1 Multiple Choice 1 point

How many parameters are needed to fully specify a Gaussian distribution?

- [ ] 1
- [x] 2
- [ ] 3
- [ ] 5

2 Multiple Choice 1 point

Given a statistical model and some observations, which of the following is NOT consistent with following the **maximum likelihood principle** as a way to select the model parameters?

- [ ] Maximize the probability of observations under the model parameters
- [ ] Maximize the logarithm of the probability of observations under the model parameters
- [x] Minimize the logarithm of the probability of observations under the model parameters
- [ ] Minimize the negative logarithm of the probability of observations under the model parameters

3 Multiple Choice 1 point

In a linear model setup with Gaussian errors, maximum likelihood estimation turns out to be equivalent to loss minimization using which loss function?

- (●) Squared error loss
- ( ) Absolute error loss
- ( ) Cross entropy loss
- ( ) Any non-negative loss function

4 Multiple Choice 1 point

The Laplace distribution with mean m and variance $2b^2$ has the probability density function:

$p(x) = 1/(2b) \exp(- |x - m| / b)$

where $|x|$ is the absolute value of x. Consider a linear model with Laplace error distribution with mean zero and variance $2b^2$. Maximum likelihood estimation in this model would be equivalent to loss minimization using which loss function?

- (●) Absolute error loss
- ( ) Squared error loss
- ( ) Cross entropy loss
- ( ) Any non-negative loss function

5 Multiple Choice 1 point

The input to output mapping in linear regression corresponds to a very simple neural network. How many layers does this simple neural network have?

- [x] 1
- [ ] 2
- [ ] Depends on the dimensions of the input features
- [ ] Depends on the number of examples in the training data set

6 Multiple Choice 1 point

Which of the following best describes the relationship between an artificial neuron and a biological neuron?

- [ ] The word "neuron" is present in both by sheer accident. They have nothing to do with each other.
- [x] An artificial neuron is an extremely simplified version of the processing a biological neuron is capable of
- [ ] An artificial neuron is an extremely precise version of the processing a biological neuron is capable of
- [ ] Real biological neurons don't compute anything whereas artificial ones do

7 Multiple Choice 1 point

Which of the following is NOT a classification problem?

- Predicting whether or not a bank customer will fail to pay the loan on time
- Assessing whether or not a credit card transaction is fraudulent
- **Predicting how long a newly admitted patient will stay at the hospital**
- Labeling a bird image with the scientific name of the species the bird belongs to

8 Multiple Choice 1 point

Consider a simple statistical model with a single parameter p which models the probability that a coin, when tossed, will land with HEADS up instead of TAILS. To get an observation, you pull a coin from your pocket and toss it once. It comes up HEADS. What is the maximum likelihood estimate of p given your observation?

- 0
- **1**
- 1/2
- Cannot be written down in any simple form. We will have to run an optimization algorithm to get the answer.

9 Multiple Choice 1 point

What is cross entropy?

- It is a neural network architecture used in classification problems
- It is a neural network architecture used in regression problems
- It is a loss function used in regression problems
- **It is a loss function used in classification problems**

10 Multiple Choice 1 point

Consider a classification problem with 3 classes. Which of the following vectors CANNOT be the output of the softmax function?

- **(0.8, 0.1, -0.1)**
- (0.4, 0.2, 0.4)
- (0.8, 0.19, 0.01)
- (0.3, 0.3, 0.4)