1

Suppose that the lifetime of Wipro light bulbs is modelled by an exponential distribution with (unknown) parameter λ. We test 4 bulbs and find they have lifetimes of 3, 2, 1, and 4 years, respectively. What is the MLE for λ?
Note: The probability density function (pdf) of an exponential distribution is

$$f(x;\lambda) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0, \\ 0 & x < 0. \end{cases}$$

- **A**. 4/5
- **B**. 3/5
- **C RIGHT**. 2/5
- **D**. 1/5

The maximum likelihood estimate for the rate parameter is:

$$\widehat{\lambda} = \frac{1}{\overline{x}} = \frac{n}{\sum_i x_i}$$

See MLE in Exponential distribution

-

2

Let $X_1, X_2, X_3, \ldots, X_n$ be a random sample from a $Uniform\,(a,\ b)$ distribution, where a and b are unknown. Select the correct maximum likelihood estimator (MLE) of a and b based on this random sample.
Note: The probability density function of the continuous uniform distribution is:

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{for } a \leq x \leq b, \\ 0 & \text{for } x < a \text{ or } x > b \end{cases}$$

- **A RIGHT**.

$$\widehat{a} = min(X_1, \ldots, X_n)\ \ \widehat{b} = max(X_1, \ldots, X_n)$$

- **B**.

$$\widehat{a} = max(X_1, \ldots, X_n)\ \ \widehat{b} = min(X_1, \ldots, X_n)$$

- **C**.

$$\widehat{a} = min(X_1, \ldots, X_n)\ \ \widehat{b} = min(X_1, \ldots, X_n)$$

- **D**.

$\hat{a} = max(X_1, ..., X_n)$ $\hat{b} = max(X_1, ..., X_n)$

3 For any particular problem, maximizing the likelihood function always leads to

- **A**. High Bias
- **B RIGHT**. Over-fitting
- **C**. Simple models
- **D**. None of the above

4

Given training data pairs (X, Y) and parameters w and the probability distribution of any random variable A denoted by p(A), which of the following statements are True with regard to Maximum A Posterior (MAP) estimation.

- **A RIGHT**.

MAP estimation depending on the choice of prior is the equivalent of a regularized maximum likelihood model

- **B**.

To find the MAP estimate of w, we find the value of w that maximizes p(y|x,w)p(x).

- **C**.

To find the MAP estimate of w we find the value of w that maximizes p(y|x,w).

- **D**.

To find the MAP estimate we need to find the value of x that maximizes p(y|x,w).p(w)

Accepted Answers: A and D

Option-D is the surprise for the week :P

---

5. What does an auto-encoder learn about the data?

- **A**. High dimensional representation of the data
- **B RIGHT**. Low dimensional representation of the data
- **C**. Average dimensional representation of the data
- **D**. No representation of the data is learned

6. Auto-encoders are able to compress the input data in its hidden representation if:

- **A RIGHT**. If the input features are correlated
- **B**. If the input features are not correlated
- **C**. If the input features are independent
- **D**. If the input features are unrelated

7. Generative Adversarial models are:

- **A RIGHT**. A. Generative models
- **B**. B. Discriminative models
- **C**. Both A and B
- **D**. None of the above.

8. The main objective of Bayesian linear regression is

- **A**. A. To find a single best value of model parameters.
- **B RIGHT**. B. To determine the posterior distribution for the model parameters
- **C**. Both A and B
- **D**.None of the above.

9.

$$P(W|y,X) = \frac{P(y|W,X) \times P(W|X)}{P(y|X)}$$

W, y, X have usual meaning.
Which of the following is/are True?

- **A RIGHT**.

$$P(W|y,X) = Posterior$$

- **B RIGHT**.

$$P(y|W,X) = Likelihood$$

- **C**.

$$P(W|X) = Normalization$$

- **D**.

$$P(y|X) = prior$$

10. Which of the following are true about variational auto-encoders?

- **A**. They are a form of denoising autoencoders
- **B**. They require a random number as input during training
- **C**. It's called variational because the output varies constantly
- **D RIGHT**. KL divergence is one of the terms in the loss function