MATH 218: Homework problems

Homework 1.

1. Two dice are rolled. Introduce the following events:

(1) E : "the sum is odd"
(2) F : "at least one number is 1"
(3) G : "the sum is 5"

List the elementary outcomes in each of the following events: $E\bigcap F$, $E\bigcup F$, $F\bigcap G$, $E\bigcap \bar{F}$, $E\bigcap F\bigcap G$. For this problem, would you care whether the dice are fair?

2. Two fair dice are rolled. Compute the probability that the number on the first is smaller that the number on the second.

3. Let A, B, C be three events such that $P(A) = 0.5$, $P(B) = 0.6$, $P(C) = 0.8$
(a) Can any two of these events be mutually exclusive? Explain your conclusion. (b) Assuming that the events are independent, compute $P(A\bigcup B\bigcup C)$.

4. Let A and B be events such that $P(A) = 0.7$ and $P(B) = 0.8$.
(a) Circle the possible values of $P(A\bigcap B)$: 0.3 0.5 0.8 0.9
(b) Circle the possible values of $P(A\bigcup B)$: 07 0.8 1
You need to explain **each** of your conclusions. For example, if you think that $P(A\bigcap B)$ can be 0.5, you draw the corresponding Venn diagram, and if you think that $P(A\bigcup B)$ cannot be 1, you support your claim with suitable formulas.

5. A student is applying to MBA programs at Harvard, Yale, and MIT. Accordingly, the student prepares three personalized application letters and three addressed envelopes. Unfortunately, after three nights of heavy studying, the student is somewhat disoriented and places the letters in the envelopes at random. What is the probability that at least one letter ended up in the correct envelope?

6. At a certain school, 60% of the students wear neither a ring nor a necklace, 20% wear a ring, 30% wear a necklace. Compute the probability that a randomly selected student wears (a) a ring OR a necklace; (b) a ring AND a necklace.

7. A school offers three language classes: Spanish (S), French (F), and German (G). There are 100 students total, of which 28 take S, 26 take F, 16 take G, 12 take both S and F, 4 take both S and G, 6 take both F and G, and 2 take all three languages.

(1) Compute the probability that a randomly selected student (a) is not taking any of the three language classes; (b) takes EXACTLY one of the three language classes.
(2) Compute the probability that, of two randomly selected students, at least one takes a language class.

8. 30% of stopped drivers are drunk. Sobriety test is 95% accurate on drunk drives and 80% accurate on sober drivers. A driver is stopped and fails the sobriety test. (a) What is the probability that the driver is drunk? (b) After failing the sobriety test on the first try, the driver somehow gets a re-test and passes it. What is the probability that the driver is not drunk? Assume that the outcomes of the two tests are independent. (c) Now supposed that the driver failed the sobriety test twice in a row. How many more times should the driver fail the test to be 99% sure that the driver is drunk? Again, assume that the outcomes of the test are independent.

9. True or false: if A and B are events such that $0 < P(A) < 1$ and $P(B|A) = P(B|\bar{A})$, then A and B are independent?

Homework 2.

1. A coin is tossed n times. Let X be the difference between the number of heads and the number of tails. Find the possible values of X. Do we care whether the coin is fair or not?

2. A fair coin is tossed n times. Let X be the difference between the number of heads and the number of tails. Find the distribution of X when (a) $n = 3$; (b) $n = 4$.

3. Consider the following strategy for paying the roulette. Bet \$1 on red. If red appears (which happens with probability 18/38), then take the \$1 profit and stop playing for the day. If red does not appear, then bet additional \$1 on red each of the following two rounds, and then stop playing for the day no matter the outcome. Let X be the net gain/loss. (a) Find the distribution of X; (b) Compute $P(X > 0)$; (c) Compute the expected value of X; (d) Would you consider this a winning strategy?

4. Two fair dice are rolled. Define the following random variables: X, the value of the first die; Y, the sum of the two values; Z, the larger of the two values; V, the smaller of the two values. Find the joint distribution of (a) Z and Y; (b) X and Y; (c) Z and V.

5. Consider the function

$$f(x) = \begin{cases} C(2x - x^2) & 0 < x < 2 \\ 0 & \text{otherwise.} \end{cases}$$

(a) Could f be a cumulative distribution function? If so, find C.
(b) Could f be a probability density function? If so, find C.

6. The joint probability density function of two random variables X and Y

$$f_{XY}(x, y) = \Big\{ Cy,\ x^2 + y^2 \le 1,\ |x| \le 1, y \ge 0, 0 \qquad \text{otherwise.}$$

Determine the value of C and then compute $EX, EY, \mathrm{Var}(X), \mathrm{Var}(Y), \mathrm{Cov}(X, Y)$, $\mathrm{Cor}(X, Y), f_X(x), f_Y(y), f_{X|Y}(x|y), f_{Y|X}(y|x), P(|X| < 1/2|Y = 1/2)$.

