UNIVERSITY OF ALBERTA
CMPUT 267 Fall 2024

Midterm Exam 1

Do Not Distribute

Duration: 75 minutes

Last Name: ______________________________________

First Name: ______________________________________

Carefully read all of the instructions and questions. Good luck!

---

1. **Do not turn this page** until instructed to begin.
2. Ensure that your exam package contains 18 pages.
3. **Only the scantron will be marked**. All of your answers must be clearly marked on the scantron.
4. Use **pencil only** to fill out the scantron (preferably an HB or #2 pencil).
5. **Erase mistakes completely** on the scantron to avoid misreading by the scanner.
6. **Mark answers firmly and darkly**, filling in the bubbles completely.
7. This exam consists of **25 questions**. Each question is worth **1 mark**. The exam is worth a total of **25 marks**.
8. Some questions may have **multiple correct answers**. To receive **full marks**, you must select **all correct answers**. If you select only **some** of the correct answers, you will receive **partial marks**. Selecting an incorrect option will cancel out a correct one. For example, if you select two answers—one correct and one incorrect—you will receive zero points for that question. If the number of incorrect answers exceeds the correct ones, your score for that question will be zero. **No negative marks** will be given.

Question 1. [1 MARK]

Which of the following represents the set of all linear functions of the form $f(x) = wx + b$ where $w, b \in \mathbb{R}$?

A. $\{f \mid f : \mathbb{R} \to \mathbb{R},\ f(x) = wx + b,\ w, b \in \mathbb{R}\}$

B. $\{f \mid f : \mathbb{R} \to \mathbb{R},\ f(x) = w^2b^2x,\ w, b \in \mathbb{R}\}$

C. $\{f \mid f : \mathbb{R} \to \mathbb{R},\ f(x) = wx,\ w \in \mathbb{R}\}$

D. $\{f \mid f : \mathbb{R} \to \mathbb{R},\ f(x) = b,\ b \in \mathbb{R}\}$

Solution 1.

Correct Answer: A.
Explanation:

A. Correctly defines the set of all functions of the form $f(x) = wx + b$ with $w, b \in \mathbb{R}$.

B. Incorrect because it defines functions of the form $f(x) = w^2b^2x+$, which are not linear functions $f(x) = wx + b$.

C. Represents only functions of the form $f(x) = wx$, missing the bias term b.

D. Represents constant functions $f(x) = b$, lacking the wx term.

Question 2. [1 MARK]

Which of the following is a valid definition of a function $\mathcal{A} : (\mathbb{R} \times \mathbb{R})^2 \to \{f \mid f : \mathbb{R} \to \mathbb{R}\}$?

A. $\mathcal{A}((a, b), (c, d)) = f$, where $f(x) = ax + b$

B. $\mathcal{A}((a, b), (c, d)) = f$, where $f(x) = cx + d$

C. $\mathcal{A}((a, b), (c, d)) = f$, where $f(x) = a + bx$

D. $\mathcal{A}((a, b), (c, d)) = f$, where $f(x) = ax - b$

Solution 2.

Correct Answers: A., B., C., D.
Explanation:

A. $f(x) = ax + b$ is a valid linear function.

B. $f(x) = cx + d$ is also a valid linear function.

C. $f(x) = a + bx$ rearranges the terms but still represents a linear function.

D. $f(x) = ax - b$ is another form of a linear function.

Therefore, all options correctly describe functions in the codomain.

Question 3. [1 MARK]

Given $\mathcal{D} = ((x_1, y_1), \ldots, (x_n, y_n))$, the loss function $\ell(\hat{y}, y) = |\hat{y} - y|$, and the predictor $f(x) = wx + b$, which of the following represents the total loss calculated by summing $\ell(f(x), y)$ over all data points $(x, y) \in \mathcal{D}$?

A. $\sum_{i=1}^{n} |wx_i + b - y_i|$

B. $\sum_{(x,y)\in\mathcal{D}} |wx + b - y|$

C. $\sum_{i=1}^{n} (wx_i + b - y_i)^2$

D. $\prod_{(x,y)\in\mathcal{D}} |wx + b - y|$

Solution 3.

