MTH 324, Homework #2, Fall 2022 Name ______________________________________

Instructions: Show work or attach R code used to perform calculations (or any other technology used). Be sure to answer all parts of each problem as completely as possible, and attach work to this cover sheet with a staple.

1. One percent of all individuals in a certain population are carriers of a particular disease. A diagnostic test for this disease has a 90% detection rate for carriers and a 5% detection rate for non-carriers. Suppose the test is applied independently to two different blood samples from the same randomly selected individual. [Hint: It may help to draw a tree diagram of the two tests.]
   a. What is the probability that both tests yield the same result?
   b. If both tests are positive, what is the probability that the selected individual is a carrier?

2. There has been a great deal of controversy over the last several years regarding what types of surveillance are appropriate to prevent terrorism. Suppose a particular surveillance system has a 99% chance of correctly identifying a future terrorist and a 99.9% chance of correctly identifying someone who is not a future terrorist. If there are 1000 future terrorists in a population of 300 million, and one of these 300 million is randomly selected, scrutinized by the system and identified as a terrorist, what is the probability that he or she is actually a future terrorist? Does the value of this probability make you uneasy about using the surveillance system? Explain.

3. Consider independently rolling two fair dice, one red and one green. Let A be the event that the red die shows 3, and let B be the event that the green die shows 4, and C be the event that the sum of the two dice is 7. Are these events pairwise independent? Are the three events mutually independent?

4. Use the two-way table to answer the questions that follow. The table shows the people attending an afternoon play.

	Stalls	Circle	Balcony	Total
Adults	36	39		112
Children	41		31	
Total		60		

   a. Complete the table.
   b. What is the probability that a randomly chosen audience member is an adult?
   c. What is the probability that a randomly chosen audience member is seated in the balcony?
   d. What is the probability that a randomly chosen audience member is an adult and is seated on the balcony?
   e. What is the probability a randomly selected person is an adult, given that they are sitting in the balcony?
   f. What is the probability a randomly selected person is seated in the balcony, given that they are an adult?
   g. What is the probability that a randomly selected audience member is either an adult or seated in the balcony?

h. Are being an adult and being seated in the balcony independent events? Why or why not? Show mathematical calculations to justify your answer.

5. If the sample space S is an infinite set, does this necessarily imply than any random variable X defined from S will have an infinite set of possible values? If yes, say why. If no, give an example.

6. For each random variable defined here, describe the set of possible values for the variable, and state whether the variable is discrete.
   a. X is the number of unbroken eggs in a randomly chosen standard egg carton.
   b. Y is the number of students on a class list for a particular course who are absent on the first day of class.
   c. U is the number of times a golfer has to swing at a golf ball before hitting it.
   d. V is the length of a randomly selected rattlesnake.
   e. W is the amount of royalties earned from the same of a first edition of $10,000 textbooks.
   f. Z is the pH of a randomly chosen soil sample.
   g. T is the tension (psi) at which a randomly selected tennis racket has been strung.
   h. S is the total number of coins tosses required for three individuals to obtain a match.

7. A mail-order telephone business has 6 telephone lines. Let X denote the number of lines in use at a specified time. Suppose the probability mass function of X is given in the accompanying table.

x	0	1	2	3	4	5	6
$p(x)$	0.10	0.15	0.20	0.25	0.20	0.06	0.04

   a. Find the cumulative distribution function (table)
   b. Calculate the following probabilities.
      i. at most three lines are in use
      ii. fewer than three lines are in use
      iii. at least three lines are in use
      iv. between two and three lines, inclusive, are in use
      v. between two and four lines, inclusive, are not in use
      vi. at least 4 lines are not in use
   c. Find the expected value and the variance.

8. Benford's law is used by the IRS to detect fraud in financial reporting. The law says that the probability of the random variable X representing the leading digit of numbers reported are given by $p(x) = \log_{10}\left(\frac{x+1}{x}\right), x = 1,2,3,\ldots,9.$
   a. Without computing individual probabilities from this formula, show that it specifies a legitimate probability pass function. [Hint: calculate the sum of the probabilities by using log properties and summation properties.]
   b. Now compute the individual probabilities and compare to the corresponding discrete uniform distribution (if, in fact, all leading digits of numbers were genuinely random).
   c. Obtain the cumulative distribution function of X.
   d. What the result in part (c), what is the probability that the leading digit is at most 3? At least 5?

