Instructor: **Nikhil Muralidhar**
October 26, 2022

Due Date: November 11, 2022

CS 556-B: Mathematical Foundations of Machine Learning
Homework 3: Probability & Calculus (100 points)

Note: All solutions methods must be fully explained. All solutions must be compiled either using Microsoft Word
or LaTeX and submitted as a PDF file. Hand-written assignment solutions will NOT be accepted.

Probability (60 points)

1. (20 points) Jane picks apples. The weight of the apples picked by Jane are normally distributed with a mean of 160g and a standard deviation of 50g. Jane does not like apples that weigh less than or equal to 95g.

   (a) (10 points) What is the percentage of apples that Jane picks and does not like? **Hint:** Employ Z-scores and the z-table to calculate the required percentage.

   (b) (10 points) Let q_3 represent the minimum weight for the upper quartile (i.e., top 25%) of weight of apples liked by Jane. What is the probability that the weight of a randomly selected apple picked by Jane is greater than or equal to q_3. **Hint:** Consider the probability that Jane *likes* an apple AND that the weight of the apple lies in the *upper quartile* (i.e., weight $\geq q_3$)

2. (10 points) A bag contains five types of balls colored red, blue, green, yellow, orange. There are 10 each of red, blue, green and yellow balls and there are 3 orange balls. We draw a ball from the bag, if the ball drawn is colored green or yellow we get $0 payoff. If it is blue the we get a $200 payoff. If the ball drawn is red, we get a $300 payoff and $500 if it is orange.

   (a) (5 points) What is our expected payoff of this experiment? **Hint:** Calculate expected value.

   (b) (5 points) What is the standard deviation and variance? **Hint:** $\sigma^2(X) = E[X^2] - (E[X])^2$

3. (10 points) Harry is 24 years old and plans to get his driver's license. The written knowledge test Harry has to give before the driving test is multiple choice where every question has y options. There are no negative marks for wrong answers and hence Harry chooses to guess an answer if he doesn't know the correct answer to the question. Let x be the probability that he knows the answer. The probability where Harry guesses at the answer and it will be correct is $\frac{1}{y}$. What is the probability that Harry knows the correct answer to a problem given that he has answered that problem correctly. (Assume that Harry does not leave any questions on the test unanswered. Also assume that for any question to which harry knows the answer, his answers will always be correct.) **Hint:** Employ Bayes rule.

4. (10 points) It is estimated that 50% of emails are spam emails. Some software has been applied to filter these spam emails before they reach your inbox. A certain brand of software claims that it can detect 99% of spam emails, and the probability for a false positive (a non-spam email detected as spam) is 5%. Now if an email is detected as spam, then what is the probability that it is in fact a non-spam email?

5. (10 points) In a study, doctors were asked what the odds of breast cancer would be in a woman, who was initially thought to have a 1% risk of cancer but who ended up with a positive mammogram result . A mammogram accurately classifies about 80% of cancerous tumors and 90% of benign tumors. What are the chances this patient has cancer?

Differential Calculus (40 points)

6. (10 points) **Limit Definition** of a function $F(x)$, is given as follows.

$$F(x) = \lim_{h \to 0} \frac{F(x+h) - F(x)}{h}$$

(a) (5 points) Given a function $F(x) = \frac{f(x)}{g(x)}$ where $f(x)$ and $g(x)$ are also functions of x, the quotient rule of derivative of $F(x)$ (denoted $F'(x)$) is given as follows.

$$F'(x) = \frac{g(x)f'(x) - f(x)g'(x)}{[g(x)]^2}$$

Use the limit definition of a function to derive the quotient rule of the derivative of function $F(x)$.

(b) (5 points) Consider a function $f(x) = \sin(x)$. Show using the limit definition that $f'(x) = \cos(x)$.

7. (5 points) Calculate the equation of a tangent line given by a function $f(x) = \frac{x^2}{e^x}$ and report the result at $x = 1$. Note: You are free to use any derivative rules you please and you DO NOT need to substitute in the value of 'e'. It will suffice to express any potential fractions in their simplest form which include the constant 'e' as is. **Hint:** Calculate slope of function and employ point slope form of a line to get the tangent equation.

8. (5 points) Consider a function $f(x) = e^x$ and a function $g(x) = \frac{x^2}{x-1}$. Let $F(x) = f(g(x))$, find an expression for $F'(x)$. **Hint:** Employ chain rule. Answer does not need to be simplified after FULL application of derivatives (the point is to test your application of derivatives not your algebra). Please enumerate the derivative rules you have employed.

9. (12 points) Given $F(x)$ in the following problems, find the derivative $F'(x)$. **Note:** Answer does not need to be simplified after FULL application of derivatives (the point is to test your application of derivatives not your algebra). Please enumerate the derivative rules you have employed per question.

(a) (2 points) $F(x) = \frac{\sqrt{x} + 2x}{7x - 4x^2}$

(b) (2 points) $F(x) = (1 + \sqrt{x^3})(\frac{1}{x^3} - 2\sqrt[3]{x})$

(c) (2 points) $F(x) = (2x + 1)^5(3x - 2)^7$

(d) (3 points) $F(x) = \frac{x\sqrt{2x+1}}{e^x \sin^3(x)}$ Hint: Employ logarithmic differentiation technique.

(e) (3 points) $F(x) = \sqrt{\frac{(x-1)(x-2)}{(x-3)(x-4)(x-5)}}$ Hint: Employ logarithmic differentiation technique.

10. (4 points) Calculate the partial derivatives with respect to each variable of the function $F(\cdot)$ in each of the following cases.

(a) (2 points) $F(x, y) = \frac{x^3 y^2 - 2x + 5}{e^x}$

(b) (2 points) $F(x, y, z) = y^2 \ln(x + 2y) - \ln(3z)(x^3 + y^2 - 4z)$

11. (4 points) The table below depicts rainfall depth (mm) in New York City beginning 7 PM on Sept. 19th 2008. **Hint:** In each of the following questions, employ numerical differentiation techniques.

Hours after 7 PM	0	5	10	15	20	25	30
Rainfal Depth (mm)	250	290	330	240	300	350	400

Table 1: Rainfall Depth in New York City Sept 19th 2008, at and after 7 PM.

(a) (2 point) Use central difference to approximate the slope of rainfall depth using numerical differentiation at hour 12.5 and at hour 22.5.

(b) (1 point) Calculate the slope of rainfall depth at hour 15 using forward numerical differentiation.

(c) (1 point) Calculate the slope of rainfall depth at hour 15 using backward numerical differentiation.