Lecture 03: 02/04/2020.

$\hat{\theta}_{MLE} := \arg\max_{\theta \in \Theta} \{ \ell(\theta, x) \}$

point estimate.
$\bar{x}$

In an advance class you will prove.

$\hat{\theta}_{MLE} \overset{d}{\approx} N(\theta, \underbrace{SE[\hat{\theta}_{MLE}]}_{\text{a function of } \theta})$

r.v "estimatore"
$\bar{X}$ big X

2nd approx

$\approx N(\hat{\theta}_{MLE}, SE[\hat{\theta}_{MLE}])$

$\hat{\theta}_{MLE}$: A estimatore

estimatore

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

$SE[\hat{\theta}_{MLE}]$: estimatore

$\hat{\theta}_{MLE}$: estimate

MLE's allow for 3 goals of inference.

1. Pt. estimation.
2. Confidence sets, $CI_{\theta, 1-\alpha} := \left[\hat{\theta}_{MLE} \pm z_{\alpha/2} \hat{SE}[\hat{\theta}_{MLE}] \Big|_{\theta = \hat{\theta}_{MLE}} \right]$
3. Hypothesis Testing $H_0 = \theta_0$ $\quad RR_\alpha := \left[\theta_0 \pm z_{\alpha/2} SE[\hat{\theta}_{MLE}] \Big|_{\theta = \theta_0} \right]$

   $H_a : \theta \neq \theta_0$
   some theory

Trouble in paradise examples.

① F = iid Bernoulli $\to x_1, x_2, x_3 \ldots x_n \overset{iid}{\sim} (\theta)$

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.

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

What is the interpretation?

Andrew's theory: $P(\theta \in CI_{\theta, 95\%}) > 95\% \to$ 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.

① If I repeat the experiment many times, (around) $\approx 95\%$ of the CI If will include θ.

exp #1 [ ] CI #1

exp #2 [ ] #2

exp #3 [ ] #3

exp #4 [ ]

θ

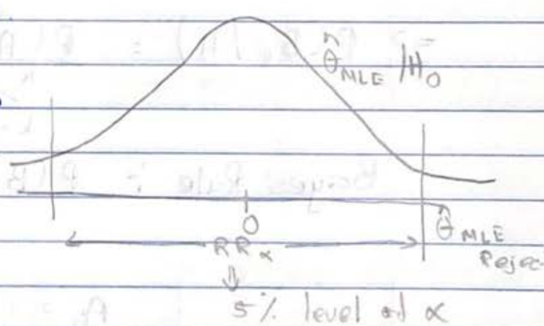
② Before you do the experiments,

$P(\theta \in CI_{\theta, 1-\alpha}) = 1-\alpha$

④ In a hypothesis, you eithere rejection of 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" is defined as.

Pval := P(seeing $\hat{\theta}_{MLE}$ or more extreme | H_0)

HOKEY!!

Pval = P(seeing $\hat{\theta}_{MLE}$ or more extreme | H_0 true)

$\neq P(H_0 | X)$ truely what you want

proba my theory is true

(5) F = iid Bernoulli $(\theta \in (0,1))$

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

$\hat{\theta}_{MLE} = \bar{X} = 1/3$

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

$= [-0.20, 0.87]$

Page ##

We know that $\theta \not< 0$

This is a bad confidence set; because it didnot represent the parameter space Θ.

Why this is bnabe!

n is small ; $\hat{\theta}_{MLE} \sim N(\quad , \quad) \to$ game over

Conditional Proba's

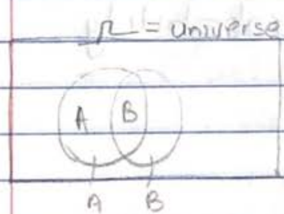
A: smoking (in an event)

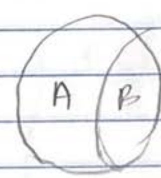
B: lung cancer (in an event)

Assume: $P(A) = 0.2$

$P(B) = 0.06$.

$P(A,B) = 0.036$





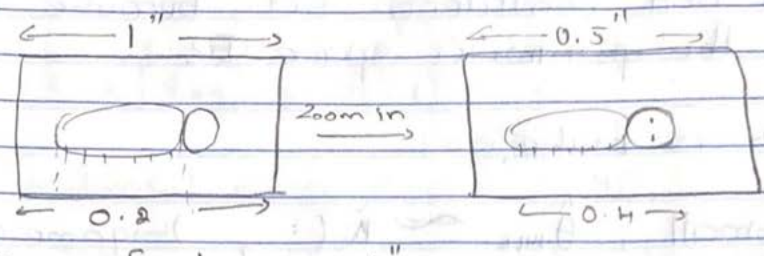
AB $\propto$ B/A $\longrightarrow$ some shape

proba of lung cancer given ~~sd~~ smoking

P(lung cancer| smoking) in a conditional probability

$= P(B|A)$

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



zoom factor $= \frac{1''}{0.5''} = 2.$

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

$$= \frac{P(1)}{P(A)} \;.\; P(A,B)$$

$$P(B|A) = \frac{P(A,B)}{P(A)}$$ def of conditional probability.

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

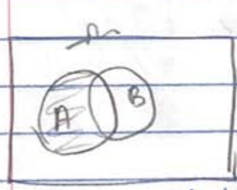
$$P(A|B) = \frac{P(A,B)}{P(B)}$$

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

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

$$\Rightarrow P(B_i|A) = \frac{P(A|B_i)\,P(B_i)}{\sum_{k=1}^{K}(A,B_k)}$$ Bayes Rule

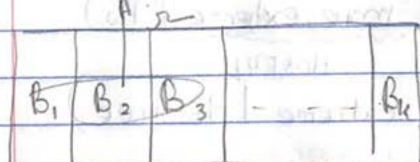
Bayes Rule :- $P(B|A) = \frac{P(B|A)\,P(B)}{P(A)}$



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

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

Addition Rule.



s.t $B_1 \cup B_2 \cup B_3 \cup \dots \cup B_k$

$= \Omega$ collecting exclusive

but $B_i \cap B_j \neq \emptyset$ mutually exclusive.

I can prove this,

$$A = (A \cap B_1) \cup (A \cap B_2) \cup \ldots \cup (A \cap B_k)$$

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

$\boxed{\text{Bayes theorm}}$

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

Let's say

Imagine two r.v's x, y

$supp[x] = \{1, 2, 3, 4\}$
$supp[y] = \{1, 2, 3, 4, 5, 6\}$

y

X	1	2	3	4	5	6
1						
2					·	
3						
4					P(y=5)	

→ table give the $p(X = x, Y = y)$

How big is this relative to entries

the marginal of the table

sum of col

marginal probability:

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

$$= \sum P(Y=5, X=x) \; \{x \in supp[X]$$

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

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

Conditional mass function (CMF)

Conditional Mass function

$$\underset{\text{CMF}}{P(x/y)} = \frac{P(x,y)}{P(y)} \longrightarrow \text{PMF} = \frac{P(y|x)\, P(x)}{P(y)}$$

Can I write the following?

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

Currently, θ is constant, i.e. discrete r.v than the formula is not useful.

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

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

$P(x)$ without knowing θ. This is unusuable knowing θ

discrete $\sum_{\theta_0 \in \Theta} P(x \mid \theta_0)\, P(\theta_0)$

$P(x)$

$$\int_{\Theta} P(x \mid \theta_0)\, P(\theta_0)\, d\theta_0$$

dinominator is a problem