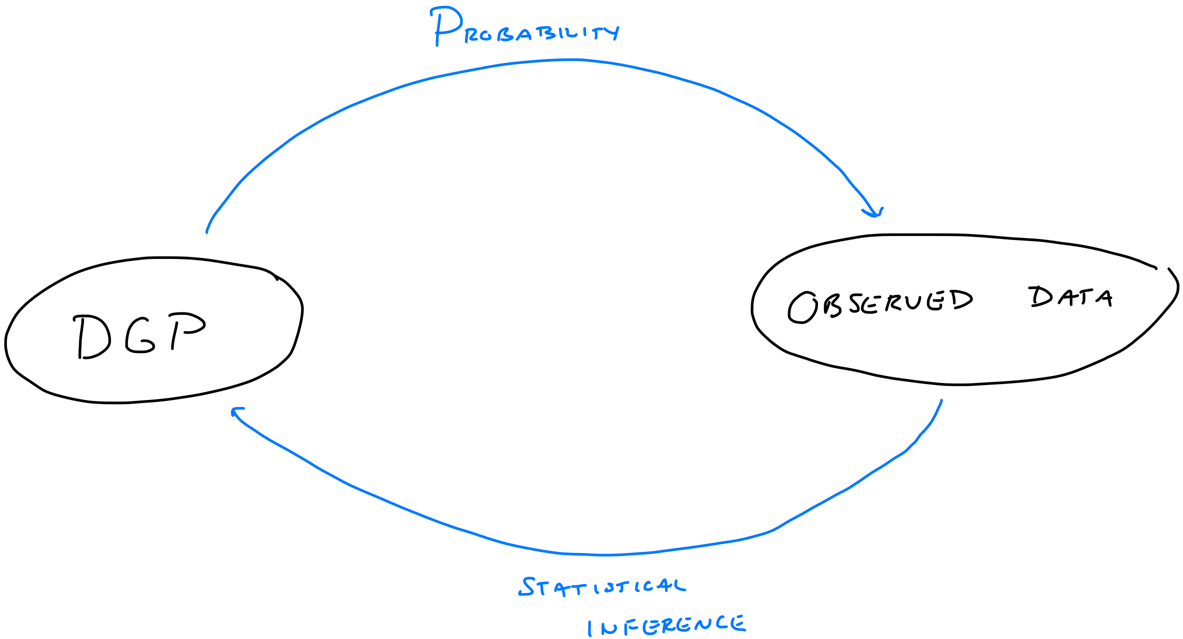
STATISTICAL INFERENCE

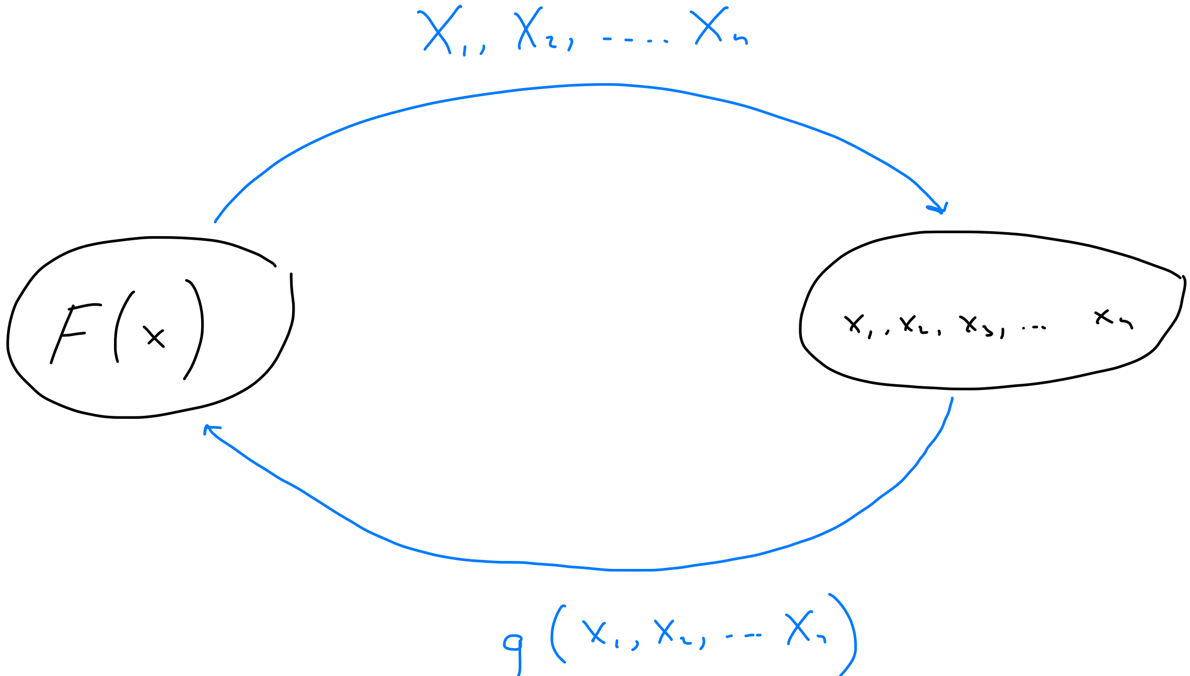
STAT 510

$$X_1, X_2, X_3, \ldots\ldots X_n \sim F$$

How do we "INFER" F?

LEARN


PROBABILITY
DGP
OBSERVED DATA
STATISTICAL INFERENCE



STATISTIC

↳ A FUNCTION OF DATA

↳ CAN BE A RANDOM VARIABLE

$X_1, X_2, \ldots \ldots X_n \sim F$

- HOW TO ESTIMATE F?

STATISTICAL MODEL

- PARAMETRIC ESTIMATION

- NONPARAMETRIC ESTIMATION

$X_1, X_2, \ldots X_n \sim F$

PARAMETRIC MODEL

STATISTICAL

$$\mathcal{F} = \{ f(x;\theta) : \theta \in \Theta \}$$

DENSITY

PARAMETER THAT DETERMINES DENSITY

PARAMETER SPACE

$X_1, X_2, \ldots X_n \sim F$

Assume $f \in \mathcal{F}$ where

$$\mathcal{F} = \left\{ f(x;p) = p^x (1-p)^{x-1} , \quad 0 < p < 1 \right\}$$

$X_1, X_2, \ldots X_n \sim \text{Bernoulli}(p)$

PROBLEM REDUCED TO ESTIMATING p

$$X_1, X_2, \ldots X_n \sim \text{Normal}(\mu, \sigma^2)$$

Assume $f \in \mathcal{F}$ where

$$\mathcal{F} = \left\{ f(x; \mu, \sigma^2) : \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{ -\frac{1}{2\sigma^2}(x-\mu)^2 \right\},\ \overbrace{\mu \in \mathbb{R},\ \sigma > 0}^{\text{Parameter space}} \right\}$$

If parameter of interest

Nuisance parameter

$$X_1, X_2, \ldots X_n \sim N(\mu, \sigma^2)$$

$X_1, X_2, \ldots\ldots X_n \sim F$

Assume $F \in \mathcal{F}_{ALL} = \{\text{ALL CDFs}\}$

NONPARAMETRIC

A NOTE ABOUT

✓ FREQUENTIST INFERENCE

VS

BAYESIAN INFERENCE

Inference Concepts

- Point Estimation
- Interval Estimation
- Hypothesis Testing

Point Estimation

Parameter of interest

Use $\hat{\theta}$ as "best guess" for θ.

