Lecture 3

The Confidence Interval

Maximum likelihood estimate,

$$\hat{\hat{\theta}}_{MLE} = s(x_1,\ x_2,\ \dots,x_n) = \bar{x}$$

is some function of the data. For a Bernoulli Random Variable, this equals $\bar{x} \in \mathbb{R}$. The following equation, theta double hat, is a random variable, and rv's have distributions.

$$\hat{\theta}_{MLE} = s(x_1,\ x_2,\ \dots,x_n = \bar{X}\tilde{\sim}\mathbb{N}(\theta,\ SE[\hat{\theta}_{MLE}])$$

We use this normality to create the confidence interval (CI). Why do CI's work? 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 aliens and UFO's exist. If they don't exist, you need to provide me sufficient evidence that aliens and UFO's don't exist.

Scenario II

I'll assume for the moment that aliens and UFO's don't exist, and I will provide you sufficient evidence to the point that you're convinced that aliens and UFO's exist.

These scenarios are kind of inverses of each other. So lets pick them apart. In both scenarios I am trying to convince you that aliens exist. In the first one I put it on you to prove your claim, but in the second scenario I use proof by contradiction to prove it. The second scenario is a more stable approach. You must assume the opposite of the theory, the null hypothesis, denoted H_0. The theory I am trying to demonstrate is called the "alternative hypothesis", denoted H_a, since its alternative to business as usual. Science lives in scenario 2.

H_0: I'll assume for the moment that aliens and UFO's don't exist
H_a: aliens and UFO's exist.

This is the "hypothesis testing" procedure. Now we will put this into our context. Theories are phrased as mathematical statements about θ, the unknown parameter. We will study three types of theories, also called H_a's.

Two-sided, or two tailed test

$$H_a : \theta \neq \theta_0 vs H_0 : \theta = \theta_0$$

Right sided test, right tailed test

$$H_a : \theta > \theta_0 vs H_0 : \theta \leq \theta_0$$

Left sided test, left tailed test

$$H_a : \theta < \theta_0 vs H_0 : \theta \geq \theta_0$$

There are two outcomes, either you are shown sufficient evidence of H_a, and thus you reject H_0. Or you are not shown sufficient evidence of H_a, and thus I "fail to reject H_0.

Example:
Imagine you're flipping a coin n times and you're counting the number of heads thus $f : iidBern(\theta)$. You want to prove the coin is unfair.

$$H_a : \theta \neq 0.5$$

$$H_0 : \theta = 0.5$$

If H_0 is true, then θ = 0.5 and we can assume

$$\hat{\theta}_{MLE} = \dot{\sim} \mathbb{N}(\theta,\ SE[\hat{\theta}_{MLE}])$$

What constitutes sufficient evidence? Its a probability of rejecting when H_0 is true denoted by alpha, α. Everyone is different. If alpha is 5%, a 2 tailed test, we put half the probability in each tail. 5% is the most common scientific standard for alpha. Thus, in this case we retain the most common 95% of the thetahathats and reject H_0 for the 5% most weird thetahathats.

$$RET = [\hat{\hat{\theta}}_{MLE} \pm Z\frac{\alpha}{2}SE[\hat{\theta}_{MLE}]]$$

$$= [0.5 \pm 1.96\sqrt{fraq0.5(1-0.5)100}] = [0.402, 0.598]$$

If $\bar{x} = 0.60, h\hat{a}t\theta = 0.60 \notin RET$, we reject H_0 and conclude that there is enough evidence to say that the coin is being unfair.

If $\bar{x} = 0.59, h\hat{a}t\theta = 0.59 \in RET$, we fail to reject H_0 and conclude there's not enough evidence of coin being unfair.

This covers the frequentest approach to statistical inference. But there are problems with it. For instance, if $F : iidBern_\theta, x =< 0, 0, 0 >$. Therefore, $h\hat{a}t\theta = \bar{x} = 0$. Is this a good point estimate? No. You shouldn't be able to say something is absolutely impossible after n=3 trials.

To begin looking at the second problem with MLE, what if you had prior knowledge that Θ was restricted to [0.1, 0.2] and not the full interval (0, 1). You can't "enter" that into your inference. There is no place to put that information when using MLE's.

The third problem goes as so. Consider the frequentest interpretation of a confidence interval

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, for example, [0.37, 0.43], and you can't say,

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

   There is no randomness, so you can't make a probability statement of it, the real value of θ is either in the interval or not.

2. 95% of CI's will cover θ. But again, I only make one. So this interpretation doesn't help me. In conclusion, any specific CI means nothing.

3. 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

$$P(H_0|x) or P(H_a|x)$$

   or in English, you cannot find the probability of the null hypothesis or alternate hypothesis being true given the data. The only thing you can do is:

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

4.

$$F : iidBern_\theta, x =< 0, 1, 0 > \rightarrow \ \theta\hat{\hat{}}_{MLE} = 0.33$$

$$CI_{\theta,\ 0.95} = [-0.20, 0.87]$$

   Is this a good hypothesis test? No because you never reject. You can never say the coin is unfair. The problem is that the asymptotic normality of the MLE does not "kick in" until n is large (MLE property 2 is not true yet).

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