Correct Answers: A., B.
Explanation:

A. $\sum_{i=1}^{n} |wx_i + b - y_i|$ correctly sums the absolute losses for each data point.

B. $\sum_{(x,y)\in\mathcal{D}} |wx + b - y|$ is another valid notation for the same sum.

C. Uses squared loss instead of absolute loss, which does not match the given loss function.

D. Represents the product of losses, not the sum.

Question 4. [1 MARK]

Let $g(x, y) = x^2 + y$ where $x, y \in \mathbb{R}$. What is

$$\int_{\mathcal{X}} \sum_{y \in \mathcal{Y}} g(x, y)\, dx$$

where $\mathcal{Y} = \{1, 2, 3\}$ and $\mathcal{X} = [0, 2]$?

A. 15

B. 18

C. 20

D. 22

Solution 4.

Correct Answer: C. 20
Explanation:

A. First, compute the sum inside the integral:

$$\sum_{y \in \mathcal{Y}} g(x, y) = \sum_{y=1,2,3} (x^2 + y) = 3x^2 + (1 + 2 + 3) = 3x^2 + 6$$

B. Next, integrate this expression over $\mathcal{X} = [0, 2]$:

$$\int_0^2 (3x^2 + 6)\, dx = 3\int_0^2 x^2\, dx + 6\int_0^2 dx$$

C. Compute each integral separately:

$$3\int_0^2 x^2\, dx = 3\left[\frac{x^3}{3}\right]_0^2 = [x^3]_0^2 = 2^3 - 0^3 = 8 - 0 = 8$$

$$6\int_0^2 dx = 6[x]_0^2 = 6(2-0) = 6 \times 2 = 12$$

D. Add the results of the two integrals:

$$8 + 12 = 20$$

E. Therefore, the value of the integral is 20, which corresponds to option C.

Question 5. [1 MARK]

Let $h(x, y) = x + y$ where $x, y \in \mathbb{R}$. What is

$$\int_{\mathcal{X}} \int_{\mathcal{Y}} h(x, y)\, dy\, dx$$

where $\mathcal{Y} = [2, 4]$ and $\mathcal{X} = [1, 3]$?

A. 18

B. 20

C. 22

D. 24

Solution 5.

Correct Answer: B. 20
Explanation:

A. Begin by computing the inner integral with respect to y:

$$\int_2^4 h(x, y)\, dy = \int_2^4 (x + y)\, dy = x\int_2^4 dy + \int_2^4 y\, dy$$

B. Evaluate each part:

$$x\int_2^4 dy = x[y]_2^4 = x(4-2) = 2x$$

$$\int_2^4 y\, dy = \left[\frac{y^2}{2}\right]_2^4 = \frac{4^2}{2} - \frac{2^2}{2} = \frac{16}{2} - \frac{4}{2} = 8 - 2 = 6$$

C. Combine the results:

$$\int_2^4 (x+y)\,dy = 2x+6$$

D. Now, integrate this expression with respect to x over $\mathcal{X} = [1,3]$:

$$\int_1^3 (2x+6)\,dx = 2\int_1^3 x\,dx + 6\int_1^3 dx$$

E. Evaluate each integral:

$$2\int_1^3 x\,dx = 2\left[\frac{x^2}{2}\right]_1^3 = [x^2]_1^3 = 3^2 - 1^2 = 9 - 1 = 8$$

$$6\int_1^3 dx = 6[x]_1^3 = 6(3-1) = 6 \times 2 = 12$$

F. Add the results of the two integrals:

$$8 + 12 = 20$$

G. Therefore, the value of the double integral is 20, which corresponds to option B.

Question 6. [1 MARK]

Let X be a discrete random variable uniformly distributed over the outcome space $\mathcal{X} = \{1,2,3,4,5\}$. The probability mass function (pmf) of X is given by $p(x) = \frac{1}{5}$ for each $x \in \mathcal{X}$. Which of the following statements are true?

