Lecture 3 2/8/21

$$\hat{\theta}_{MLE} = \underbrace{s(x_1, \dots, x_n)}_{\text{some function of the data } x_1, \dots, x_n} = \bar{x} \in \mathbb{R} \quad \text{e.g. } 0.1984$$

$\mathcal{F}$: iid Bern(θ)   by property 2

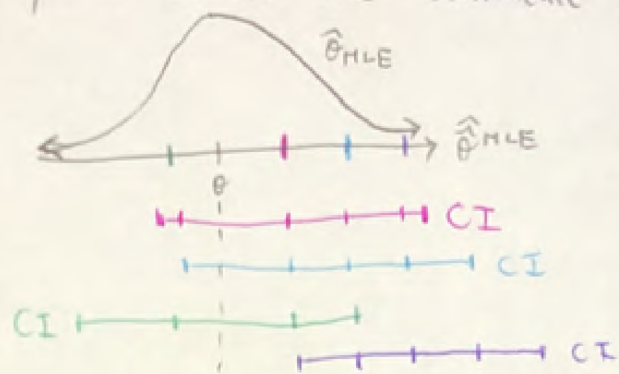
$$\hat{\theta}_{MLE} = \underbrace{S(X_1, X_2, \dots, X_n)}_{\text{the same function except of the r.v.'s}} = \bar{X} \mathrel{\dot\sim} N(\theta, SE[\hat{\theta}_{MLE}]^2)$$

A r.v., not a value, r.v.'s have distributions

We use the normality to create the confidence interval (CI). Why do CI's work?

$\alpha = 5\%$ $1-\alpha = 95\%$

$z_{\frac{\alpha}{2}} = 1.96$



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

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

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

Scenario I:

I declare

H_a: Aliens and UFO's exist.

If they don't exist *you* need to provide *me* *sufficient* evidence that:

H_0: Aliens and UFO's don't exist.

Scenario II:

I'll assume for the moment that

H_0: Aliens and UFO's don't exist

and I will provide *you* sufficient evidence to the point that *you're* convinced that

H_a: Aliens 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 opposite of the theory which is called the "null hypothesis" (H_0). This is the "Hypothesis Testing" procedure.

In our context, theories are phrased as mathematical statements about θ, the unknown parameter. We will study 3 types of H_a's:

$H_a: \theta \neq \theta_0$ ← constant value vs. $H_0: \theta = \theta_0$ ← 2-sided / 2 tailed test

$H_a: \theta > \theta_0$ vs. $H_0: \theta \leq \theta_0$ ← Right-sided / Right tailed test

$H_a: \theta < \theta_0$ vs. $H_0: \theta \geq \theta_0$ ← Left-sided / Left tailed test

There are two outcomes of a hypothesis test:

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

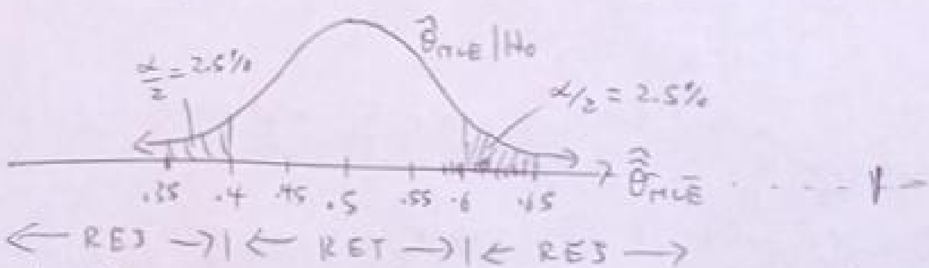
(B) You were indeed shown sufficient evidence of H_a. Thus, you "reject H_0" or "accept H_a."

Imagine you're flipping a coin $n=100$ times and you're counting the number of HEADS, then $F: \text{iid Bern}(\theta)$. You want to prove the coin is unfair.

$H_a: \theta \neq 0.5$

$H_0: \theta = 0.5$

If H_0 is true, $\theta = 0.5 \Rightarrow \underbrace{\hat{\theta}_{MLE}}_{=\bar{x}} \sim N(\underbrace{\theta}_{=0.5}, \underbrace{SE[\hat{\theta}_{MLE}]}_{=\sqrt{\frac{0.5(1-0.5)}{n}} = 0.05}^2)$



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

If $\alpha = 5\%$, in a 2-tailed test, we put $\frac{1}{2}$ the probability in each tail. 5% is the most common scientific standard. Thus, we retain H_0 for the most non-weird 95% of the $\hat{\theta}$'s and reject H_0 for the 5% most weird $\hat{\theta}$'s.

$$RET = [\theta_0 \pm z_{\frac{\alpha}{2}} SE[\hat{\theta}_{MLE}]] = [0.5 \pm 1.96 \underbrace{\sqrt{\frac{0.5(1-0.5)}{100}}}_{=0.05}] = [0.402, 0.590]$$

e.g. if $\bar{x} = \frac{61}{100} = 0.61$ $\alpha = 5\%$

$\hat{\hat{\theta}}_{MLE} = 0.61 \notin RET \Rightarrow$ Rej H_0 and conclude coin is unfair.

e.g. if $\bar{x} = \frac{59}{100} = 0.59$

$\hat{\hat{\theta}}_{MLE} = 0.59 \in RET \Rightarrow$ 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 with it...

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

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

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

$$CI_{\theta, 1-\alpha} = \left[0 \pm 1.96 \underbrace{\left(\sqrt{\frac{0(1-0)}{3}}\right)}_{0}\right] = \{0\} \rightarrow$$ Is this a 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 $(0,1)$. You can't "enter that into" your inference.

③ Consider the frequentist interpretation of a CI:

(1) Before you do the experiment, you have a 95% probability of capturing θ. 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 you can't say: $P(\theta \in \underbrace{[0.37, 0.43]}_{\text{no randomness!!}}) = 0.95$

(2) 95% of CI's will cover θ. But again, I only make one!!! So interpretation doesn't help me!

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

You cannot!!! One thing you can do is:

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

⑤ $\mathcal{F}$: iid Bern(θ), $x < 0, 1, 0 > \Rightarrow \hat{\hat{\theta}}_{MLE} = 0.33$

$$CI_{\theta, 95\%} = \left[0.33 \pm 1.96 \sqrt{\frac{0.33 \cdot 0.67}{3}}\right] = [-0.20, 0.87]$$

Is this a reasonable confidence set? No. It's outside of the legal parameter space which is $(0,1)$.

$\alpha = 5\%$

$H_0: \theta = 0.5 \Rightarrow RET = \left[0.5 \pm 1.96\sqrt{\frac{0.5 \cdot 0.5}{3}}\right] = [-0.066, 1.066]$

Is this a good hypothesis? No, because you NEVER reject!!!

The problem in #5 is because the asymptotic normality of the MLE doesn't "kick in" until n is large (MLE property 2 is not true yet).

We will solve all these problems with Bayesian Inference. Unfortunately, we will get other problems instead. It's a tradeoff and a personal decision.