Homework 1

Math for AI

Instructions: Please solve the following problems showing all your work. State any assumptions you make.

1. **Data Analysis**
   You are analyzing the performance of a new classification model. Let X be the random variable representing the model's confidence score (discretized into Low, Medium, High), and let Y be the variable representing whether the prediction was Correct ($Y = 1$) or Incorrect ($Y = 0$). The joint probability distribution $P(X, Y)$ is given below:

	Confidence (X)		
Prediction (Y)	Low (x_1)	Medium (x_2)	High (x_3)
Incorrect ($Y = 0$)	0.15	0.10	0.05
Correct ($Y = 1$)	0.05	0.15	0.50

   (a) Calculate the marginal distribution of the confidence scores, $P(X)$.

   (b) What is the overall accuracy of the model, $P(Y = 1)$?

   (c) Calculate the conditional probability that a prediction is correct given that the confidence is High, $P(Y = 1|X = \text{High})$.

   (d) Are the random variables X and Y independent? Justify your answer mathematically.

2. **System Reliability**
   A cloud computing request must pass through three independent layers of security checks to be processed: a Firewall (F), an Authentication Service (A), and an Authorization Policy (Z).

   - The probability the Firewall passes the request is 0.98
   - The probability the Authentication Service passes the request given that the Firewall has passed is 0.95.
   - The probability the Authorization Policy passes the request given that both the Firewall and the Authentication Service have passed is 0.99.

   (a) What is the probability that a legitimate request is successfully processed (passes all three checks)?

   (b) What is the probability that a request is rejected specifically by the Authentication Service?

3. **Spam Detection**
Let S denote the event that an email is spam ($S = 1$) or not ($S = 0$). Based on prior data, the probability that an incoming email is spam is $P(S = 1) = 0.4$.

   You have a spam filter that looks for the keyword "Congratulations". Let $K = 1$ indicate the presence of the keyword.

   - If an email is spam, the probability it contains the keyword is $P(K = 1|S = 1) = 0.25$.
   - If an email is not spam, the probability it contains the keyword is $P(K = 1|S = 0) = 0.05$.

   If an incoming email contains the word "Congratulations", what is the posterior probability that it is actually spam?

4. **Attracting events**
Prove that for two events in a sample space, if $P(A|B) > P(A)$ then $P(B|A) > P(B)$.

5. **Engines**
Customers who purchase a certain make of car can order an engine in any of three sizes. Of all cars sold, 45% have the smallest engine, 35% have the medium sized one, and 20% have the largest. Of cars with the smallest engine, 10% fail an emissions test within the warranty period, while 12% of those with medium size and 15% of those with largest engine fail. A friend of yours tells you that her car, which is of this make, failed an emissions test within the warranty period. Compute the probabilities that your friend's car has small, medium or large engine.

6. **Server Pings**
A monitoring node sends "pings" to a remote server. Each ping has a probability $\mu = 0.8$ of receiving a successful echo reply, independent of other pings.

   (a) Let X be a random variable for a single ping ($X = 1$ for success, $X = 0$ for failure). What is the variance of X?
   (b) You send a batch of $N = 10$ pings. Let Y be the number of successful replies. What kind of distribution does Y follow?
   (c) Calculate the probability of receiving exactly 8 successful replies.
   (d) What is the expected number of successful replies in the batch?

7. **API Traffic**
Requests arrive at a microservice API at an average rate of $\lambda = 4$ requests per second. Assume the arrival of requests follows a Poisson process.

   (a) Write the expression for the probability of receiving k requests in a given second.
   (b) What is the probability that the service receives exactly 0 requests in a specific second?
   (c) What is the probability that the service receives 2 or more requests in a specific second?

8. **LLM Token generation**
A simplified Large Language Model has a vocabulary of only 3 tokens: "cat", "dog", "fish". For a specific prompt, the model outputs a categorical distribution with probabilities:

$$\mu_{\text{cat}} = 0.5, \quad \mu_{\text{dog}} = 0.3, \quad \mu_{\text{fish}} = 0.2$$

(a) If we sample $N = 50$ tokens independently, what is the expected number of times the token "dog" will appear? What is the variance for this count?

(b) We sample $N = 5$ tokens independently from this distribution. What is the probability of generating exactly: 3 "cats", 1 "dog", and 1 "fish"?

9. **Expectation and Variance of a Custom Variable**
A GPU cluster job has three possible states after execution:

- State A: Success (Probability 0.6). Reward $R = 10$.
- State B: Soft Failure (Probability 0.3). Reward $R = 1$ (partial results).
- State C: Hard Failure (Probability 0.1). Reward $R = -5$ (wasted compute time).

(a) Calculate the Expected Reward $E[R]$.

(b) Calculate the Variance of the Reward $Var(R)$.

10. **Maximum Likelihood Estimation**
Derive the formula for the MLE of the parameter λ of the Poisson distriburion.

11. **Quiz in two parts**
An instructor has given a short quiz consisting of two parts. For a randomly selected student, let X be the number of points earned on the first part and Y be the number of points earned on the second part. Suppose that the joint pmf of X and Y is given by the accompanying table:

		y			
	$p(x, y)$	0	5	10	15
	0	0.02	0.06	0.02	0.10
x	5	0.04	0.15	0.20	0.10
	10	0.01	0.15	0.14	0.01

a) Compute the marginal probability distributions of X and Y.

b) Compute the conditional probability mass function of Y given that $X = 10$.

c) Compute the conditional mean and variance of X given that $Y = 5$.

d) Are X and Y independent; justify your conclusion numerically.

12. **Voters**
In a certain town, 40% of the eligible voters prefer candidate A, 10% prefer candidate B, and the remaining 50% have no preference. You randomly sample 12 eligible voters.

a) What is the probability that 6 will prefer candidate A, 2 will prefer candidate B, and the remaining 4 will have no preference?

b) What is the marginal distribution for the number of voters who prefer candidate A in your sample?

c) What is the marginal joint pmf for the number of voters who prefer candidates A, B in your sample?

d) What is the conditional joint distribution for the number of voters who prefer candidates A, B in your sample given that 5 voters in the sample have no preference?