STA 260 S 2020

Indicator functions

$$I_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases} = I(x \in A)$$

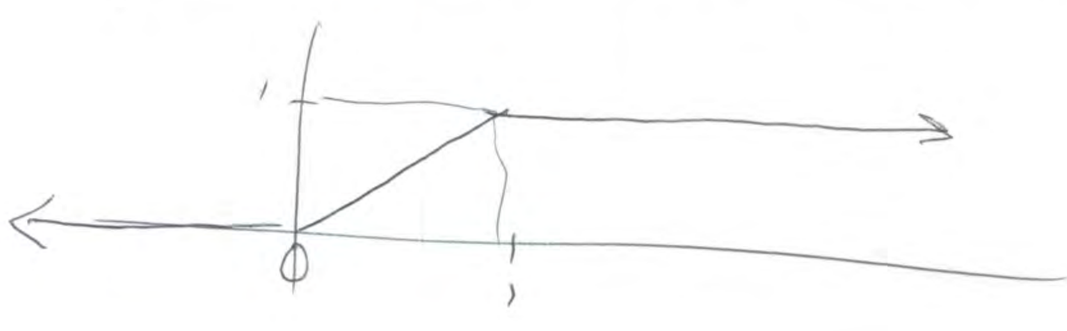
$$P(A) = \int_A f_X(x)\,dx = \int_{-\infty}^{\infty} \underbrace{I_A(x)}_{g(x)} f_X(x)\,dx = E(I_A(X))$$

Use Indicators to write prob density functions

$$f_X(x) = \begin{cases} e^{-x} & \text{for } x \geq 0 \\ 0 & \text{for } x < 0 \end{cases} = e^{-x} I(x \geq 0)$$

$$P_X(x) = \frac{e^{-\lambda}\lambda^x}{x!} I(x = 0, 1, \ldots)$$

$$F_X(x) = \cancel{0}\, x\, I(0 < x < 1) + I(x \geq 1)$$



Ex $X \sim U(0,1)$ so $f_X(x) = I(0<x<1)$

$Y = 2X+3$

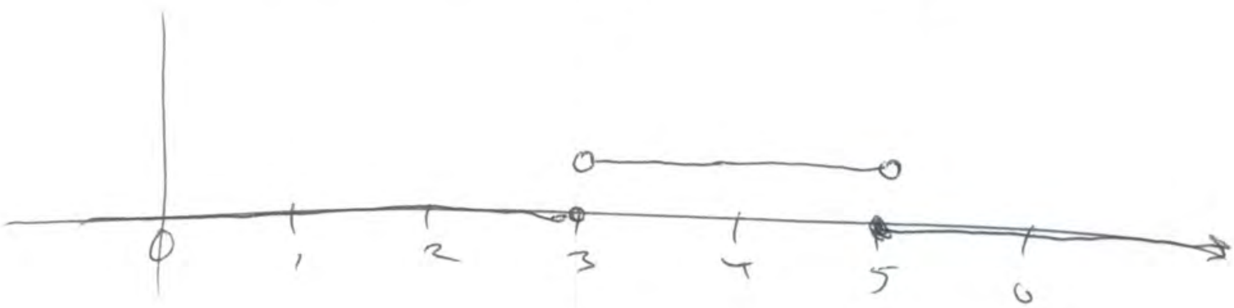
$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{d}{dy} P(Y \le y) = \frac{d}{dy} P(2X+3 \le y)$$

$$= \frac{d}{dy} P(2X \le y-3) = \frac{d}{dy} P\left(X \le \frac{y-3}{2}\right)$$

$$= \frac{d}{dy} F_X\left(\frac{y-3}{2}\right) = f_X\left(\frac{y-3}{2}\right) \cdot \frac{1}{2}$$

$$= I\left(0 < \frac{y-3}{2} < 1\right) \cdot \frac{1}{2} = I(0 < y-3 < 2) \cdot \frac{1}{2}$$

$$= \frac{1}{2} I(3 < y < 5) \quad U(3,5)$$



CLT: If $X_1, \ldots X_n \overset{iid}{\sim} ?(\mu, \sigma^2)$

Then $Z_n = \dfrac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \xrightarrow{d} Z \sim N(0,1)$

Theorem If $\hat{\sigma}_n^2 \xrightarrow{P} \sigma^2$ then

$$Z_n = \frac{\sqrt{n}(\bar{X}_n - \mu)}{\hat{\sigma}_n} \xrightarrow{d} Z \sim N(0,1)$$

Def A "random sample" is a collection of independent and identically distributed random variables.

For ex $X_1, \dots X_n \overset{iid}{\sim} N(\mu, \sigma^2)$

n is the sample size

Ex $X_1, \dots X_n \overset{iid}{\sim} \text{Bernoulli}(\theta)$

In general $X_1, \dots X_n \overset{iid}{\sim} P_\theta \, , \theta \in \Omega$

Omega

or ind   theta

θ is the parameter: Whatever is unknown about the problem

Ω is the parameter space $\theta \in \Omega$

n is the sample size

$X_i \sim \text{Bernoulli}(\theta)$, $\Omega = \{\theta : 0 < \theta < 1\}$

$X_i \sim \text{Normal}(\mu, \sigma^2)$ $\theta = (\mu, \sigma^2)$

$\Omega = \{(\mu, \sigma^2) : -\infty < \mu < \infty, \sigma^2 > 0\}$

Def A <u>statistic</u> is a function of the sample data that does not depend (<u>functionally</u>) on any unknown parameters.

