Quiz 2

- Due Feb 2 at 11:59pm
- Points 10
- Questions 10
- Available Jan 29 at 10am - Feb 2 at 11:59pm
- Time Limit 60 Minutes

Instructions

General Instructions:

- The quiz contains 10 multiple choice questions. You have 1 hour to finish it. Once submitted, you cannot re-take the quiz.
- The syllabus for this quiz are all the content covered last week in the Monday (01/22) and Wednesday (01/24) lectures.
- You are allowed to consult is lecture slides and discussion notes, which you can download in advance and refer to if helpful. No other online or offline resource is permitted.
- The quiz is open till 11:59pm on Friday, Feb 02 2024. There are no late submissions allowed.
- Please follow the UCLA honor code. Any evidence of sharing questions and answers relating to the quiz with other students will lead to an immediate F grade. You are also barred from posting any questions relating to the quizzes on Campuswire until the deadline for submitting the quiz has passed.

This quiz was locked Feb 2 at 11:59pm.

Attempt History

	Attempt	Time	Score
LATEST	Attempt 1	30 minutes	8 out of 10

Score for this quiz: 8 out of 10

Submitted Feb 2 at 7:03pm

This attempt took 30 minutes.

Question 1

1 / 1 pts

Which of the following statements about Kullback-Leibler (KL) divergence is true?

Correct!

- [x] $D_{KL}(p, q)$ is undefined when the support of p is strictly greater than the support of q

If the support of p is strictly greater than the support of q, then there exists x for which $p(x) > 0$ and $q(x) = 0$, which will lead to an undefined expression in the log.

- [ ] KL divergence is always symmetric
- [ ] KL divergence is only defined for discrete probability distributions
- [ ] $D_{KL}(p, q)$ is undefined when the support of q is strictly greater than the support of p

Question 2

1 / 1 pts

When does the KL divergence between two distributions become zero?

- [ ] When the two probability distributions have disjoint supports
- [ ] When the entropy of the first distribution is zero

Correct!

- [x] When the two probability distributions are identical
- [ ] When the entropy of the second distribution is zero

Question 3

1 / 1 pts

The expected log-likelihood of a model p_θ under a data distribution p_{data} is equal to the KL divergence between p_{data} and p_θ.

Correct!

- [x] False

$$D_{KL}(p_{\text{data}}, p_\theta) = \mathbb{E}_{p_{\text{data}}}[\log p_{\text{data}}] - \mathbb{E}_{p_{\text{data}}}[\log p_\theta]$$

- [ ] True

Question 4

0 / 1 pts

Given an input $\mathbf{x}$, the conditionals in GPT can be evaluated in parallel for density estimation.

Correct Answer

- True

You Answered

- False

Question 5

1 / 1 pts

Given an input $\mathbf{x}$, the conditionals in RNN can be evaluated in parallel for density estimation.

- True

Correct!

- False

False, the hidden state for current timestep depends on the hidden state for the previous timestep creating a sequential dependency.

Question 6

1 / 1 pts

For generating a new sample $\mathbf{x}$, the conditionals in GPT can be sampled in parallel.

Correct!

- False

False, sampling a conditional in autoregressive models depend on the previous sampled token creating a sequential dependency.

- True

Question 7

1 / 1 pts

For generating a new sample $\mathbf{x}$, the conditionals in RNN can be sampled in parallel.

True

Correct!

False

False, sampling a conditional in autoregressive models depend on the previous sampled token creating a sequential dependency.

Question 8

0 / 1 pts

For two arbitrary random vectors $\mathbf{x}, \mathbf{z}$, there does not exist a joint distribution p such that $p(\mathbf{x}) < p(\mathbf{x}, \mathbf{z})$.

You Answered

False

Correct Answer

True

Question 9

1 / 1 pts

For any arbitrary random vector $\mathbf{x}$, we have $\mathbb{E}_{\mathbf{x}\sim p}[f(\mathbf{x})] = \mathbb{E}_{\mathbf{x}\sim q}\left[\frac{f(\mathbf{x})}{q(\mathbf{x})}\right]$.

Correct!

False

False, by definition of expectation, $\mathbb{E}_{\mathbf{x}\sim p} = \mathbb{E}_{\mathbf{x}\sim q}\left[\frac{p(\mathbf{x})f(\mathbf{x})}{q(\mathbf{x})}\right]$.

True

Question 10

1 / 1 pts

If an autoregressive model assigns high likelihoods to a training set of images, it will also necessarily assign high likelihoods to unseen examples from the same distribution.

- True

Correct!

- False

False, the model could have overfitted to the training set.

Quiz Score: 8 out of 10