Lecture 3:

$$\hat{\theta}_{MLE} := \underset{\theta \in \Theta}{\text{argmax}} \{ l(\theta, x) \}$$

↑ Point estimate

$\bar{x}$ - small x

In advance class You will Prove

$$\underbrace{\hat{\theta}_{MLE}}_{\substack{\text{r.v "estimator"} \\ \bar{X}\text{ - big X}}} \overset{d}{\approx} N(\theta, \underbrace{SE[\hat{\theta}_{MLE}]}_{\text{a function of } \theta})$$

impossible to know

$$\overset{\text{2nd approx}}{\approx} N(\underset{\text{A estimator}}{\hat{\theta}_{MLE}}, \; SE[\underset{\text{estimator}}{\hat{\theta}_{MLE}}])$$

$$SE[\hat{\theta}_{MLE}] \Big|_{\theta = \hat{\theta}_{MLE}}$$

↑ estimator ↳ estimate.

MLE'S allow 3 goals of inference

1. pt estimator
2. Confidence sets, $CI_{\theta,1-\alpha}$

$$CI_{\theta,1-\alpha} := \left[\hat{\theta}_{MLE} \pm z_{\alpha/2}\,\hat{SE}[\hat{\theta}_{MLE}]\right]$$

3. Testing: $H_0 := \theta_0$

$$H_a : \underset{\substack{\uparrow \\ \text{some theory}}}{\theta \neq \theta_0}$$

$$RR_\alpha := \left[\theta_0 \pm z_{\alpha/2}\, SE[\hat{\theta}_{MLE}]\Big|_{\theta=\theta_0}\right]$$

Trouble in Paradise Examples:-

1. $\mathcal{F}$ = iid Bernoulli $\rightarrow$ $x_1, x_2, \dots x_n \overset{iid}{\sim} (\theta)$

Parametric model

$$X = \langle 0, 0, 0 \rangle.$$

$$\hat{\theta}_{MLE} = \bar{X} = 0$$

$$CI_{\theta,1-\alpha} = \left[\bar{X} \pm z_{\alpha/2} \sqrt{\frac{\bar{x}(1-\bar{x})}{3}} \right] = \{0\}$$

$RR_\alpha = \{\theta_0\}$ all test are rejected

(2) What if you know

$$\theta \in [0.1, 0.2]$$

is there any way to make use of this information?

NO!

(3) let's interpret the confidence input;

$$CI_{0,95\%} = [0.37, 0.43]$$

What is the interpretation?

$P(\theta \in CI_{0,95\%}) \to$ wrong

Our assumption θ is a fixed value (Parameter)

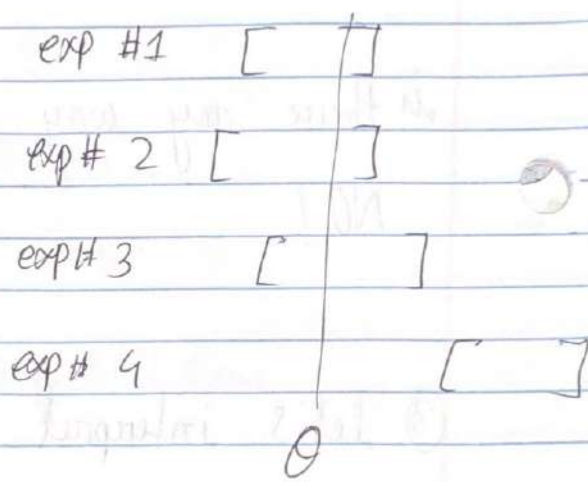
$$P(0.392 \in [0.37, 0.43]) = 1 \; \boxed{\to \text{True}}.$$

$$P(0.36 \in [0.37, 0.43]) = 0$$

Valid Interpretation:-

(I) if I repeat the experiment many times (around)

$\approx 95\%$ of the CI it will include θ.