Fixed, unknown

Given $X_1, X_2, \ldots X_n \sim F$

Point estimator

$$\hat{\theta}_n = g(\underbrace{X_1, X_2, \ldots X_n}_{\text{Data}})$$

Function

Random Variable

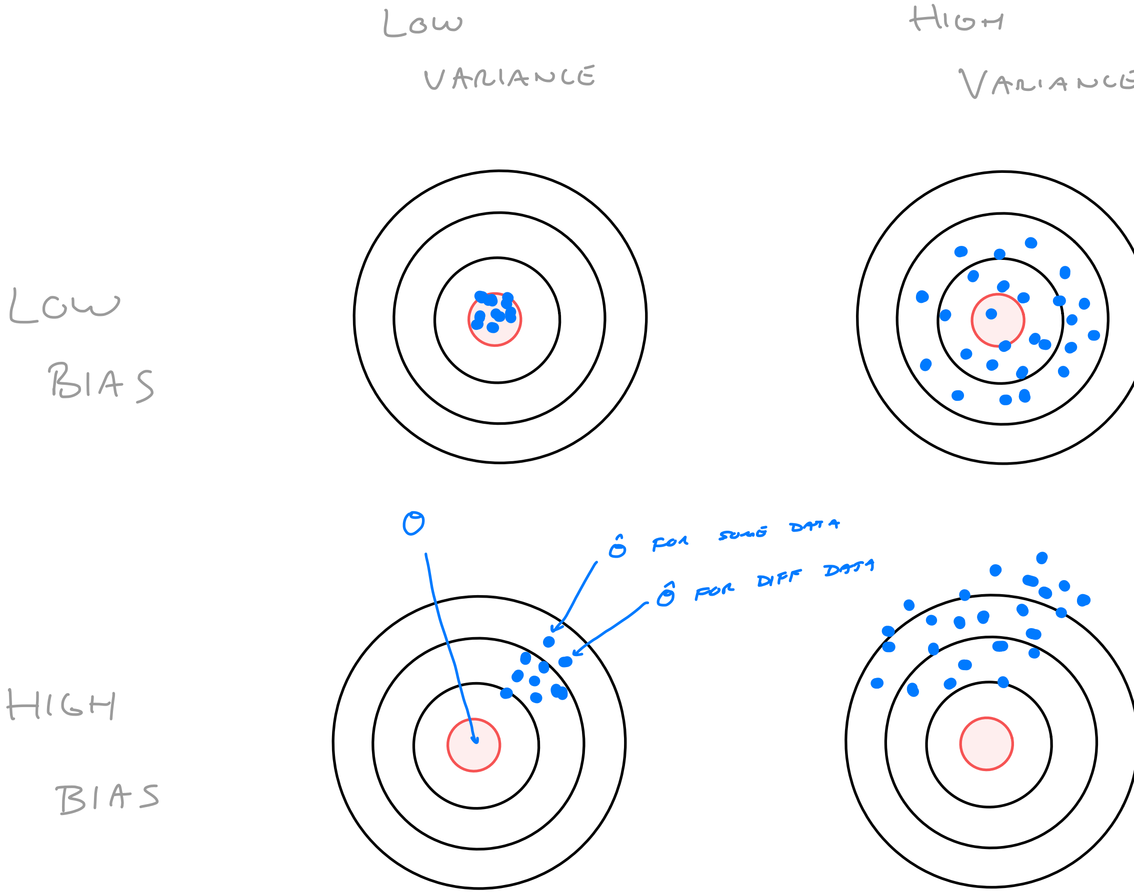
Sampling distribution...

Defn

$$\mathrm{BIAS}[\hat{\theta}] = \mathbb{E}[\hat{\theta}] - \theta$$

$$\mathrm{VAR}[\hat{\theta}] = \mathbb{E}\left[\left(\hat{\theta} - E[\bar{\theta}]\right)^2\right]$$

$$\mathrm{MSE}[\bar{\theta}] = \mathbb{E}\left[(\hat{\theta} - \theta)^2\right] = \left(\mathrm{BIAS}[\hat{\theta}]\right)^2 + \mathbb{V}[\hat{\theta}]$$


Low Variance
High Variance
Low Bias
High Bias
θ
θ̂ for same data
θ̂ for diff data

DEFN

A POINT ESTIMATOR $\hat{\Theta}_n$ OF A PARAMETER Θ

IS CONSISTENT IF

$$\hat{\Theta}_n \xrightarrow{P} \Theta$$

ESTIMATOR GETS "BETTER" WITH MORE DATA!

THM

IF BIAS $\to 0$

AND VAR $\to 0$

AS $n \to \infty$

THEN $\hat{\theta}_n$ IS CONSISTENT.

PROOF

IF BIAS AND VARIANCE GO TO ZERO, THEN

MSE $\to 0$

THUS $\hat{\theta}_n \xrightarrow{qm} \theta$, SO $\hat{\theta}_n \xrightarrow{P} \theta$.

