Section 9

Slides by Claris Winston

ADMINISTRIVIA

- hw 7 due on wednesday
- hw 8 released last night, due on March 6
    - the coding part is due on Friday, March 8
- (hw8 is the last 312 hw!!)

Content Review

CONTINUOUS LAW OF TOTAL PROBABILITY

This is the law of total probability for continuous random variables!

Instead of conditioning on a set of discrete events, and finding the probability of a discrete random variable, we might want to find the probability of continuous random variables, partitioning on the values in the range of a different continuous random variable.

$$\mathbb{P}(A) = \int_{x \in \Omega_X} \mathbb{P}(A|X = x) f_X(x) dx$$

LAW OF TOTAL EXPECTATION

Computed the expected value by partitioning on a set of events.

This is very similar in idea to law of total probability except that we're doing it for expectation.

$$\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[X|A_i]\,\mathbb{P}(A_i)$$

$$\mathbb{E}[X] = \sum_{y} \mathbb{E}[X|Y=y]\,p_Y(y)$$

CONTINUOUS LOTE

$$\mathbb{E}[X] = \int_{y\in\Omega_Y} \mathbb{E}[X|Y=y]\,f_Y(y)dy$$

Same thing as above, but for continuous RVs!

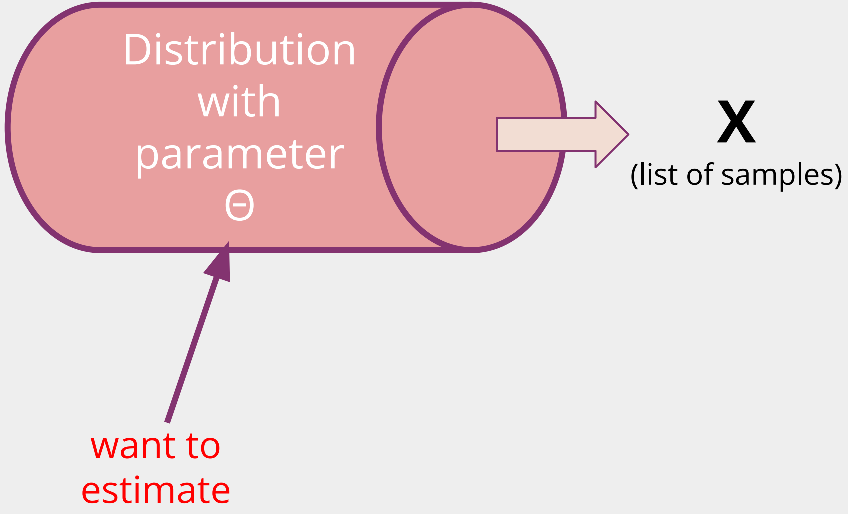
MLE

MLE

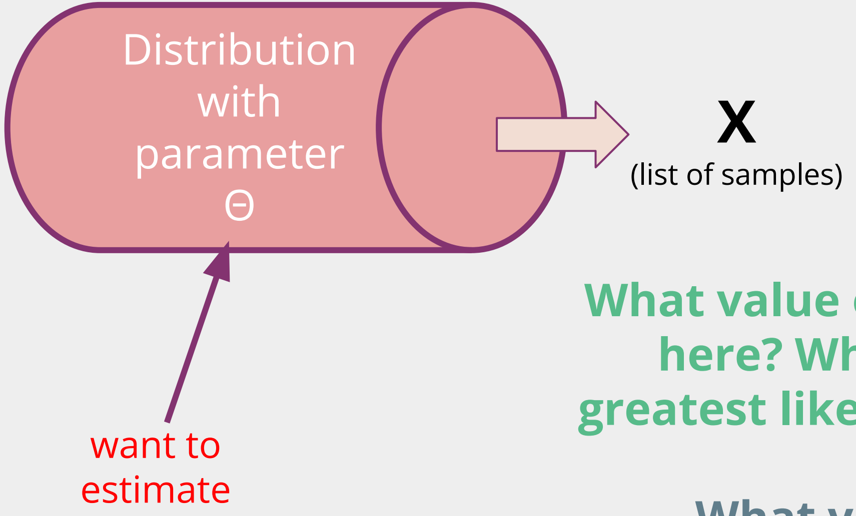
Sometimes, we don't know enough about a particular distribution, and we want to estimator a parameter to that distribution!

MLE helps us provide as <u>estimate</u> for some parameter(s) to a distribution based on some samples observed from that distribution

MLE


Distribution with parameter Θ
X
(list of samples)
want to estimate

MLE



What value of Θ would make the most sense here? What value of Θ would have the greatest likelihood of producing this sample?

What value of Θ maximizes L(X;Θ)?

MLE

Likelihood

Let x_1,...,x_n be iid samples from pmf p_X (x;θ) where θ are the distribution's parameters. The likelihood function is the probability of seeing the data given the parameters as

$$\mathcal{L}(x_1, \dots, x_n; \theta) = \prod_{i=1}^{n} p_X(x_i; \theta)$$

MLE

Likelihood

Let x_1,...,x_n be iid samples from pmf p_X (x;θ) where θ are the distribution's parameters. The likelihood function is the probability of seeing the data given the parameters as

$$\mathcal{L}(x_1, \ldots, x_n; \theta) = \prod_{i=1}^{n} p_X(x_i; \theta)$$

Represents the samples observed - can also just use X to represent the list of observations

multiplying bc we want the likelihood *all* of these are observed

in the *continuous case*, the PMF is replaced with the PDF

MLE

MLE

MLE of θ is \hat{θ} - the value of theta that will ***maximize*** the likelihood function

$$\hat{\theta} = \arg\max_{\theta} \mathcal{L}(x_1, \ldots, x_n; \theta)$$

We're given that there's a distribution with some unkown parameter(s) θ, and there are some observations x1,..,xn from this distribution

1. Write the likelihood function

even if there are multiple parameters to the distirbution, only one likelihood function - remember to multiply (not sum) the probability of each of the observations!

2. Take the log of the likelihood function (usually ln, not log)

we typically want to take natural log (ln) of the likelihood function in order to make finding the derivative much easier (remember ln of a product is a sum of the ln's)

3. Take the derivative of the log likelihood function

if you're finding an MLE where there are multiple parameter, in this step, take the partial derivative with respect to the parameter you're trying to solve for

4. Set the derivative of the log likelihood function to 0 and solve for θ

this step where you set the derivate equal to 0 is where you'll want to add the hat on top of the theta since at this step, we're solving for the maximum likelihood estimator

5. Verify the estimator is a maximizer via the 2nd derivate test (usually ignore for 312)

UNBIASED ESTIMATOR

Bias: The bias of an estimator $\hat{\theta}$ for a ***true*** parameter θ is defined as $E[\hat{\theta}] - \theta$. An estimator is unbiased iff $E[\hat{\theta}] = \theta$.