As $N \to \infty$, our estimator $\overline{x}$ gets closer and closer to a Normal distribution for its sampling distribution

- Q: What is the Sampling Distribution?
  - The statistical distribution of an estimator

- Q: Estimator:

Statistical Inference.



Biasedness

$$\bar{X} = \frac{1}{n}\sum x_i$$

Sample mean

$E(\bar{x}) = X$

$E(\hat{\beta}) = \beta$

? → statement the sampling dist of $\bar{X}$

Sampling dist ↔ likelihood

Visualize the Frequentist Prob.

Multiverse
?
parallel universe

Population
(μ, σ or θ)
probability density

? inference about this

Sampling process

Population ≈ Data Generating Process

Universe $j=1$

Universe $j=2$

Universe $j=3$

Universe $j=M=100,000$

Bootstrap

obs $i=1 \ldots n$

Sample 1 $\bar{x}, s, m$ → $j=1$

Sample 2 → $j=2$

Sample 3 → $j=3$

Sample M → $j=M$

Yield

$\bar{x}_1$ = sample mean in universe 1 from sample 1

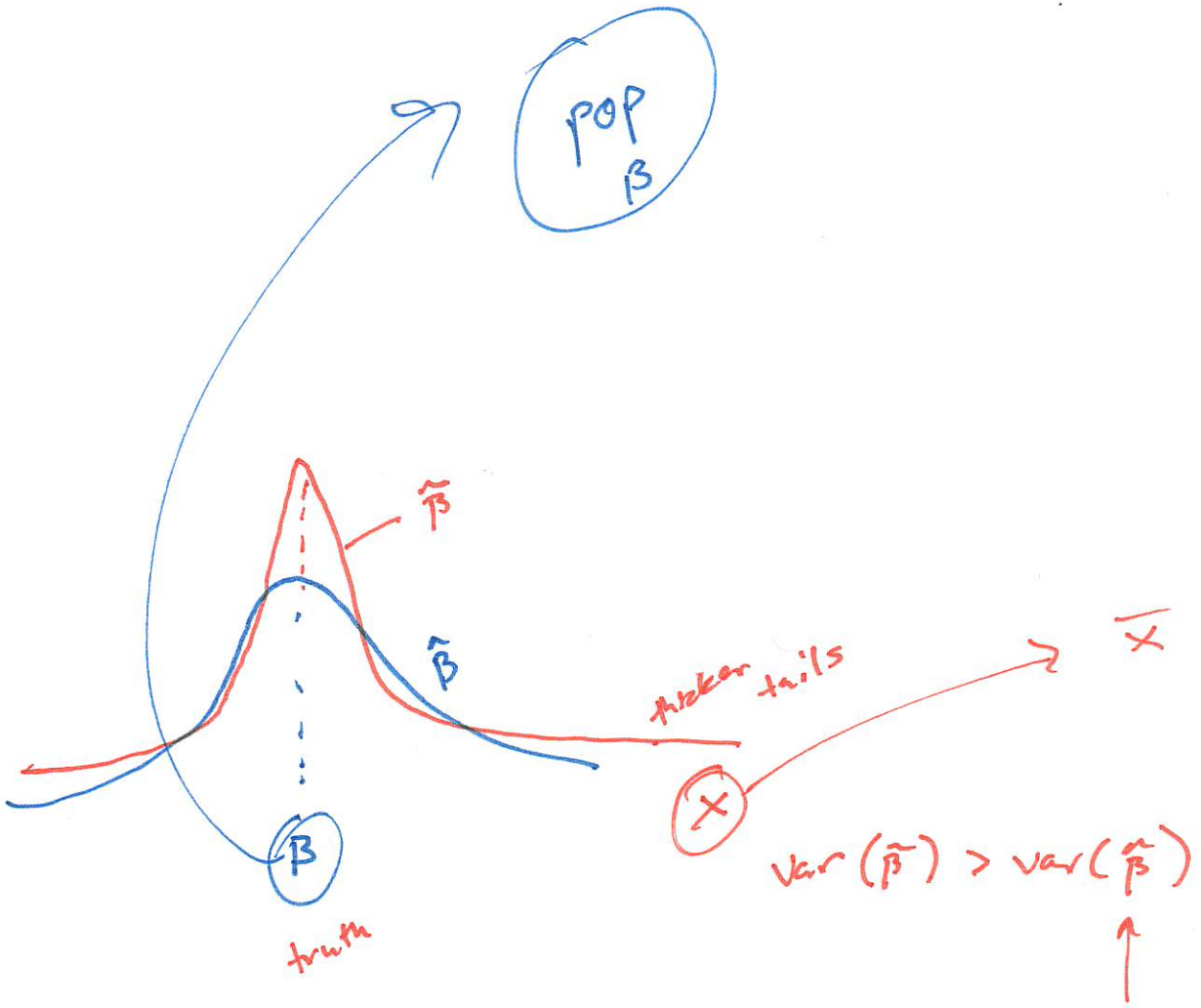
$\bar{x}_2$

$\bar{x}_3$

$\bar{x}_M$

CLT

$E(\bar{x}) = \mu$ unbiased


pop
β
β̃
β̂
thicker tails
x̄
x
β
truth
Var(β̃) > Var(β̂)