A. The probability that X is less than 4 is 0.5.

B. The expected value $\mathbb{E}[X] = 3$.

C. The probability that X is an even number is 0.4.

D. The probability that X is greater than or equal to 4 is 0.4.

Solution 6.

Correct Answers: B., C., D.
Explanation:

A. The probability that X is less than 4 is 0.5. This is false because $\mathbb{P}(X < 4) = \mathbb{P}(1,2,3) = \frac{3}{5} = 0.6$.

B. The expected value $\mathbb{E}[X] = 3$. This is true because $\mathbb{E}[X] = \frac{1+2+3+4+5}{5} = 3$.

C. The probability that X is an even number is 0.4. This is true because $\mathbb{P}(X \text{ even}) = \mathbb{P}(2,4) = \frac{2}{5} = 0.4$.

D. The probability that X is greater than or equal to 4 is 0.4. This is true because $\mathbb{P}(X \geq 4) = \mathbb{P}(4,5) = \frac{2}{5} = 0.4$.

Question 7. [1 MARK]

Suppose you roll two fair six-sided dice. Let the first die be represented by the random variable $D_1 \in \{1, 2, 3, 4, 5, 6\}$ and the second die by $D_2 \in \{1, 2, 3, 4, 5, 6\}$. Which of the following sets represents the outcome space of the random variable $D = (D_1, D_2)$?

A. $\{(d_1, d_2) \mid d_1, d_2 \in \{1, 2, 3, 4, 5, 6\}\}$

B. $\{2, 3, 4, \ldots, 12\}$

C. $\{1, 2, 3, 4, 5, 6\}$

D. $\{d_1 + d_2 \mid d_1, d_2 \in \{1, 2, 3, 4, 5, 6\}\}$

Solution 7.

Correct Answer: A.
Explanation:

A. Option A correctly represents all possible ordered pairs (d_1, d_2) where each die can land on any number from 1 to 6.

B. Option B represents the possible sums of the two dice, not the ordered pairs themselves.

C. Option C represents the outcomes of a single die, not the pair.

D. Option D also represents the sums of the dice, similar to Option B.

Therefore, Option A is the correct outcome space of $D = (D_1, D_2)$.

Question 8. [1 MARK]

Suppose you randomly select a letter from the English alphabet. The outcome space is $\mathcal{X} = \{\mathrm{A}, \mathrm{B}, \mathrm{C}, \ldots, \mathrm{Z}\}$. Which of the following is an event?

A. $\{\mathrm{A}, \mathrm{E}, \mathrm{I}, \mathrm{O}, \mathrm{U}\}$

B. $\{\mathrm{AA}, \mathrm{BB}\}$

C. M

D. $\{\mathrm{B}, \mathrm{C}, \mathrm{D}, \mathrm{F}, \mathrm{G}\}$

Solution 8.

Correct Answers: A., D.
Explanation:

A. Option A is a subset of the outcome space $\mathcal{X}$, thus it is an event.

B. Option B contains elements not in $\mathcal{X}$ (e.g., "AA"), so it is not an event in this context.

C. Option C is a single element without braces, making it ambiguous whether it's an event (set) or an element.

D. Option D is also a subset of $\mathcal{X}$, so it is an event.

Therefore, Options A and D are valid events.

Question 9. [1 MARK]

Suppose you have a random variable X representing the time (in minutes) it takes to commute to work. You know X is distributed according to the continuous uniform distribution over the interval $\mathcal{X} = [30, 60]$. Let p be the pdf of X. Which of the following statements are true?

A. $p(45)$ is the probability that $X = 45$.

B. $p(x)$ represents the probability density at point x.

C. The probability that X is between 40 and 50 is $\frac{1}{3}$.

D. The probability that $X = 30$ is $\frac{1}{30}$.

Solution 9.

Correct Answers: B., C.
Explanation:

A. False. For continuous random variables, $P(X = x) = 0$ for any specific x. So $p(45)$ is not the probability that $X = 45$; it is the probability density at $x = 45$.

