lec. 3

Feb 8

Some function of the data $X_1, \ldots, X_n$

$\hat{\theta} = S(x_1, \ldots, x_n) = \bar{x} \in \mathbb{R}$ e.g 0.198

$F \overset{iid}{\sim} Bern(\theta)$

Some function of the data $X_1, \ldots X_n$

$F \overset{iid}{\sim} Bern(\theta)$

by property z

$\hat{\theta}_{MLE} = S(X_1, X_2, \ldots X_n) = \bar{X} \overset{\cdot}{\sim} N(\theta, SE[\hat{\theta}_{MLE}])$

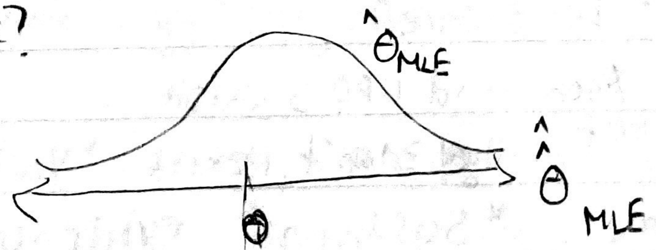
the same function except of the random variable

a random variable, not a value
random variable have distribution

we use this normality to create the confident interval (CI)

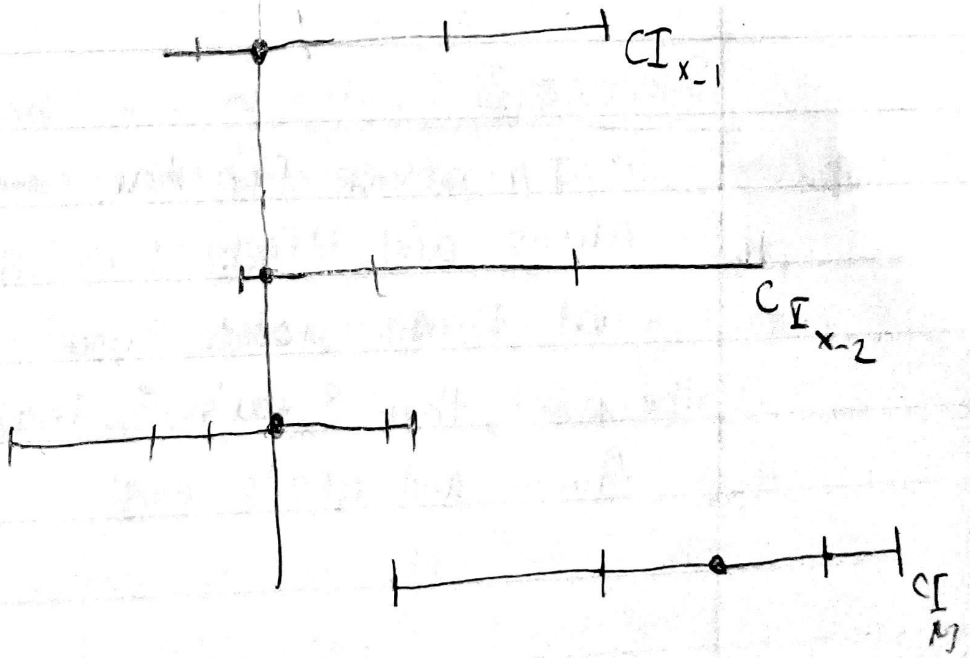
Why do CI work?

if the property is write there is a distribution



$\alpha = 5\%$

$z_{\alpha/2} = 1.96$



CI's have 95% probability (before they're realized) of capturing (i.e. including) the real value of theta

Inference goal #3: "hypothesis testing" also called "theory testing"

Consider a situation where "I" am trying to convince "you" of something

Scenario I:

"I" declare

H_a = Alien and UFO's exist

If they don't exist "you" need to provide "me" "sufficent" evidence that:

H_0 = Alien and UFO's do not exist.

Scenario II:

"I'll assume for the moment that

H_0 = Aliens and UFO's don't exist

and I will provide "you" sufficent evidence to the point that "you're" that

H_a = Alien and UFO's exist

Scenario II: is more convincing and it is how science generally works the theory I'm trying to demonstrate is called the "alternative hypothesis" (H_a Since it's alternative to maybe business-as-usual. In scenario II, you assume the oppositive of the theory which is called the "null hypothesis" (H_0). This is the hypothesis testing" procedure.

In our context, theories are phrases as mathamatical statements about theta the unknown parameter. We will study three types of H_a's:

$H_a: \theta \neq \theta_0$ ← constant value vs $H_0: \theta = \theta_0$ → two sided test/ two tail test

$H_a: \theta > \theta_0$ $H_0: \theta \leq \theta_0$ → right-side test/ right-side tail

$H_a: \theta < \theta_0$ $H_0: \theta \geq \theta_0$ → left-sided test/ Left-tailed test

There are two outcome of a hypothesis test

You were not shown sufficient evidence of H_a, Thus, you "fail to reject H_0," or "retain H_0".