9. An appliance dealer sells three different models of upright freezers having 13.5, 15.9 and 19.1 cubic feet of storage space, respectively. Let X = the amount of storage space purchased by the next customer to a buy a freezer. Suppose that X has a probability mass function show.

x	**13.5**	**15.9**	**19.1**
$p(x)$	0.2	0.5	0.3

   a. Compute $E(X), E(X^2), V(X)$.
   b. If the price of a freezer having capacity X cubic freezer is $2.5X - 8.5$, what is the expected price paid by the next customer to buy a freezer?
   c. What is the variance of the price $25X - 8.5$ paid by the next customer?
   d. Suppose that although the rated capacity of a freezer is X, the actual capacity is $h(X) = X - 0.01X$. What is the expected actual capacity of the freezer purchased by the next customer?

10. A geologist has collected 10 specimens of basaltic rock and 10 specimens of granite. The geologist instructs a lab assistant to randomly select 15 of the specimens for analysis.
   a. What is the probability mass function of the number of granite specimens selected for analysis?
   b. What is the probability that all specimens of one of two types of rock are selected for analysis?
   c. What is the probability that the number of granite specimens selected for analysis is within one standard deviation of its mean value?

11. Suppose that $p = P(male\ birth) = 0.49$. A couple wishes to have exactly two female children in their family. They will have children until this condition is fulfilled.
   a. What is the probability that the family has x male children?
   b. What is the probability that the family has four children?
   c. What is the probability that the family has at most four children?
   d. How many male children would you expect this family to have? How many children would you expect this family to have?

12. Suppose that small aircraft arrive at a certain airport according to a Poisson process with rate $\alpha = 8$ per hour, so that the number of arrivals during a time period t hours is a Poisson random variable with parameter $\mu = 8t$.
   a. What is the probability that exactly six small aircraft arrive during a 1-hour period? At least 6? At least 10?
   b. What is the expected value and the standard deviation of the number of small aircraft that arrive during a 90-minute period?
   c. What is the probability that at least 20 small aircraft arrive during a 2.5-hour period? That at most 10 arrive during this period?

13. The current in a certain circuit as measured by an ammeter is a continuous random variable X with the following density function: $f(x) = \begin{cases} 0.075x + 0.2, & 3 \le x \le 5 \\ 0, & otherwise \end{cases}$.
   a. Graph the probability density function and verify that the total area under the density curve is indeed 1.
   b. Calculate $P(X \le 4)$. How does this probability compare to $P(X < 4)$?
   c. Calculate $P(3.5 \le X \le 4.5)$ and also $P(4.5 < X)$.
   d. Find $E(X), V(X)$.

14. The actual tracking weight of a stereo cartridge that is set to track at 3 g on a particular changer can be regarded as a continuous random variable X with probability density function $f(x) = \begin{cases} k[1-(x-3)^2], 2 \le x \le 4 \\ 0, otherwise \end{cases}$.
   a. Sketch the graph of $f(x)$.
   b. Find the value of k.
   c. What is the probability that the actual tracking weight is greater than the prescribed weight?
   d. What is the probability that the actual weight is within 0.25 g of the prescribed weight?
   e. What is the probability that the actual weight differs from the prescribed weight by more than 0.5 g?
   f. Find $E(X), V(X), \sigma_X$.

15. Let X denote the amount of time a book on two-hour reserve is actually checked out, and suppose the cumulative distribution function $F(x) = \begin{cases} 0, x < 0 \\ \frac{x^2}{4}, 0 \le x < 2 \\ 1, 2 \le x \end{cases}$. Use the cdf to obtain the following:
   a. $P(X \le 1)$
   b. $P(0.5 \le X \le 1)$
   c. $P(X > 1.5)$
   d. The median checkout duration.
   e. What is the 75th percentile?
   f. $F'(x)$ to obtain the density function $f(x)$.
   g. $E(X), V(X)$.

16. Let $f(x) = \frac{c}{1+x^2}$. For what values of c is $f(x)$ a probability density function?

17. If $f(x) = 30x^2(1-x)^2, 0 \le x \le 1$ is a probability density function. Find:
   a. $P(0 \le X \le 1)$
   b. The mean, i.e. calculate $\int_0^1 xf(x)dx = \mu$
   c. Use that information to find the variance $\sigma^2 = \int_0^1 (x-\mu)^2 f(x)dx$
   d. $P\left(X \le \frac{1}{3}\right)$

18. Let X have a binomial distribution with parameters $n = 25$ and p. Calculate each of the following probabilities using the normal approximation (with the continuity correction) for the cases $p = 0.5, 0.6, 0.8$, and compare each to the exact probability obtained from the binomial distribution.
   a. $P(15 \le X \le 20)$
   b. $P(X \le 15)$
   c. $P(20 \le X)$

19. Sandy's daughter Caroline got a 92 on a test with a mean of 84 and a standard deviation of 8 points. Her twin sister Lisa has a different teacher and got an 89 on a test with a mean of 75 and a standard deviation of 10. Which sister got the higher score? Assume the scores on both tests are normally distributed.