B. True. $p(x)$ gives the probability density at point x.

C. True. The probability is calculated as:

$$P(40 \le X \le 50) = \frac{50 - 40}{60 - 30} = \frac{10}{30} = \frac{1}{3}.$$

D. False. $P(X = 30) = 0$ for continuous variables.

Question 10. [1 MARK]

Suppose that $Y \in \mathbb{R}$ is distributed according to Normal$(0, 1)$. Which of the following statements are true?

A. The probability density function of Y is $p(y) = \frac{1}{\sqrt{2\pi}} e^{-y^2/2}$ for $y \in \mathbb{R}$.

B. The probability that $Y = 0$ is $\frac{1}{\sqrt{2\pi}}$.

C. Y is a continuous random variable.

D. The variance of Y is 1.

Solution 10.

Correct Answers: A., C., D.
Explanation:

A. True. This is the probability density function (pdf) of the standard normal distribution.

B. **False.** For continuous random variables, the probability of any specific value is zero, so $P(Y=0)=0$, not $\frac{1}{\sqrt{2\pi}}$.

C. True. Y is continuous because it can take any real value within $\mathbb{R}$.

D. True. The variance of a Normal$(0,1)$ distribution is 1.

Question 11. [1 MARK]

Suppose you have two discrete random variables $X \in \{0,1,2\}$ and $Y \in \{1,2\}$. The joint probability mass function (pmf) of X and Y is given by the following values:

$$\begin{aligned} p(0,1) &= \frac{1}{12}, \quad p(0,2) = \frac{1}{12} \\ p(1,1) &= \frac{1}{4}, \quad p(1,2) = \frac{1}{6} \\ p(2,1) &= \frac{1}{6}, \quad p(2,2) = \frac{1}{4} \end{aligned}$$

Which of the following is the marginal pmf of X?

A. $p_X(0) = \frac{1}{6}, \quad p_X(1) = \frac{5}{12}, \quad p_X(2) = \frac{5}{12}$

B. $p_X(0) = \frac{1}{6}, \quad p_X(1) = \frac{1}{3}, \quad p_X(2) = \frac{1}{2}$

C. $p_X(0) = \frac{1}{3}, \quad p_X(1) = \frac{1}{2}, \quad p_X(2) = \frac{1}{6}$

D. $p_X(0) = \frac{1}{2}, \quad p_X(1) = \frac{1}{4}, \quad p_X(2) = \frac{1}{4}$

Solution 11.

Correct Answer: A.
Explanation:
For $X=0$:

$$p_X(0) = p(0,1) + p(0,2) = \frac{1}{12} + \frac{1}{12} = \frac{2}{12} = \frac{1}{6}.$$

For $X=1$:

$$p_X(1) = p(1,1) + p(1,2) = \frac{1}{4} + \frac{1}{6} = \frac{3}{12} + \frac{2}{12} = \frac{5}{12}.$$

For $X=2$:

$$p_X(2) = p(2,1) + p(2,2) = \frac{1}{6} + \frac{1}{4} = \frac{2}{12} + \frac{3}{12} = \frac{5}{12}.$$

Question 12. [1 MARK]

Let X_1, X_2, X_3 be independent and identically distributed random variables representing three coin flips, where each coin lands heads (1) with probability 0.6 and tails (0) with probability 0.4. What is the probability that all three flips result in heads or all three flips result in tails?

A. 0.216

B. 0.064

C. 0.28

D. 0.512

Solution 12.

Correct Answer: C. 0.28
Explanation:

A. Calculate the probability of all three heads:

$$\mathbb{P}(\{(H, H, H)\}) = (0.6)^3 = 0.216$$

B. Calculate the probability of all three tails:

$$\mathbb{P}(\{(T, T, T)\}) = (0.4)^3 = 0.064$$

C. Sum the probabilities:

$$\mathbb{P}(\{(H, H, H), (T, T, T)\}) = \mathbb{P}(\{(H, H, H)\}) + \mathbb{P}(\{(T, T, T)\}) = 0.216 + 0.064 = 0.28$$

Question 13. [1 MARK]

Let X be a random variable taking values in $\mathbb{R}$, and let $Y = 2X + 3$. Which of the following statements are true?