<u>Ex</u> $X_1 \cdots X_n \overset{iid}{\sim} N(\mu, \sigma^2)$

- $\overline{X}_n$ Yes
- $\dfrac{\sqrt{n}\,(\overline{X}_n - \mu)}{\sigma}$ No
- $S^2 = \dfrac{\sum_{i=1}^{n}(X_i - \overline{X}_n)^2}{n-1}$ Yes
- X_3 Yes

Def Estimator: A statistic that is a formula for estimating θ or sometimes a function $\psi(\theta)$ of θ.

When formula is applied to set of numerical data values $x_1, \ldots, x_n$ it yields an <u>estimate</u>.

Example Diagnosis with a particular type of cancer. What proportion will be alive 5 years after diagnosis?

Model $X_1, \ldots X_n \overset{iid}{\sim} \text{Bernoulli}(\theta)$

$1 = \text{alive} \quad 0 = \text{dead}$

Estimate θ with $\hat{\Theta}_n = \overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$

Estimator

If 60/200 in the sample survive, the **Estimate** is $\bar{x}_n = \frac{1}{n} \sum_{i=1}^{n} x_i = \frac{60}{200} = 0.3$

Desirable properties of Estimators

- Unbiased
- Consistent
- Efficient (precise)
- Sufficent

__Def__ An estimator $\hat{\Theta}$ or $\hat{\Theta}_n$ is said to be unbiased for θ if

$$E(\hat{\Theta}) = \theta \quad \text{for all } \theta \in \Omega$$

Bernoulli cancer example

$\hat{\Theta}_n = \bar{X}_n$ is unbiased because

$$E(\hat{\Theta}_n) = E(\bar{X}_n) = \cancel{E(X_i)}$$
$$= E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i)$$
$$= \frac{1}{n}\sum_{i=1}^{n} \theta = \frac{1}{n} n\theta = \theta$$

~~Also~~

__Def__ An estimator $\hat{\Theta}_n$ is said to be ~~const~~ consistent for θ if $\hat{\Theta}_n \xrightarrow{P} \theta$ for all $\theta \in \Omega$, meaning

For all $\varepsilon > 0$ $\lim_{n\to\infty} P\{|\hat{\Theta}_n - \theta| \geq \varepsilon\} = 0$

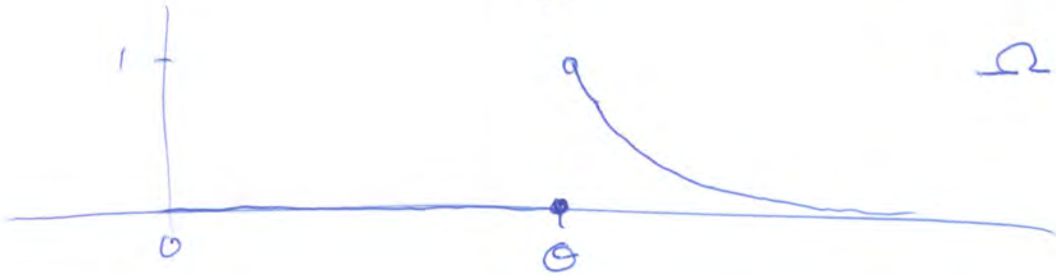
or $\lim_{n\to\infty} P\{|\hat{\Theta}_n - \theta| < \varepsilon\} = 1$

$\hat{\Theta}_n = \bar{X}_n$ is consistent for Bernoulli–cancer ex by the Law of Large Numbers

Ex Shifted Exponential

$$X_1, \dots, X_n \overset{iid}{\sim} f_X(x|\theta) = e^{-(x-\theta)} I(x>\theta)$$

$\Omega = \{\theta : \theta > 0\}$



Consider $\hat{\Theta}_1 = \bar{X} - 1$, $\hat{\Theta}_2 = Min(X_i)$

Are they unbiased?

$$E(\hat{\Theta}_1) = E(\bar{X} - 1) = E(\bar{X}) - 1$$

$$= E(X_i) - 1$$

$$E(X_i) = \int_\theta^\infty x e^{-(x-\theta)} dx$$

$u = x - \theta$
$x = u + \theta$
$dx = du$

x	u
∞	∞
θ	0

$$= \int_0^\infty (u+\theta) e^{-u} du$$

$$= \int_0^\infty u e^{-u} du + \int_0^\infty \theta e^{-u} du = 1 + \theta$$

Exp$(\lambda = 1)$

$$E(\hat{\Theta}_1) = E(X_i) - 1 = 1 + \theta - 1 = \theta$$ <u>unbiased</u>

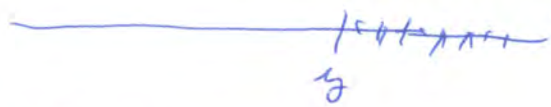
To check $E(\hat{\Theta}_c) = E(Min(X_i))$,
need density of the minimum

