Quiz 1 Solution

CSE 4/574

Fall 2024

Question 1

Consider the problem discussed in the class about learning a generative model for a discrete random variable which emulates acoin toss. Modeling the random variable as a Bernoulli distribution with theta indicating the probability of observing a head, answer the following questions:

Correct Choice

For D = {H,T,H,H,H} and using the MLE approach, the probability of observing a tails will be 0.2.

Choice Explanation:
Let N_1 be the number of heads, and N_0 be the number of tails. Estimate θ in Bernoulli distribution:

$$\theta_{MLE,Tail} = \frac{N_0}{N} = \frac{1}{5} = 0.2$$

Incorrect Choice 1

If D = {H} (just one head), the MLE for theta will be 0.

Choice Explanation:

$$\theta_{MLE} = \frac{N_1}{N} = \frac{1}{1} = 1$$

Incorrect Choice 2

If D = {H} (just one head), the MLE for theta will be 0.5.

Choice Explanation:
Refer to incorrect choice 1 explanation.

Incorrect Choice 3

For D = {H,T,H,H,H} and using the MLE approach, the probability of observing a heads will be 0.4.

Choice Explanation:

$$\theta_{MLE} = \frac{N_1}{N} = \frac{4}{5} = 0.8$$

Question 2

In this problem you will use the prior and MLE estimates of a univariate Gaussian distribution to assign probability of a new instance to be in one of two classes (each class modeled as a Gaussian distribution). You will need to use the discussion in class about MLE for Gaussian distribution. You can also use the shared notebook RandomVariablesIntroduction.ipynb to compute the exact values for the probability density. Please note that the norm function takes the standard deviation (and not the variance) as a parameter.
Consider the example of analyzing weights (at-birth) of human babies (female) for different cities. WHO estimates state that the weights of girls born in Mexico City are normally distributed with mean weight of 4.4 lbs and standard deviation of 0.5 lbs. The weights of girls born in Boston are also normally distributed with mean weight of 5 lbs and standard deviation of 0.7 lbs.

A recent sample from hospitals in both cities is shown below. Each row corresponds to weights of 10 girls born in each city in the last 1 day.

Mexico City: $4.72, 4.15, 4.46, 4.23, 5.28, 4.07, 4.67, 4.87, 4.54, 3.75$
Boston: $5.76, 5.12, 5.64, 5.59, 5.54, 5.09, 5.46, 5.87, 5.71, 5.56$

Given the prior WHO estimates and the new data, your task is to infer the city of birth for two newborn girls born with following weights: Baby A - 4.6 lbs and Baby B - 4.9 lbs.
Hint: You need to compare the probability density for each baby's weight under two distributions for each city (class): (1) Using the prior WHO estimates, and (2) Using the MLE estimates using the observed data.

Correct Choice

Baby B is more likely to be born in Boston by WHO estimates but is more likely to be born in Mexico City by the new data.

Choice Explanation:
By WHO estimates:
PDF for $Normal(5, 0.7)$ at $x = 4.9$ is approximately 0.56 (Boston)
PDF for $Normal(4.4, 0.5)$ at $x = 4.9$ is approximately 0.48 (Mexico)

Similarly for the new data, calculate the mean, standard deviation or variance for both Mexico City and Boston.
Mexico City: mean = 4.474, Standard Deviation = 0.442
Boston: mean = 5.534, Standard Deviation = 0.255

By new data:
PDF for $Normal(5.534, 0.255)$ at $x = 4.9$ is approximately 0.07 (Boston)
PDF for $Normal(4.474, 0.442)$ at $x = 4.9$ is approximately 0.57 (Mexico)

Incorrect Choice 1

Both babies are more likely to be born in Mexico city by WHO estimates.

Choice Explanation:
We already know that Baby B is more likely to be born in Boston.
For Baby A:
PDF for $Normal(5, 0.7)$ at $x = 4.6$ is approximately 0.48 (Boston)
PDF for $Normal(4.4, 0.5)$ at $x = 4.6$ is approximately 0.74 (Mexico)
Therefore, only Baby A is more likely to be born in Mexico City by WHO estimates.

Incorrect Choice 2

None of the two babies are more likely to be born in Mexico city by using the new data.

Choice Explanation:
For baby A:
PDF for $Normal(5.534, 0.255)$ at $x = 4.6$ is approximately 0.002 (Boston)
PDF for $Normal(4.474, 0.442)$ at $x = 4.6$ is approximately 0.87 (Mexico)
By new data, baby A is more likely to be born in Mexico city.

Incorrect Choice 3

Baby B is more likely to be born in Boston according to WHO estimates and new data.

Choice Explanation:
According to explanation for correct choice, by new data, Baby B is more likely to be born in Mexico City.