A. Y is a random variable.

B. The outcome space of Y is $\mathbb{R}$.

C. Y is not a random variable.

D. The outcome space of Y is $[3, \infty)$.

Solution 13.

Correct Answers: A., B.
Explanation:

A. **True.** Y is a function of the random variable X, so Y is also a random variable.

B. **True.** Since X can take any real value and $Y = 2X + 3$, the outcome space of Y is also $\mathbb{R}$.

C. **False.** Y is indeed a random variable.

D. **False.** The outcome space of Y is not restricted to $[3, \infty)$; since X can be any real number, Y can be any real number.

Question 14. [1 MARK]

Suppose X_1, X_2, X_3 are independent random variables, each with $X_i \sim \mathcal{N}(5, 4)$. Let $\bar{X} = \frac{1}{3}(X_1 + X_2 + X_3)$. Which of the following statements are true?

A. The expected value $\mathbb{E}[\bar{X}] = 5$.

B. The variance $\mathrm{Var}(\bar{X}) = \frac{4}{3}$.

C. The variance $\mathrm{Var}(X_1) = 4$.

D. The expected value $\mathbb{E}[\bar{X}] = \frac{5}{3}$.

Solution 14.

Correct Answers: A., B., C.
Explanation:

A. **True.**

$$\mathbb{E}[\bar{X}] = \mathbb{E}\left(\frac{1}{3}(X_1 + X_2 + X_3)\right) = \frac{1}{3}\left(\mathbb{E}[X_1] + \mathbb{E}[X_2] + \mathbb{E}[X_3]\right) = \frac{1}{3}(5 + 5 + 5) = 5$$

B. **True.**

$$\mathrm{Var}(\bar{X}) = \mathrm{Var}\left(\frac{1}{3}(X_1 + X_2 + X_3)\right) = \left(\frac{1}{3}\right)^2 \left(\mathrm{Var}(X_1) + \mathrm{Var}(X_2) + \mathrm{Var}(X_3)\right) = \frac{1}{9}(4{+}4{+}4) = \frac{12}{9} = \frac{4}{3}$$

C. **True.** Each X_i has variance $\sigma^2 = 4$.

D. **False.** $\mathbb{E}[\bar{X}] = 5$, not $\frac{5}{3}$.

Question 15. [1 MARK]

Let X be a random variable with outcome space $\mathcal{X} = \{a, b, c\}$ and pmf $p(a) = 0.2, \quad p(b) = 0.3, \quad p(c) = 0.5$. The function $f(x)$ is given by:

$$f(a) = 4, \quad f(b) = 6, \quad f(c) = 2$$

What is $\mathbb{E}[f(X)]$?

A. 3.0

B. 3.6

C. 4.0

D. 2.4

Solution 15.

Correct Answer: B. 3.6
Explanation:
Compute $\mathbb{E}[f(X)]$ using the definition of expected value:

$$\mathbb{E}[f(X)] = f(a) \cdot p(a) + f(b) \cdot p(b) + f(c) \cdot p(c) = (4)(0.2) + (6)(0.3) + (2)(0.5) = 0.8 + 1.8 + 1.0 = 3.6$$

Question 16. [1 MARK]

Let X represent the outcome of a biased coin flip, where $\mathbb{P}(X = 1) = 0.8$ (heads) and $\mathbb{P}(X = 0) = 0.2$ (tails). Given that $X = 1$, the conditional distribution of a random variable $N \in \{1, 2, 3\}$ is:

$$\mathbb{P}_{N|X}(N = 1|X = 1) = \frac{1}{5}, \quad \mathbb{P}_{N|X}(N = 2|X = 1) = \frac{2}{5}, \quad \mathbb{P}_{N|X}(N = 3|X = 1) = \frac{2}{5}.$$

What is the conditional expectation $\mathbb{E}[N \mid X = 1]$?

