Quiz 1

- Due Jan 26 at 11:59pm
- Points 10
- Questions 10
- Available Jan 22 at 4pm - Jan 26 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 Discussion 1 (basic probability review) and Lecture 2 (Autoregressive Models, everything before RNNs).
- 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, Jan 26 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 Jan 26 at 11:59pm.

Attempt History

	Attempt	Time	Score
LATEST	Attempt 1	25 minutes	8 out of 10

Score for this quiz: 8 out of 10

Submitted Jan 26 at 12:17pm

This attempt took 25 minutes.

Question 1

1 / 1 pts

Consider a collection of n discrete random variables $\{X_i\}_{i=1}^n$, where the number of outcomes for X_i is $|\mathbf{val}(X_i)| = k_i$. Under the full independence assumption (i.e., every variable is independent of every other variable), what is the total number of parameters needed to describe the joint distribution over $(X_1, \ldots, X_n)$?

- $(\sum_{i=1}^n k_i) - 1$
- $(\prod_{i=1}^n k_i) - 1$
- n

Correct!

- $\sum_{i=1}^n (k_i - 1)$

There are $\prod_{i=1}^n k_i$ unique configurations. With full independence assumption, the number of independent parameters needed is $\sum_{i=1}^n (k_i - 1)$.

Question 2

1 / 1 pts

Consider a collection of n discrete random variables $\{X_i\}_{i=1}^n$, where the number of outcomes for X_i is $|\mathbf{val}(X_i)| = k_i$. Without any (conditional) independence assumptions, what is the total number of parameters needed to describe the joint distribution over $(X_1, \ldots, X_n)$?

Correct!

- $(\prod_{i=1}^n k_i) - 1$

There are $\prod_{i=1}^n k_i$ unique configurations. Without independence assumptions, the number of independent parameters needed is $(\prod_{i=1}^n k_i) - 1$.

- $(\sum_{i=1}^n k_i) - 1$
- n

- $\sum_{i=1}^{n}(k_i - 1)$

Question 3

0 / 1 pts

Which of the following signifies a valid factorization for an autoregressive generative model over an input $\mathbf{x} \in \mathbb{R}^3$?

Correct Answer

- $p(\mathbf{x}) = p(x_1|x_2, x_3)p(x_2)p(x_3|x_2)$

- $p(\mathbf{x}) = p(x_1|x_2, x_3)p(x_2|x_1, x_3)p(x_3|x_1, x_2)$

You Answered

- $p(\mathbf{x}) = p(x_1)p(x_2|x_1)p(x_3|x_2)$

- $p(\mathbf{x}) = p(x_1)p(x_2)p(x_3)$

Question 4

1 / 1 pts

Consider a NADE generative model over a 3D input $\mathbf{x} \in \mathbb{R}^3$. Each $x_i \in \{0, 1\}$ is binary. The first marginal $p(x_1)$ is simply a Bernoulli distribution requiring 1 parameter. Each conditional $p(x_i \mid x_{<i})$ for $i > 1$ is parameterized via a 1 hidden layer neural network of dimensionality 10. Assume no bias parameters for the hidden or the output layers. With **no sharing** of parameters across the weight matrices for the two conditionals, what is the total number of learnable parameters for NADE?

Correct!

- 51

51. 1 (for x_1) + 10 (hidden layer for x_2) + 20 (hidden layer for x_3) + 2x10 (output layer for x_2, x_3).

- 31

- 41

- 21

Question 5

0 / 1 pts

Consider a NADE generative model over a 3D input $\mathbf{x} \in \mathbb{R}^3$. Each $x_i \in \{0, 1\}$ is binary. The first marginal $p(x_1)$ is simply a Bernoulli distribution requiring 1 parameter. Each conditional $p(x_i \mid x_{<i})$ for $i > 1$ is parameterized via a 1 hidden layer neural network of dimensionality 10. Assume no bias parameters for the hidden or the output layers. With **sharing** of parameters, what is the total number of learnable parameters for NADE?

Correct Answer

- 41

You Answered

- 31
- 21
- 51

Question 6

1 / 1 pts

If X is a random variable with a mean of μ and variance σ^2, what is the variance of the random variable $Y = aX + b$, where a and b are constants?

Correct!

- $a^2\sigma^2$
- $a^2\sigma^2 + b^2$
- $a\sigma^2 + b$
- σ^2

Question 7

1 / 1 pts

Which of the following scenarios is LEAST likely to be modeled accurately by a Gaussian distribution?

- Scores on a well-designed math test.

Correct!

- Income distribution in a highly skewed economy.

Income distribution in a highly skewed economy will not be symmetric around the mean. This should follow an exponential distribution.

- Heights of adult individuals in a large population.
- Errors in measurements made by a precise sensor.

Question 8

1 / 1 pts

Let X be a continuous random variable with a probability density function $f(x)$. What is the expected value of a function $g(X)$?

Correct!

- $\int g(x)f(x)dx$ over the range of X.
- $\int xg(f(x))dx$ over the range of X.
- $\int g(x)dx$ over the range of X.
- $g(\int xf(x)dx)$ over the range of X.

Question 9

1 / 1 pts

If the covariance between two random variables X and Y is zero, what can we conclude?

- ○ X and Y are independent.

Correct!

- ◉ X and Y have no linear relationship.

- ○ The variance of X and Y is zero.

- ○ The mean of X and Y is zero.

Question 10

1 / 1 pts

A medical test for a disease has a 95% probability of giving a positive result when the person actually has the disease and a 5% probability of giving a positive result when the person does not have the disease. If 1% of the population actually has the disease, what is the probability that a person has the disease given that they have tested positive?

- ○ Approximately 64%
- ○ Approximately 32%
- ○ Approximately 48%

Correct!

- ◉ Approximately 16%

Approximately 16%. $P(D \mid P) = \frac{P(P|D)P(D)}{P(P|D)P(D)+P(P|\bar{D})P(\bar{D})} = \frac{\mathbf{0.95 \times 0.01}}{\mathbf{0.95 \times 0.01 + 0.05 \times 0.99}} \approx \mathbf{0.16}$.

Quiz Score: 8 out of 10