7. The joint probability density function of two random variables X and Y is

$$f_{XY}(x, y) = c(y^2 - x^2)e^{-y},\ -y \le x \le y,\ 0 < y < +\infty.$$

Find (a) the value of c; (b) the marginal densities of X and Y; (c) expected value of X.

8. Given the joint density $f_{XY} = f_{XY}(x, y)$ of two random variables X, Y, decide whether the random variables are independent:

$$(a) \quad f_{XY}(x, y) = \begin{cases} xe^{-(x+y)}, & x > 0, y > 0; \\ 0, & \text{otherwise} \end{cases} \qquad (b) \quad f_{XY}(x, y) = \begin{cases} 2, & 0 < x < y,\ 0 < y < 1; \\ 0, & \text{otherwise} \end{cases}$$

9. Five men and five women are ranked according to their performance on a test. Assume that there are no ties (that is, no two people can have the same ranking) and that all possible ranking are equally likely. Let X be the ranking of the top woman. Find the distribution of X.

Homework 3.

1. A "traditional" three-digit telephone area code is constructed as follows. The first digit is from the set $\{2, 3, 4, 5, 6, 7, 8, 9\}$, the second is either 0 or 1, the last is from the set $\{1, 2, 3, 4, 5, 6, 7, 8, 9\}$. (a) How many area codes like this are possible? (b) How many such area codes start with 5?

2. In how many ways can three novels, two mathematics books and one chemistry book be arranged on a shelf if (a) any arrangement is allowed; (b) math books must be together and the novels must be together; (c) only the novels must be together.

3. Seven different gifts are distributed among 10 children so that no child gets more than one gift. How many different outcomes are possible.

4. In a certain jurisdiction, it takes at least 9 votes of a 12-member jury to get a conviction. Assume that
(1) 65% of all defendants are guilty;
(2) the probability that a juror will convict an innocent is 0.1;
(3) the probability that a juror will acquit a guilty is 0.2;
(4) each juror votes independently of the rest of the panel;

Compute the probabilities of the following events: (a) the panel renders a correct decision; (b) the defendant is convicted.
Note: You do need a computer to evaluate the probabilities numerically.

5. The number of fire alarms in a certain city has Poisson distribution. On average, there are 6 alarms every 24 hours. Compute the following probabilities: (a) To have three or four alarms in 24 hours. (b) To have no alarms in 4 hours. (c) To have more than one alarm in 36 hours. (d) To have at least one alarm in 20 minutes.

6. (a) A stick is broken into two pieces at random. Compute the probability that the ratio of the longer part to the shorter is at least a, where $a > 1$. (The length of the stick does not matter; put it equal to 1 if you want).
(b) A stick is broken into three pieces at random. What is the probability that the pieces are sides of a triangle?

7. Given a normal random variable X with mean 10 and variance 36, compute the following probabilities: (a) $P(X > 5)$; (b) $P(4 < X < 16)$; (c) $P(X < 8)$; (d) $P(X < 20)$; (e) $P(X > 16)$. Please use a table of the standard normal distribution.

8. Compute the variance of the normal random variable X if $E(X) = 5$ and $P(X > 9) = 0.2$. Please use a table of the standard normal distribution.

9. An urn contains four black and four white balls. Four balls are taken out of the urn. If two are black and two are white, the experiment ends. Otherwise, the balls are returned to the urn and the experiment is repeated. Denote by X the number of experiments conducted. Find the probability distribution of X. (Note: the probability of success is 18/35; start by verifying this).

Homework 4.

1. Let X be binomial random variable with parameters $n = 100$ and $p = 0.65$. Use normal approximation with the continuity correction to compute the following probabilities: (a) $P(X \geq 50)$; (b) $P(60 \leq X \leq 70)$; (c) $P(X < 75)$.

2. Let X be the number of cavities that develop in a 6-month period in the mouth of a child that uses the new brand of toothpaste "Cavifree". The distribution of X is shown below.

c	0	1	2	3
$P(X = c)$	0.4	0.3	0.2	0.1

a) A family has three children and they all use Cavifree. Assuming that the number of cavities acquired by any one child is independent of the number acquired by any other child, find the probability that between them they acquire at most one cavity in a 6-month period.
b) Find the expected value and the standard deviation of X.
c) A boarding school has 150 students and they all use Cavifree. Use the CLT to approximate the probability that the students acquire more than a total of 200 cavities in a 6-month period. (Again, you may assume that the number of cavities acquired by the different students are independent.)

3. The waiting time T, in minutes, for the green light at a certain intersection is a random variable with the following probability density function:

$$f_T(t) = \begin{cases} 3t^2, & 0 < t < 1, \\ 0, & \text{otherwise} \end{cases}$$

Using the CLT, find the approximate probability that, after driving through the intersection 60 times, you will have spent the total of more than 45 minutes waiting for the green light.