A. 2.0

B. 2.2

C. 1.8

D. 2.4

Solution 16.

Correct Answer: B. 2.2
Explanation:
Calculate $\mathbb{E}[N \mid X = 1]$:

$$\begin{aligned}
\mathbb{E}[N \mid X = 1] &= \sum_n n \cdot p_{N|X}(n \mid 1) \\
&= (1)\left(\frac{1}{5}\right) + (2)\left(\frac{2}{5}\right) + (3)\left(\frac{2}{5}\right) \\
&= \frac{1}{5} + \frac{4}{5} + \frac{6}{5} = \frac{11}{5} = 2.2
\end{aligned}$$

Question 17. [1 MARK]

Suppose you are tasked with predicting the price of a car based on the following information: the engine size in liters, the number of doors, and the year it was manufactured. Which of the following options correctly specifies the features and the label for this problem?

A. Features: engine size, number of doors; Label: year manufactured

B. Features: year manufactured, price of the car; Label: engine size

C. Features: engine size, number of doors, year manufactured; Label: price of the car

D. Features: price of the car, engine size; Label: number of doors

Solution 17.

Correct Answer: C.
Explanation:

A. **A.** Incorrect. The label is year manufactured, but we are supposed to predict the price.

B. **B.** Incorrect. The features include the price, but we are supposed to predict the price.

C. **C.** Correct. We are using engine size, number of doors, and year manufactured to predict the price of the car.

D. **D.** Incorrect. The label is number of doors, but we are supposed to predict the price.

Question 18. [1 MARK]

Suppose you have a dataset where each instance represents a smartphone. The features are: screen size (in inches), battery capacity (in mAh), and RAM size (in GB). The label is the brand of the smartphone (e.g., Apple iPhone, Samsung Galaxy, Google Pixel). Which of the following statements are true?

A. The features can be represented as an element of $\mathbb{R}^3$.

B. This is a regression problem.

C. The label set is a finite set.

D. This is a classification problem.

Solution 18.

Correct Answers: A., C., D.
Explanation:

A. **True.** The features (screen size, battery capacity, RAM size) are numerical values and can be represented as a vector in $\mathbb{R}^3$.

B. **False.** Regression involves predicting continuous numerical values, but the label (brand) is categorical.

C. **True.** The label set consists of a finite number of brands (e.g., Apple iPhone, Samsung Galaxy, Google Pixel).

D. **True.** Predicting the brand of the smartphone is a classification problem.

Question 19. [1 MARK]

Suppose you are working with a dataset where the features $X \in \mathbb{R}$ and the labels $Y \in \{0, 1\}$. You are using a predictor $f(X)$ which outputs predicted labels. The loss function is defined as:

$$\ell(f(X), Y) = \begin{cases} 1 & \text{if } f(X) \neq Y \\ 0 & \text{if } f(X) = Y \end{cases}$$

Which of the following expressions correctly represent the expected loss?

A. $\mathbb{E}[\ell(f(X), Y)] = \int_{\mathbb{R}} \sum_{y \in \{0,1\}} \ell(f(x), y)\, p(x, y)\, dx$

B. $\mathbb{E}[\ell(f(X), Y)] = \int_{\mathbb{R}} \ell(f(x), y)\, p(x)\, dx$

C. $\mathbb{E}[\ell(f(X), Y)] = \int_{\mathbb{R}} \sum_{y \in \{0,1\}} \ell(f(x), y)\, p(y \mid x)\, p(x)\, dx$

D. $\mathbb{E}[\ell(f(X), Y)] = \int_{\mathbb{R}} \ell(f(x), y)\, dx$

Solution 19.

Correct Answers: A., C.
Explanation:
The expected loss $\mathbb{E}[\ell(f(X), Y)]$ is calculated by averaging the loss over all possible values of X and Y, weighted by their probabilities.

A. **Correct.** This expression sums over y and integrates over x, using the joint probability $p(x, y)$.

B. **Incorrect.** This expression does not properly account for y, as y is not integrated or summed over.

C. **Correct.** This expression sums over y and integrates over x, using the conditional probability $p(y \mid x)$ and $p(x)$.