A NOTE ABOUT "STANDARD ERROR"

$$se = se(\hat{\theta}_n) = \sqrt{\mathbb{V}[\hat{\theta}_n]} = sd(\hat{\theta}_n)$$

$$\hat{se} = \hat{se}(\hat{\theta}_n) = \sqrt{\hat{\mathbb{V}}[\hat{\theta}_n]} = \hat{sd}(\hat{\theta}_n)$$

EXAMPLE

$X_1, X_2, \ldots\ldots, X_n \sim \text{BERNOULLI}(p)$

$$\hat{p} = \frac{1}{n} \sum X_i$$

$$\mathbb{E}[\hat{p}] = \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}[x_i] = p \implies \hat{p} \text{ UNBIASED}$$

$$\mathbb{V}[\hat{p}] = \frac{p(1-p)}{n} \qquad \hat{sd}(\hat{p}) = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

ESTIMATED STANDARD ERROR

SINCE $\hat{p}$ UNBIASED AND

$\mathbb{V}[\hat{p}] \to 0$ THEN $\hat{p}$ IS CONSISTENT.

AS $n \to \infty$

DEFN

AN ESTIMATOR IS

ASYMPTOTICALLY NORMAL

IF

$$\frac{\hat{\theta} - \theta}{sd(\hat{\theta})} \xrightarrow{d} N(0,1)$$

Confidence Intervals

NOT RANDOM

$1-\alpha$ CI FOR θ IS AN INTERVAL

$$C_n = (a, b)$$

Random! $\longrightarrow$ $a = a(X_1, X_2, \dots X_n)$

Random! $\longrightarrow$ $b = b(X_1, X_2, \dots X_n)$

COVERAGE

SUCH THAT

$$P[a < \theta < b] = P[\theta \in C_n] \geq \underline{1-\alpha} \quad \text{FOR ALL } \theta \in \Theta$$

INTERPRETATION

CAN MAKE PROB STATEMENTS

RANDOM DEPENDING ON DATA

$$\hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

NOTHING RANDOM

$$(0.48 - 0.07,\ 0.48 + 0.07)$$

DO NOT MAKE PROBABILITY STATEMENTS!!!

Example

Berger and Wolpert

X_1	X_2	
-1	-1	✗
-1	1	✓
1	-1	✓
1	1	✓

X_1, X_2 IND $\quad P[X_i = 1] = P[X_i = -1] = 0.5$

Define $Y_i = \theta + X_i$ $\quad$ Observe Y_1, Y_2

Define "interval" $\quad C = \begin{cases} Y_1 - 1 & \text{if } Y_1 = Y_2 \\ (Y_1 + Y_2)/2 & \text{if } Y_1 \neq Y_2 \end{cases}$

$P[\theta \in C] = 3/4$ for any $\theta \Rightarrow 75\%$ CI

$Y_1 = 15$
$Y_2 = 17$ $\Rightarrow \theta = 16$ $\qquad P[\theta \in C \mid Y_1 \neq Y_2] = 1$

Bayesian....

Normal Based CIs

Given $\hat{\theta}_n \approx N(\theta, \hat{se}^2)$ $\qquad \hat{se} = \hat{sd}(\hat{\theta}_n)$

$$P[Z > z_{\alpha/2}] = \alpha/2$$

Define $C_n = (\hat{\theta}_n - z_{\alpha/2} \cdot \hat{se},\ \hat{\theta}_n + z_{\alpha/2}\, \hat{se})$

Then $P(\theta \in C_n) \longrightarrow 1-\alpha$

$\alpha = 0.05 \Rightarrow z_{\alpha/2} = 1.96$ "=" 2 $\qquad \hat{\theta} \pm 2 \cdot \hat{se}$

$$Z_n = \frac{\hat{\theta} - \theta}{\hat{se}}$$

WE ASSUME $Z_n \xrightarrow{d} Z$

$Z \sim N(0,1)$

$$P(\theta \in C_n) = P\left(\hat{\theta} - z_{\alpha/2}\,\hat{se} < \theta < \hat{\theta} + z_{\alpha/2}\,\hat{se}\right)$$

$$= P\left(-z_{\alpha/2} < \frac{\hat{\theta} - \theta}{\hat{se}} < z_{\alpha/2}\right)$$

$$\approx P\left(-z_{\alpha/2} < Z < z_{\alpha/2}\right)$$

$$= 1 - \alpha$$

Hypothesis Testing

Assume $X_1, X_2, \ldots X_n \sim \text{Bernoulli}(p)$

Null $\longrightarrow$ $H_0 : \; p = 1/2$

Alternative $\longrightarrow$ $H_1 : \; p \neq 1/2$

Define $T = \left| \hat{p} - 1/2 \right|$

"Reject H_0" when T is "large"

How large?