UNIVERSITY OF CALIFORNIA, LOS ANGELES
Cyrus Asasi
506014946

Computer Science M146 **Prof. Suhas Diggavi**

Homework 0

Due Sunday, April 7, 2024 11:59pm **via Gradescope**

1. Consider $y = xsin(z)e^{-x}$. What is the partial derivative of y with respect to x?
   The partial derivative of y with respect to x is

$$\frac{\partial y}{\partial x} = \frac{\partial}{\partial x}\left[xsin(z)e^{-x}\right] = sin(z)(e^{-x} - xe^{-x})$$

2. Consider the matrix $\mathbf{X}$ and the vectors $\mathbf{y}$ and $\mathbf{z}$ below:

$$\mathbf{X} = \begin{pmatrix} 2 & 4 \\ 1 & 3 \end{pmatrix} \qquad \mathbf{y} = \begin{pmatrix} 1 \\ 3 \end{pmatrix} \qquad \mathbf{z} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

   (a) What is the inner product $\mathbf{y}^T\mathbf{z}$?

$$\mathbf{y}^T\mathbf{z} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}^T \cdot \begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 & 3 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 3 \end{pmatrix} 1 \cdot 2 + 3 \cdot 3 = 11$$

   (b) What is the product $\mathbf{Xy}$?

$$\mathbf{Xy} = \begin{pmatrix} 2 & 4 \\ 1 & 3 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 2 \cdot 1 + 4 \cdot 3 \\ 1 \cdot 1 + 3 \cdot 3 \end{pmatrix} = \begin{pmatrix} 14 \\ 10 \end{pmatrix}$$

   (c) Is $\mathbf{X}$ invertible? If so, give the inverse; if not, explain why not.
   In order for a matrix to be invertible, it must satisfy the following two conditions:
      i. It must be a square matrix.
      ii. The determinant must be non-zero.

$$\det \begin{pmatrix} 2 & 4 \\ 1 & 3 \end{pmatrix} = 6 - 4 = 2 \neq 0 \ \checkmark$$

   Since both conditions are satisfied, the matrix $\mathbf{X}$ is invertible.

$$\begin{pmatrix} 2 & 4 \\ 1 & 3 \end{pmatrix}^{-1} = \frac{1}{\det(\mathbf{X})} \cdot \begin{pmatrix} 3 & -4 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 1.5 & -2 \\ -0.5 & 1 \end{pmatrix}$$

   (d) What is the rank of $\mathbf{X}$
   Since the second row cannot be derived from the first row, the rank of $\mathbf{X}$ is 2. Checking both columns, we also see that the second column cannot be derived from the first row, which means the rank of $\mathbf{X}$ is 2.

3. Consider a sample of data S obtained by flipping a coin five times. $X_i, i \in \{1, \ldots, 5\}$ is a random variable that takes a value 0 when the outcome of coin flip i turned up heads, and 1 when it turned up tails. Assume that the outcome of each of the flips does not depend on the outcomes of any of the other flips. The sample obtained $S = (X_1, X_2, X_3, X_4, X_5) = (1, 1, 0, 1, 0)$.
   (a) What is the sample mean for this data?

$$\text{sample mean} = \frac{1+1+0+1+0}{5} = \frac{3}{5} = 0.6$$

(b) What is the unbiased sample variance?

$$
\begin{aligned}
\text{sample variance} &= \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \text{sample mean})^2 \\
&= \frac{1}{4}\left[(1-0.6)^2 + (1-0.6)^2 + (0-0.6)^2 + (1-0.6)^2 + (0-0.6)^2\right] \\
&= \frac{1}{4}\left[0.4^2 + 0.4^2 + 0.6^2 + 0.4^2 + 0.6^2\right] \\
&= \frac{1}{4}\left[0.16 + 0.16 + 0.36 + 0.16 + 0.36\right] \\
&= \frac{1}{4}\left[1.2\right] = 0.3
\end{aligned}
$$

(c) What is the probability of observing this data assuming that a coin with an equal probability of heads and tails was used? (i.e., The probability distribution of X_i is $P(X_i = 1) = 0.5$, $P(X_i = 0) = 0.5$.)

$$P(11010) = P(1) \cdot P(1) \cdot P(0) \cdot P(1) \cdot P(0) = 0.5^5 = 0.03125$$

(d) Note the probability of this data sample would be greater if the value of the probability of heads $P(X_i = 1)$ was not 0.5 but some other value. What is the value that maximizes the probability of the sample S ? [Optional: Can you prove your answer is correct?]