4. Using either a computer or pencil, paper and your knowledge, draw the normal probability plots for the following distributions. Then use your knowledge to explain why the graphs look the way they do.

$$f_1(x) = \frac{1}{2\sqrt{\pi}} e^{-x^2/4}; \ f_2(x) = \frac{1}{\pi(x^2+1)}; \ f_3(x) = cx^2 e^{-x^2}; \ f_4(x) = cx^4 e^{-x}; \ f_5(x) = cx^{-1/2} e^{-x^3}.$$

For f_3, f_4, and f_5, we assume the function to be zero for $x \leq 0$ and choose c so that the function is a probability density.

Homework 5.

1. The lifetime of a toaster from the company Toaster's Choice has a normal distribution with standard deviation 1.5 years. A random sample of 400 toasters was drawn yielding the sample lifetime average of 6 years.

a) Compute a 90% confidence interval for the mean lifetime of the toasters.
b) What sample size is needed to find the mean lifetime of the toasters to within plus or minus 0.05 years at the same 90% confidence level?
c) How will the answers in parts a) and b) change if, instead of knowing the standard deviation to be 1.5 years, it was estimated to be 1.5 years, based on the same sample of 400 devices.
d) Do parts a) and b) under the assumption that the lifetime has normal distribution, but with unknown standard deviation, and that a sample of 10 devices produced sample lifetime average of 6 years and sample standard deviation of 1.5 years.
e) Compare the intervals from parts a) and d). Which one is longer? Does it make sense? Why?
f) Compare the sample sizes in parts b) and d). Which one is larger? Does it make sense? Why?

2. In a survey of 100 people from a certain city, 20 claimed to have seen a UFO.
(a) Find the point estimate of the proportion of people in that city who claim to have seen a UFO.

(b) Construct the 98% confidence interval for the proportion of people in that city who claim to have seen a UFO.

(c) How many more people in that city must be surveyed to estimate the proportion of the people who claim to have seen a UFO to within $\pm 2\%$ with 98% confidence.

Homework 6.

1. A weight-loss company "Sleek and Slender" claims that the average weight loss of its customers is at least 25 pounds. After a bad experience with this company, Fred, an unsatisfied customer, wants to perform sampling in order to reject the claim of Sleek and Slender and possibly sue them. He decides to take 4 randomly chosen customers and to use a level of significance of 5%. He has obtained the data

$$13 \quad 10 \quad 20 \quad 25$$

representing their weight loss in pounds. He assumes that the weight-loss data for customers are normally distributed.

a) Formulate an appropriate null and alternative hypotheses for Fred to use.
b) State the rejection rule.
c) Compute the value of the test statistic.
d) Should Fred reject the null hypothesis at the 5% level of significance? Explain.
e) What can you say about the p-value for for this experiment. Circle one, *and explain.*
   (1) The p-value is less than 0.01.
   (2) The p-value is between 0.01 and 0.025.
   (3) The p-value is between 0.025 and 0.05.
   (4) The p-value is between 0.05 and 0.1.
   (5) The p-value is greater than 0.1.

2. After losing a game with friends, Alice suspects that the die which was used was not fair. She suspects that the probability of "1" appearing is not 1/6 and she decides to test this by rolling the die 300 times and using a level of significance of 5%.

a) Formulate an appropriate null and alternative hypotheses for Alice to use.
b) State the rejection rule.
c) After rolling the die 300 times, she noticed that "1" appeared 38 times. Compute the value of the test statistic in part b).
d) Should Alice reject the null hypothesis at the 5% level of significance? Explain.
e) Compute the p-value.

3. Consider the problem of testing the null hypothesis $a = 0$ against the alternative $a = 1$, where a is the mean of the normal population with standard deviation σ. Let n be the sample size and let y be the value of the test statistic.

(a) Sketch the graph of the p-value as the function of (i) y (ii) σ (iii) n
(b) Sketch the graph of the power of the test as the function of (i) y (ii) σ (iii) n

4. Suppose that the distribution of the test statistic to test the null hypothesis $a = 0$ against the alternative $a = 1/2$ is $f_0(x) = 2(1-x),\ 0 < x < 1$, under the null hypothesis and $f_1(x) = 2x,\ 0 < x < 1$, under the alternative. Suppose that the critical region is $[c, 1]$ and the observed value of the test statistic is y, where c and y are numbers between 0 and 1.

Compute the p-value of the experiment and the power of the test, as functions of y and sketch the corresponding graphs. What do you expect from the power and p-value as $y \to 0$? As $y \to 1$? Do you expect the functions to be monotone? Are you getting the behavior you expect?