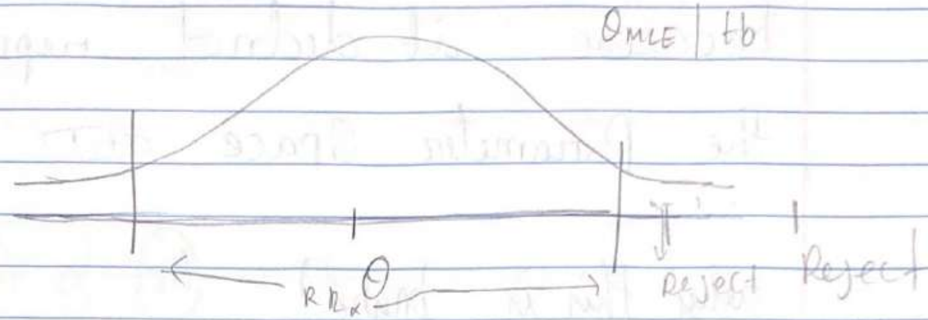


(II) In a hypothesis, You either reject H_0 or use H_0

$\hat{\theta}_{MLE} \in RR_\alpha$ = Retain H_0

$\hat{\theta}_{MLE} \notin RR_\alpha \Rightarrow$ Reject H_0

The smaller p value the strong rejection



p value defined as

$$P_{val} := P(\text{ seeing } \hat{\theta}_{MLE} \text{ "or more extreme" } | H_0 \text{ true})$$

$$\neq P(H_0 | X) \text{ truly what you want}$$

Proba my theory is true.

(5) $\mathcal{F}$ = iid Bernoulli $(\Theta = (0,1))$

$$X = \langle 0, 1, 0 \rangle$$

$$\hat{\theta}_{MLE} = \frac{0+1+0}{3} = \frac{1}{3}.$$

$$CI_{\theta, 95\%} = \left[\frac{1}{3} \pm 2\sqrt{\frac{1/3 \cdot 2/3}{3}}\right] = [-0.20, 0.87]$$

We know that $\theta \notin \Theta'$

This is a bad confidence set because it didnot represent the Parameter Space Θ

why this is brabe!

n is small

$\hat{\theta}_{MLE} \sim N(\, , \,) \rightarrow$ game over.

<u>Conditional Proba:</u>

A: Smoking (is an event)

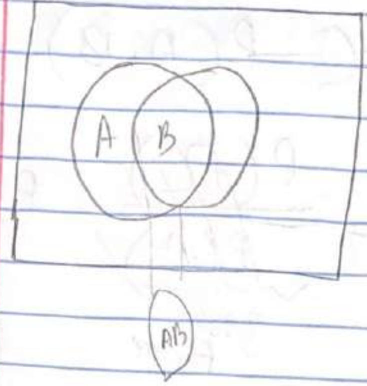
B: lung Cancer (" ")

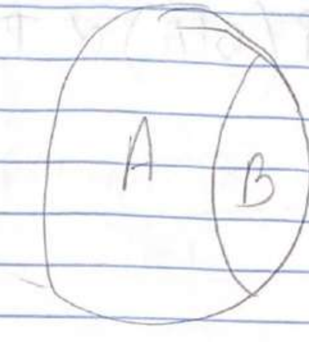
Assume; $P(A) = 0.2$

$P(B) = 0.06$

$P(A, B) = 0.036$

Ω = Universe





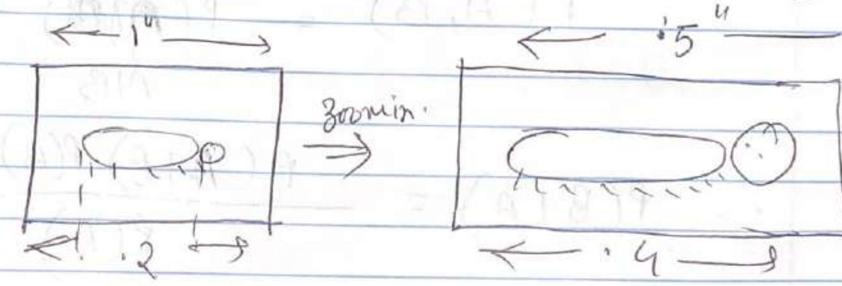
$(AB) \propto (B|A) \rightarrow$ same shape.

Proba of lung Cancer given Smoking

P(lung Cancer | Smoking), is a Conditional Probability.