D. **Incorrect.** This expression lacks the probability terms and does not correctly represent the expected loss.

Question 20. [1 MARK]

Suppose you roll a six-sided die 4 times and observe the following outcomes: $X_1 = 2$, $X_2 = 5$, $X_3 = 3$, $X_4 = 6$.
What is the sample mean of these 4 rolls?

A. 3.5

B. 5.0

C. 4.0

D. 4.5

Solution 20.

Correct Answer: C. 4.0
Explanation:
Calculate the sample mean:

$$\text{Sample Mean} = \frac{X_1 + X_2 + X_3 + X_4}{4} = \frac{2+5+3+6}{4} = \frac{16}{4} = 4.0$$

Question 21. [1 MARK]

Suppose you are given a predictor $f(x) = 5 + 0.5x$, which models the relationship between the number of years of experience x and the salary y (in tens of thousands of dollars) of an employee. You are provided with the following dataset of 4 (x, y) pairs:

$$\mathcal{D} = ((x_1, y_1), (x_2, y_2), (x_3, y_3), (x_4, y_4)) = ((1, 5.5),\ (2, 6.0),\ (3, 6.0),\ (4, 7.5)).$$

Let the loss function be the squared loss $\ell(f(x), y) = (f(x) - y)^2$.
Calculate $\hat{L}(f) = \frac{1}{4}\sum_{i=1}^{4} \ell(f(x_i), y_i)$ (which is an estimate of the expected squared loss $L(f) = \mathbb{E}[\ell(f(x), y)]$).

A. 0.25

B. 0.5

C. 0.0625

D. 0.125

Solution 21.

Correct Answer: D. 0.125
Explanation:
Compute $f(x_i)$ and $\ell(f(x_i), y_i)$ for each i:

- For $x_1 = 1$:

$$f(1) = 5 + 0.5 \times 1 = 5.5$$
$$\ell(f(1), y_1) = (5.5 - 5.5)^2 = 0$$

- For $x_2 = 2$:

$$f(2) = 5 + 0.5 \times 2 = 6.0$$
$$\ell(f(2), y_2) = (6.0 - 6.0)^2 = 0$$

- For $x_3 = 3$:

$$f(3) = 5 + 0.5 \times 3 = 6.5$$
$$\ell(f(3), y_3) = (6.5 - 6.0)^2 = (0.5)^2 = 0.25$$

- For $x_4 = 4$:

$$f(4) = 5 + 0.5 \times 4 = 7.0$$
$$\ell(f(4), y_4) = (7.0 - 7.5)^2 = (-0.5)^2 = 0.25$$

Sum the losses:

$$\sum_{i=1}^{4} \ell(f(x_i), y_i) = 0 + 0 + 0.25 + 0.25 = 0.5$$

Compute $\hat{L}(f)$:

$$\hat{L}(f) = \frac{1}{4} \times 0.5 = 0.125$$

Therefore, the estimated loss is 0.125, which corresponds to option D.

Question 22. [1 MARK]

Consider the function $g(w) = e^{3w} + 2w^2$, where $w \in \mathbb{R}$. What is the second derivative $g''(w)$?

A. $g''(w) = 3e^{3w} + 4w$

B. $g''(w) = 9e^{3w} + 4$

C. $g''(w) = e^{3w} + 4$

D. $g''(w) = 9e^{3w} + 2$

Solution 22.

Correct Answer: B.
Explanation:
First, compute the first derivative $g'(w)$:

$$g'(w) = \frac{d}{dw}\left(e^{3w} + 2w^2\right) = 3e^{3w} + 4w$$

Then, compute the second derivative $g''(w)$:

$$g''(w) = \frac{d}{dw}\left(3e^{3w} + 4w\right) = 9e^{3w} + 4$$

Question 23. [1 MARK]

Consider the convex function $g(w) = 3w^2 - 12w + 7$, where $w \in \mathbb{R}$.
What is the minimum value of $g(w)$ and at what w is it achieved?