Let $Y = Min(X_i)$

$$f_Y(y) = \frac{d}{dy}F_Y(y) = \frac{d}{dy}P(Y \le y)$$

$$= \frac{d}{dy}P(Min(X_i) \le y) = \frac{d}{dy}\left(1 - P(Min(X_i) > y)\right)$$

$$= -\frac{d}{dy}P(\text{All } X_i > y)$$



$$= -\frac{d}{dy}P\left(\bigcap_{i=1}^{n}\{X_i > y\}\right) \overset{ind}{=} -\frac{d}{dy}\prod_{i=1}^{n}P(X_i > y)$$

$$= -\frac{d}{dy}\prod_{i=1}^{n}\left(1 - F_X(y)\right) = -\frac{d}{dy}\left(1 - F_X(y)\right)^n$$

$$= (+1)\, n\left(1 - F_X(y)\right)^{n-1} \cdot (+1)\, f_X(y)$$

$$= n\left(1 - F_X(y)\right)^{n-1} e^{-(y-\theta)} I(y > \theta)$$

Need $F_X(y)$

For $y > \theta$

$$F_X(y) = \int_\theta^y e^{-(x-\theta)}\,dx \qquad \begin{array}{l} u = x-\theta \\ du = dx \end{array}$$

x	u
y	$y-\theta$
θ	0

$$= \int_0^{y-\theta} e^{-u}\,du$$

$$= 1 - e^{-(y-\theta)} \quad \text{, so}$$

$$f_Y(y) = n\left(1-\left(1-e^{-(y-\theta)}\right)\right)^{n-1} e^{-(y-\theta)} I(y>\theta)$$

$$= n\left(e^{-(y-\theta)}\right)^{n-1} e^{-(y-\theta)} I(y>\theta)$$

$$= n e^{-n(y-\theta)} I(y>\theta)$$

And $E(Y) = E(\hat{\theta}_2) = \displaystyle\int_\theta^\infty y\, n e^{-n(y-\theta)}\,dy$

$$= \int_0^\infty (u+\theta) n e^{-nu}\,du \qquad \begin{array}{l} u = y-\theta \quad y = u+\theta \\ du = dy \end{array}$$

y	$u = y-\theta$
∞	∞
θ	0

$$= \int_0^\infty u \underbrace{n e^{-nu}}_{Exp(\lambda=n)}\,du + \theta \int_0^\infty n e^{-nu}\,du$$

$= \frac{1}{n} + \theta \cdot 1 = \theta + \frac{1}{n}$ Biased

$\neq \theta$

Def Bias of an estimator $\hat{\Theta}$ is

$b(\theta) = E(\hat{\Theta}) - \theta$ here bias $= \frac{1}{n}$

$\lim_{n \to \infty} b(\theta) = 0$ then $\hat{\Theta}$ is called asymptotically unbiased

$\Theta_3 = \Theta_2 - \frac{1}{n}$ is unbiased

$\uparrow$

$Min(X_i) - \frac{1}{n}$

$\hat{\Theta}_1$ & $\hat{\Theta}_3$ are both unbiased. Which is "better"?

$Var(\hat{\Theta}_1) = Var(\bar{X}_n - 1) = Var(\bar{X}_n) = \frac{\sigma^2}{n}$

What is $\sigma^2 = Var(X_i)$

Rather than calculating $Var(X_i) = \sigma^2$ by $E(X_i^2) - E(X_i)^2$, note if

$X \sim Exp(\lambda)$ & $Y = X + \theta$ $\quad f_Y(y) = \frac{d}{dy} P(Y \le y)$

$$= \frac{d}{dy} P(X + \theta \le y) = \frac{d}{dy} P(X \le y - \theta)$$

$$= \frac{d}{dy} F_X(y - \theta) = f_X(y - \theta) \cdot 1$$

$$= \lambda e^{-\lambda(y-\theta)} I(y - \theta) > 0$$

$$= \lambda e^{-\lambda(y-\theta)} I(y > \theta)$$ Shifted exponential.

So X_i is just $Exp(1) + \theta$

$$Var(X_i) = Var \text{ of } exp(1) = \frac{1}{1^2} = 1$$

$$Var(\hat{\Theta}_1) = Var(\bar{X} - 1) = Var(\bar{X}) = \frac{1}{n}$$

Also note $f_Y(y) = n e^{-n(y-\theta)} I(y>\theta)$

& $\hat{\Theta}_2 = \min(X_i)$ is shifted exponential, shifted by θ & $\lambda = n$

$$Var(\hat{\Theta}_3) = Var(Y_1) = Var(X+\theta)$$

$X \sim Exp(\lambda = n)$

$$Var(\hat{\Theta}_3) = \frac{1}{\lambda^2} = \frac{1}{n^2}$$

compare

$$Var(\hat{\Theta}_1) = \frac{1}{n}$$