02/04/2020 Lecture 3

$$\hat{\theta}_{MLE} := \underset{\theta \in \Theta}{\operatorname{argmax}} \{ \ell(\theta; x) \}$$

pt. estimate
"estimate"
$\bar{x}$

In an advance class, you'll prove

estimator

$$\underset{\substack{\text{r.v} \\ \text{"estimator"} \\ \bar{X}}}{\hat{\theta}_{MLE}} \overset{d}{\approx} \underbrace{N(\theta, \underbrace{SE[\hat{\theta}_{MLE}]}_{\text{a function of } \theta})}_{\text{impossible to know}} \underset{\text{2nd approx.}}{\approx} N(\hat{\theta}_{MLE}, \underbrace{\hat{SE}[\hat{\theta}_{MLE}]}_{\underset{\text{estimator}}{SE[\hat{\theta}_{MLE}]}\big|_{\underset{\text{estimate}}{\theta = \hat{\theta}_{MLE}}}})$$

MLE's allow for 3 goal of inference

① Pt. estimator $\hat{\theta}_{MLE}$

② Confidence sets $CI_{\theta, 1-\alpha} := [\hat{\theta}_{MLE} \pm z_{\frac{\alpha}{2}} \hat{SE}[\hat{\theta}_{MLE}]]$

③ Testing $H_0: \theta = \theta_0$ $\quad RR_\alpha := [\theta_0 \pm z_{\frac{\alpha}{2}} SE[\hat{\theta}_{MLE}]\big|_{\theta=\theta_0}]$
$H_a: \theta \neq \theta_0$
some theory

Trouble in Paradize Examples

① $\mathcal{F}$ = iid Bernoulli $X_1, \dots, X_n \overset{iid}{\sim} Bern(\theta)$

$x = \langle 0, 0, 0 \rangle$

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

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

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

② What if you know

$\theta \in [0.1, 0.2]$ Is there any way to make use of this info?

<u>No</u>

③ Let's interpret the confidence interval

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

what is the interpretation?

$\Rightarrow P(\theta \in CI_{\theta, 95\%}) = 95\%$. <u>Wrong</u>, but it's what you want to say.

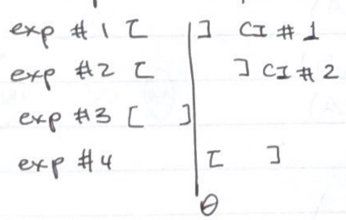
Our assumption was θ is a fixed value (parameter)

$P(0.392 \in [0.37, 0.43]) = 1$

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

Valid Interpretation

Ⓘ If I repeat the experiment many times, $\approx 93\%$ of the CI's will include θ



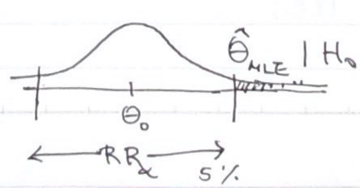
Ⅱ Before you do the experiment...

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

④ In a hypothesis test you either reject H_0 or retain H_0

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

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



smaller p value stronger the rejection.

33:00

"p-value" is defined as -

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

$$\neq P(H_0 \mid X) \leftarrow \text{prob. my theory is true} \quad \text{(what you want)}.$$

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

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

$\hat{\theta}_{MLE} = \bar{x} = \frac{1}{3}$

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

we know that $\theta \notin 0$! This is a bad confidence set because it the param space Θ.

why did this happen?

n is small ... therefore $\hat{\theta}_{MLE} \not\sim N(\ ,\ ) \Rightarrow$ game over.

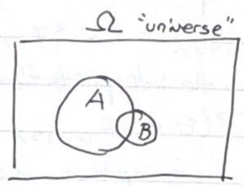
A: smoking

B: lung cancer

Assume: $P(A) = 0.2$

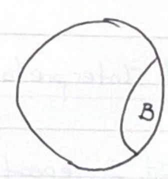
$P(B) = 0.06$

$P(A, B) = 0.036$

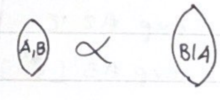


$\Downarrow$

$A = \Omega' \subset \Omega$



$\Downarrow$

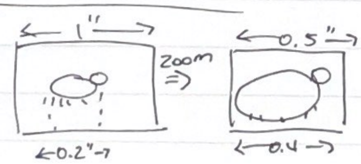


P(lung cancer | smoking) a "conditional probability"
$= P(B \mid A)$

$$P(B \mid A) \propto P(A, B) = \underset{\text{"multiple"}}{c}\, P(A, B) = \overbrace{\frac{P(\Omega)}{P(A)}}^{\text{zoom}} P(A, B)$$

$$\underset{\text{definition of cond. prob.}}{=} \frac{P(A, B)}{P(A)}$$

ex.



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

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

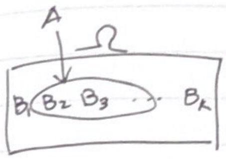
$P(A \mid B) = \frac{P(A, B)}{P(B)} \Rightarrow P(A, B) = P(A \mid B)\, P(B)$.

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

"Bayes Rule"

Addition Rule

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



I can prove this...

$$A = (A \cap B_1) \cup (A \cap B_2) \cup \ldots \cup (A \cap B_K)$$
$$\Rightarrow P(A) = \sum_{k=1}^{K} P(A, B_K)$$

s.t $B_1 \cup B_2 \cup \ldots \cup B_K = \Omega$ collectively exclusive.
but $B_i \cap B_j = \emptyset$ mutually exclusive

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

$$P(B_i \mid A) = \frac{P(A \mid B_i)\, P(B_i)}{\sum_{k=1}^{K} P(A \mid B_K)\, P(B_K)}$$ "Bayes Theorem"

Imagine 2 r.v.s X, Y.

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

Table gives the $P(X=x, Y=y)$ i.e the joint mass function

X \ Y	1	2	3	4	5	6
1						
2						
3						
4						

Marginal Probability

$$P(Y=5) = P(Y=5, X=1) + P(Y=5, X=2) + P(Y=5, X=3) + P(Y=5, X=4)$$
$$= \sum_{x \in Supp[X]} P(Y=5, X=x)$$

$\Sigma P(Y=5)$
the margin of the table.

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

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

$$P(X \mid Y) = \frac{P(X, y) \leftarrow \text{JMF}}{P(x) \longleftarrow \text{PMF}} = \frac{P(y \mid x)\, p(y)}{p(x)}$$

conditional mass function. CMF

Can I write the following ?

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

currently Θ is a constant,
i.e a degenerate r.v

$\Theta \sim \{ \Theta$ w.p 1

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

Thus the formula is not simple.

$$P(x) = \sum_{\Theta_0 \in \Theta} P(X | \Theta_0) P(\Theta_0)$$

$$\int_{\Theta} P(X | \Theta_0)\, P(\Theta_0)\, d\Theta_0$$

denominator is a problem.

$P(x)$ without knowing Θ this is unassumable without knowing Θ.

$$\mathcal{F} = \{ p(X; \theta) ; \theta \in \Theta \}.$$