A. The minimum value is 1 at $w = 2$.

B. The minimum value is -5 at $w = 2$.

C. The minimum value is -5 at $w = -2$.

D. The minimum value is 7 at $w = 0$.

Solution 23.

Correct Answer: B.
Explanation:
To find the minimum value of the quadratic function $g(w) = 3w^2 - 12w + 7$, we find the vertex of the parabola.
The vertex occurs at:

$$w = -\frac{b}{2a} = -\frac{-12}{2 \times 3} = \frac{12}{6} = 2$$

Compute the minimum value:

$$g(2) = 3(2)^2 - 12(2) + 7 = 12 - 24 + 7 = -5$$

Therefore, the minimum value is -5 at $w = 2$.

Question 24. [1 MARK]

Consider the convex function $g(w) = \sum_{i=1}^{4}(w - x_i)^2$, where $x_1 = 1$, $x_2 = 3$, $x_3 = 5$, $x_4 = 7$, and $w \in \mathbb{R}$. What value of w minimizes $g(w)$?

A. $w = 5$

B. $w = 3$

C. $w = 4$

D. $w = 0$

Solution 24.

Correct Answer: C. $w = 4$
Explanation:
To find the value of w that minimizes $g(w) = \sum_{i=1}^{4}(w - x_i)^2$, we take the derivative of $g(w)$ with respect to w, set it to zero, and solve for w.
Step 1: Compute the derivative $g'(w)$

$$g'(w) = \frac{d}{dw}\left(\sum_{i=1}^{4}(w - x_i)^2\right) = \sum_{i=1}^{4}\frac{d}{dw}\left((w - x_i)^2\right) = \sum_{i=1}^{4}2(w - x_i)$$

$$g'(w) = 2\sum_{i=1}^{4}(w - x_i)$$

Step 2: Set the derivative equal to zero and solve for w

$$g'(w) = 0 \implies 2\sum_{i=1}^{4}(w - x_i) = 0$$

Divide both sides by 2:

$$\sum_{i=1}^{4}(w - x_i) = 0$$

Simplify the summation:

$$\sum_{i=1}^{4}(w - x_i) = 4w - \sum_{i=1}^{4} x_i = 0$$

Compute $\sum_{i=1}^{4} x_i$:

$$\sum_{i=1}^{4} x_i = x_1 + x_2 + x_3 + x_4 = 1 + 3 + 5 + 7 = 16$$

Set up the equation:

$$4w - 16 = 0$$

Solve for w:

$$4w = 16 \implies w = \frac{16}{4} = 4$$

Question 25. [1 MARK]

Consider the convex function $g(w_1, w_2) = w_1^2 + 4w_2^2 - 6w_1 - 16w_2 + 50$, where $w_1, w_2 \in \mathbb{R}$.
At what values of w_1 and w_2 is $g(w_1, w_2)$ minimized?

A. $w_1 = 2, \quad w_2 = 3$

B. $w_1 = 3, \quad w_2 = 2$

C. $w_1 = 2, \quad w_2 = 2$

D. $w_1 = 3, \quad w_2 = 3$

Solution 25.

Correct Answer: B. $w_1 = 3, \quad w_2 = 2$
Explanation:
To find the values of w_1 and w_2 that minimize $g(w_1, w_2)$, we take the first partial derivatives with respect to w_1 and w_2, set them equal to zero, and solve for w_1 and w_2.

Step 1: Compute the partial derivatives

$$\frac{\partial g}{\partial w_1} = 2w_1 - 6$$

$$\frac{\partial g}{\partial w_2} = 8w_2 - 16$$

Step 2: Set the partial derivatives equal to zero
Set $\frac{\partial g}{\partial w_1} = 0$:

$$2w_1 - 6 = 0 \implies w_1 = \frac{6}{2} = 3$$

Set $\frac{\partial g}{\partial w_2} = 0$:

$$8w_2 - 16 = 0 \implies w_2 = \frac{16}{8} = 2$$