$= P(B/A)$

$P(B|A) \propto P(A,B) = \overset{\text{multiple}}{C}\; P(A,B)$



Zoom factor, $\frac{1''}{.5''} = 2.$

$$P(B|A) \propto P(A,B) = C\, P(A,B)$$

$$= \cancel{\underbrace{\frac{P(\Omega)}{P(A)}}_{\text{zoom factor}} \Big/ P(A,B)}$$

$$= \frac{P(\Omega)}{P(A)} \cdot P(A,B)$$

$$P(B|A) = \frac{P(A,B)}{P(A)} \quad \text{def of conditional prob}$$

$$\Rightarrow P(A,B) = P(B|A)\, P(A)$$

$$P(A|B) = \frac{P(A,B)}{P(B)}$$

$$P(A,B) = \underset{A|B}{P(A|B)}\, P(B)$$

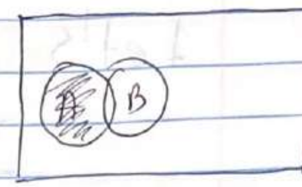
$$\therefore P(B|A) = \frac{P(A|B)\, P(B)}{P(A)}$$

$$\Rightarrow P(B_i|A) = \frac{P(A|B_i)P(B_i)}{\sum_{K=1}^{K}(A, B_K)} \quad \text{Bayes rule}$$

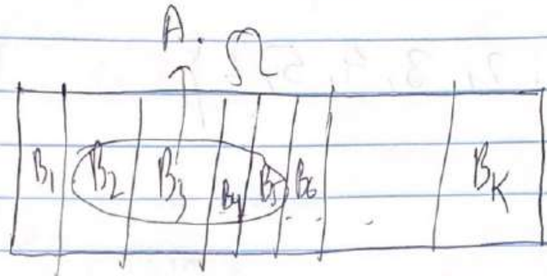
$$\text{Bayes Rule:- } P(B|A) = \frac{P(B|A)P(B)}{P(A)}$$

$$A = (A \cap B) \cup (A \cap B^c)$$

$$P(A) = P(A \cap B) + P(A \cap B^c)$$



Addition rule



S.t. $B_1 \cup B_2 \cup B_3 \cup \dots \cup B_K = \Omega$ Collecting exclusive

But $B_i \cap B_j \neq 0$ mutually exclusive

I can Prove this;

$$A = (A \cap B_1) \cup (A \cap B_2) \cup \dots \cup (A \cap B_K)$$

$$P(A) = \sum_{K=1}^{K} P(A, B_K)$$

Bayes theorem:-

$$P(B_i \mid A) = \frac{P(A \mid B_i)\, P(B_i)}{\sum_{k=1}^{K} P(A \mid B_k)\, P(B_k)}$$

Let's say,

Imagine two r.v's X, Y

$Sup[X] = \{1, 2, 3, 4\}$

$Sup[Y] = \{1, 2, 3, 4, 5, 6\}$

→ table gives the $P(X=x, Y=y)$ i.e. JmF

	1	2	3	4	5	6
1						
2						
3						
4					$P(Y=5)$	

How big is this relative to everything

the marginal of the table

Sum of col

marginal Probability:

$$P(y=5) = P(y=5, X=1) + P(y=5, X=2) + P(y=5, X=3) + P(y=5, X=4)$$

$$= \sum_{x \in Supp[X]} P(y=5, X=x)$$

$$P(X=2 \mid Y=5) = \frac{P(X=2, Y=5)}{P(Y=5)}$$

$$P(X=x \mid Y=y) = \frac{P(X=x, Y=y)}{P(Y=y)}$$

Conditional mass function (Cmf)

$$P(X \mid Y) = \frac{P(X,Y) \rightarrow \text{Jmf}}{P(Y) \rightarrow \text{Pmf}} = \frac{P(Y \mid X)\, P(X)}{P(Y)}$$

Can I write the following;

$$P(\theta|x) = \frac{P(x|\theta)\, P(\theta)}{P(x)}$$

Currently, θ is constant, i.e discrete r.v.

thus the formula is not useful.

$\theta \sim \{ \theta$ w.p 1 $\quad F = \{ P(x;\theta) ; \theta \in \Theta \}$

$\theta | X \sim \{ \theta$ w.p 1.

→ $P(x)$ without knowing θ. This is unusable knowing θ -

$P(X) =$ discrete $\sum_{\theta_0 \in \Theta} P(X|\theta_0)\, P(\theta_0)$

Cont. $\int_{\Theta} P(X|\theta_0)\, P(\theta_0)\, d\theta_0$.

denominator is a problem.