Homework 0: Basic Probability Review Problems **Name:**
ECE/CS 498 DS Spring 2020 **NetID:**
Issued: 01/27/2020
Due: 02/03/2020 (11:59 PM)
Please submit your homework on Compass 2G.

Please submit Problem 1, 4, 5, 7, 9, 10, 12 for grading. The rest are for your practice.

Problem 1 (Basic Concepts)

(a) **(5 points)** Write down the Probability Axioms.

(b) **(5 points)** Explain the differences between a probability mass function (pmf) at a point and a probability density function (pdf) at a point.

(c) **(5 points)** If A and B are independent events with $P(A) = 0.8$, and $P(B) = 0.5$, find $P(A \bigcup B)$.

(d) **(5 points)** Prove $P(A, B|C) = P(A|B, C) \times P(B|C)$.
Hint: Start from $P(A, B|C) = \frac{P(A,B,C)}{P(C)}$.

Problem 2 (Counting)

(a) Find the number of solutions of $x+y+z = 15$ where x, y, z are all positive integers.

(b) Find the number of solutions of $x+y+z < 15$ where x, y, z are all positive integers.

(c) Find the number of solutions of $x + y + z = 15$ where x, y, z are all nonnegative integers.

Problem 3 (Independence)

There are 9 identical balls in an urn. 2 balls are marked “none”, 2 balls are marked “1”, 2 balls are marked “2”, 2 balls are marked “3”, and 1 ball is marked “123”. “none” means no number is marked on that ball. Suppose a ball is taken from the urn at random, event $A_i = \{$“i” is on the ball$\}$. For example, A_1 occurs when ball “1” is picked or when ball “123” is picked. We can find $P(A_i) = \frac{1}{3}$.

(a) What is the difference between pairwise independence and mutual independence? Illustrate your answer with respect to three random variables $X, Y,$ and Z.

(b) Are A_1, A_2, A_3 pairwise independent? (Show your calculation.)

(c) Are A_1, A_2, A_3 mutually independent? (Show your calculation.)

Problem 4 (Exponential Distributions and Poisson Distributions)

(a) Exponential distribution is often used to model the lifetime of electronic components in autonomous vehicles. An exponential random variable X can be parameterized by its *rate* $\lambda\,(\lambda > 0)$ via the probability density function (pdf):

$$f(x) = \lambda e^{-\lambda x}, \qquad\qquad x > 0$$

(i) **(5 points)** Derive the cumulative distribution function (cdf) of the exponential distribution.

(ii) **(5 points)** Explain the memoryless property of the exponential distribution and provide the mathematical expression.

(iii) **(5 points)** Derive the mean and variance of the exponential distribution.

(b) **(10 points)** The Poisson distribution can be seen as a limiting case of the binomial distribution as the number of trials goes to infinity and the expected number of successes remains fixed. Derive the Poisson distribution from the Binomial distribution.

Problem 5 (Marginal/Joint Distributions)

Let X and Y be jointly continuous random variables with the following joint PDF

$$f_{X,Y}(x,y) = \begin{cases} 10e^{-(2x+5y)} & x \geq 0, y \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

(a) **(5 points)**Find the marginal distribution of the random variable X.

(b) **(5 points)**Find the marginal distribution of the random variable Y.

(c) **(5 points)**Are X and Y independent? Explain your answer.

(d) **(5 points)**What is the conditional PDF $f_{Y|X}(y|x)$. Include the values of x and y for which it is (i) well defined and (ii) zero .

(e) **(5 points)**Determine $P\{Y > X\}$.

Problem 6 (Inequalities)

(a) A coin is weighted so that the probability of landing on heads is 40%. Suppose the coin is flipped a 100 times. Find the upper bound on the probability the coin lands on heads at least 80 times.

(b) Using the same coin in (a), find the upper bound on the probability that the coin lands on heads at least 50 times or at most 30 times.

(c) Derive Chebyshev's Inequality from Markov's Inequality.

Problem 7 (Covariance and Correlation Coefficient)

(a) Suppose X and Y are random variables, $\mathrm{Var}(X+Y)=7$, $\mathrm{Var}(2X-2Y)=12$.

  (i) **(5 points)**Find the covariance $\mathrm{Cov}(X,Y)$.

  (ii) **(5 points)**In addition to (i), given Var(X)=1, find the correlation coefficient $\rho_{X,Y}$.

(b) **(10 points)** Suppose random variables $X_1, X_2, \cdots, X_{10}$ are uncorrelated. For each i in $\{1,2,\cdots,10\}$, $\mathrm{E}[X_i]=i$ and $\mathrm{Var}(X_i)$=5. Find $\mathrm{Var}(\frac{S_{10}}{\sqrt{10}})$, where $S_{10} = X_1 + X_2 + \cdots + X_{10}$.

Problem 8 (Continuous Random Variable)

Here we define a probability density function $f(x)$. $f(x) = \frac{x^3}{\alpha}$ when $0 < x < 6$, $f(x) = 0$ otherwise. $X_1, ..., X_{50}$ are independent, continuous random variables and each one has probability density function $f(x)$.

(a) Find the valid value of α.

(b) Find the expectation $\mathrm{E}[X_i]$.

(c) Find the variance $\mathrm{Var}(X_i)$.

(d) Find a good estimation for $P(X_1 + X_2 + \cdots + X_{50} < 230)$. (Hint: Central Limit Theorem)

Problem 9 (Central Limit Theorem)

An autonomous vehicle consists of 400 independent components. Assume the probability that each component functions properly is 0.98.

1. **(5 points)** Random variable X is the number of properly functioning components. Find the distribution of X.

2. **(5 points)** The vehicle requires at least 390 properly functioning components to work. Use the Central Limit Theorem to find the probability that the system works.

Problem 10 (Bayes Theorem and Conditional Probabilities)

(10 points) When autonomous vehicles have malfunctions, the probability of a disengagement is 0.85. When autonomous vehicles do not have malfunctions, the probability of a disengagement is 0.002. If the probability of a malfunction is 0.0002, evaluate the probability that a given disengagement is due to a malfunction.
FYI: A disengagement is a failure that causes the control of the vehicle to switch from the software to the human driver.

Problem 11 (Bayes Theorem and Conditional Probabilities)

Timely patching is important for server security. Suppose an organization has 3 servers. Two of them have been patched, while one of them still remains unpatched. The probability an unpatched server gets compromised is $\frac{1}{2}$, and the probability that a patched server get compromised is p. If an attacker randomly attacks one of the three servers, the probability he compromises the server is $\frac{2}{3}$.

(a) What is the value of p?

(b) This time the attacker randomly chose two serves to attack, and observed only one of them got compromised. What is the conditional probability that both attacked servers are patched?

Problem 12 (Uniform Distribution)

Suppose you are waiting for buses at a bus stop. In the next 10 minutes, both bus A and bus B are expected to arrive, and their arrival time are independent to each other. We use X to denote the arrival time of bus A, and Y to denote the arrival time of bus B. X and Y are continuous variables, and each follows the uniform distribution over [0,10].

(a) **(5 points)** Find the probability that bus A and bus B arrive at exactly the same time.

(b) **(5 points)** Find the probability that bus A arrives earlier than bus B.

(c) **(5 points)** Denote Z as the arrival time of the later of the two. Find the pdf of Z and the expectation E[Z]. (Use minute as the unit, and leave the answer in decimal or fraction)

(d) **(5 points)** Denote W as the arrival time of the earlier of the two. Find the expectation E[W].

(e) **(5 points)** Suppose both bus A and bus B will wait for 5 minutes at the bus stop. Find the probability that bus A and bus B are together at the bus stop.