You were indeed shown sufficient evidence of H_a. Thus, you "reject H_0," or "accept H_a". (also cheat sheet

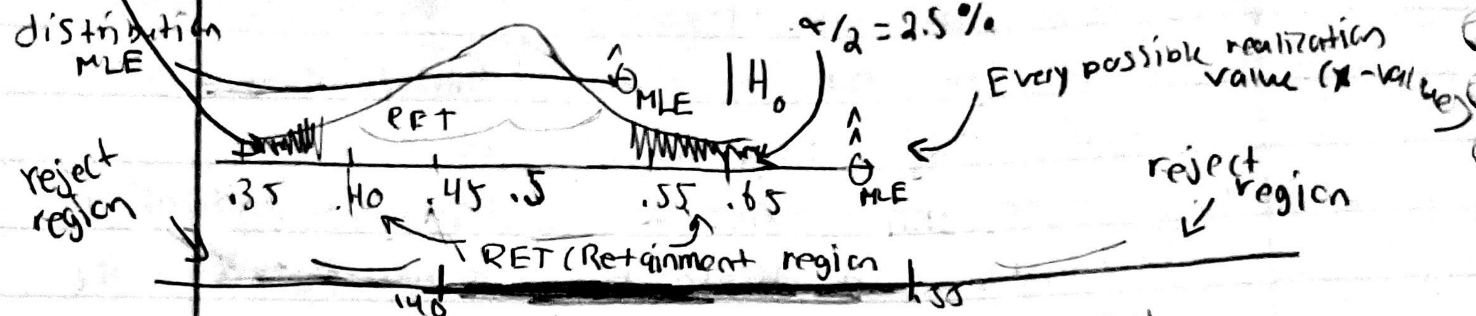
Imagine you're flipping a coin n times and you're counting the number of HEADS, then F: iid Bern(theta). You want prove the coin is unfair. →

$H_a: \theta \neq 0.5$

$H_0: \theta = 0.5$ $\bar{x}$ $0.5 = \rightarrow = \sqrt{\frac{0.5(1-0.5)}{n}} = 0.05$

$\alpha/2 = 2.5\%$

If H_0 is true, $\theta = 0.5 \Rightarrow \hat{\theta}_{MLE} \sim N(\theta, SE[\hat{\theta}_{MLE}]^2)$



What constitue "sufficient evidence". It's a probability of rejecting when H_0 is true. (Denoted α, alpha). Everyone is different.

If alpha = 5% in a 2-tailed test, we put 1/2 the probability in each tail. 5% is the most common scientific standard. Thus, we retain the H_0 for the most non-weird 95% of the thetahathats and reject H_0 for the 5% most weird thetahathats.

Write cheat sheet →

$$RET = [\theta_0 \pm z_{\alpha/2} SE[\hat{\theta}_{MLE}]] = \left[0.5 \pm 1.96\sqrt{\frac{0.5(1-.5)}{100}}\right]$$

$$= [0.402, .590]$$

ex. if $\bar{x} = \frac{61}{100} = .61$ $\alpha = 5\%$

→

NOT in the RET

$\hat{\theta}_{MLE} = .61 \notin RET \Rightarrow$ reject H_0 and conclude coin is unfair

<u>e.x.</u> if $\bar{x} = \frac{59}{100} = .59$

$\hat{\theta}_{MLE} = .59 \in RET \Rightarrow$ fail to reject H_0 and conclude there's not enough evidence of coin being unfair

We've covered the "frequentist" approach to statistical inference. But there are problems w/ it....

① $\mathcal{F}$ = iid Bern(θ), $x = \langle 0,0,0 \rangle$

$\vec{x}$ it's a vector, it has magnitude

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

Is that a good point estimate? No. You shouldn't be able to say something is absolutely impossible after $n=3$ trials.

$$CI_{\theta, 1-\alpha} = \left[0 \pm 1.96\sqrt{\frac{0(1-0)^2}{3}}\right] = \{0\}$$

Is this good confidence set? No. This is not a good set of "reasonable values."

(2) what if you had prior knowledge that Θ was restricted to e.g. [0.1, 0.2] and not the full (0,1)? Ans. ⇒ you can't "enter that into" your inference, by using MLE

(3) Consider the frequentist interpretation of a CI.

(1) Before you do the experiment, you have a 95% probability of capturing theta. But this doesn't tell you anything about after your experiment. After your experiment you have an interval e.g. [0.37, 0.43] and can't say:

$$P(\theta \in [0.37, 0.43]) = 0.95$$

can't say = NO randomness; no probability

(2) 95% of CI's will cover theta. But again, I only make one. So this interpretation doesn't help! In conclusion, any specific CI means NOTHING.

(4) Hypothesis tests result in a binary outcome: either you reject H_0 or you fail to reject H_0. What if you want to know

data $P(H_0 | X)$ or $P(H_a | X)$ data

you cannot!!! one thing you can do:

$$P_{val} := P(\text{seeing } \hat{\theta} \text{ or more extreme} | H_0) \neq P(H_0 | X)$$

⑤ F: iid Bern(θ), $x < 0,1,0 > \Rightarrow \hat{\theta} = .33$

$CI_{\theta, 95\%} = \left[.33 \pm 1.96\sqrt{\frac{.33 \cdot .67}{3}} \right] = [-0.20, .87]$

This is not legal confident set, because it's outside of the legal parameter space which is $(0,1)$

$\alpha = 5\%$

$H_0: \theta = 0.5\% = RET = \left[0.5 \pm 1.96\sqrt{\frac{.5 \cdot 0.5}{3}} \right] = [-.066, 1.066]$

This is a good hypothesis test because it's never going to reject.

The problem in #5 is b/c the asympototic normally of the MLE doesn't "kick in" until n is large (MLE property 2 is not true yet).