Taking $P(X_i = 1)$ to be the sample mean of the given outcomes will maximize the probability of the sample S. This is because in this scenario, the best we can do is ensure that the ratio of getting a 1 to getting a 0 is 3:2.

$$P(11010) = P(1) \cdot P(1) \cdot P(0) \cdot P(1) \cdot P(0) = 0.6^3 + 0.4^2 = 0.376$$

(e) Given the following joint distribution between X and Y, what is $P(X = T \mid Y = b)$?

$P(X,Y)$		Y a	b	c
X	T	0.2	0.1	0.2
	F	0.05	0.15	0.3

Bayes rule gives us $P(A \mid B) = \frac{P(A \cap B)}{P(B)}$ Therefore

$$P(X = T \mid Y = b) = \frac{P(X = T \cap Y = b)}{P(Y = b)} = \frac{0.1}{0.1 + 0.15} = 0.4$$

4. Match the distribution name to its formula.

(a) Gaussian (i) $p^x(1-p)^{1-x}$, when $x \in \{0,1\}; 0$ otherwise

(b) Exponential (ii) $\frac{1}{b-a}$ when $a \leq x \leq b; 0$ otherwise

(c) Uniform (iii) $\binom{n}{x} p^x (1-p)^{n-x}$

(d) Bernoulli (iv) $\lambda e^{-\lambda x}$ when $x \geq 0; 0$ otherwise

(e) Binomial (v) $\frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right)$

(a) -> (v)
(b) -> (iv)
(c) -> (ii)
(d) -> (i)
(e) -> (iii)

5. (a) What is the mean and variance of a *Bernoulli(p)* random variable?

   The expectation of a bernoulli random variable $\mathbb{E}[X] = p$.
   The variance of a bernoulli random variable is $\text{VAR}(X) = p(1-p)$.

   (b) If the variance of a zero-mean random variable X is σ^2, what is the variance of $2X$? What about the variance of $X + 2$?

   Variance is defined as $\mathbb{E}[X^2] - \mathbb{E}[X]^2$. Since the mean is zero, we have $\mathbb{E}[X^2] = \sigma^2$.

$$\begin{aligned}\mathbb{E}[(2X)^2] - \mathbb{E}[2X]^2 &= \mathbb{E}[4x^2] - 2\mathbb{E}[X]^2 \\ &= 4\mathbb{E}[X^2] - 0 \\ &= 4\sigma^2\end{aligned}$$

$$\begin{aligned}\mathbb{E}[(X+2)^2] - \mathbb{E}[X+2]^2 &= \mathbb{E}[X^2 + 2X + 4] - (\mathbb{E}[X] + \mathbb{E}[2])^2 \\ &= E[X^2] + \mathbb{E}[2X] + \mathbb{E}[4] - 4 = \sigma^2 + 4\sigma^2 + 4 - 4 \\ &= 5\sigma^2\end{aligned}$$

6. For each pair (f, g) of functions below, list which of the following are true: $f(n) = O(g(n))$, $g(n) = O(f(n))$, or both. Briefly justify your answers.

   (a) $f(n) = \ln(n), g(n) = \lg(n)$. Note that ln denotes log to the base e and lg denotes log to the base 2.

   Usign L'Hopital's rule:

$$\begin{aligned}\lim_{n\to\infty}\frac{ln(n)}{lg(n)} &= \lim_{n\to\infty}\frac{\frac{1}{n}}{\frac{1}{nln(2)}} = ln(2) \neq 0 \\ \lim_{n\to\infty}\frac{lg(n)}{ln(n)} &= \lim_{n\to\infty}\frac{\frac{1}{nln(2)}}{\frac{1}{n}} = \frac{1}{ln(2)} \neq 0\end{aligned}$$

   Thus, $f(n) \neq O(g(n))$ and $g(n) = O(f(n))$

   (b) $f(n) = 3^n, g(n) = n^{10}$

   Applying L'Hopital's rule 10 times:

$$\begin{aligned}\lim_{n\to\infty}\frac{n^{10}}{3^n} &= \lim_{n\to\infty}\frac{10!}{3^n \cdot ln(3)^{10}} = 0 \\ \lim_{n\to\infty}\frac{3^n}{n^{10}} &= \lim_{n\to\infty}\frac{3^n \cdot ln(3)^{10}}{10!} = \infty\end{aligned}$$

   Thus, $f(n) \neq O(g(n))$, but $g(n) = O(f(n))$

   (c) $f(n) = 3^n, g(n) = 2^n$

$$\begin{aligned}\lim_{n\to\infty}\frac{3^n}{2^n} &= \infty \\ \lim_{n\to\infty}\frac{2^n}{3^n} &= 0\end{aligned}$$

   Thus, $f(n) \neq O(g(n))$, but $g(n) = O(f(n))$

7. If X and Y are independent random variables, show that $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$.