CSCI2244-Randomness and Computation
Final Exam with Solutions

May 11, 2018

Each part of Problems 1 and 4 is worth 5 points. Each part of Problems 2 and 3 is worth 10 points. That makes a total of 135 possible points. Show work with *every problem*: In many cases this will be a computation, in some just an account of your reasoning. But just giving a number is not a correct response. You may cite any result that was discussed in class as part of your justification.

1 Some questions with pictures.

These questions are designed to test familiarity with fundamental concepts and involve very little or no calculation.

(a) Figure 1 shows a sample space S whose elements are the black dots, and two events, E and F. Assume a uniform probability function. Find $P(E)$, $P(F)$, $P(E|F)$ and $P(F|E)$. Are E and F independent?

Solution. There are 9 dots in all, so

$$P(E) = \frac{6}{9} = \frac{2}{3}, P(F) = \frac{5}{9}, P(E \cap F) = \frac{3}{9} = \frac{1}{3}.$$

As a result,

$$P(E|F) = \frac{P(E \cap F)}{P(F)} = \frac{3}{5}, P(F|E) = \frac{P(E \cap F)}{P(E)} = \frac{1}{2}.$$

Also

$$P(E) \cdot P(F) = \frac{10}{27} \neq \frac{1}{3} = P(E \cap F),$$

so the E and F are not independent.

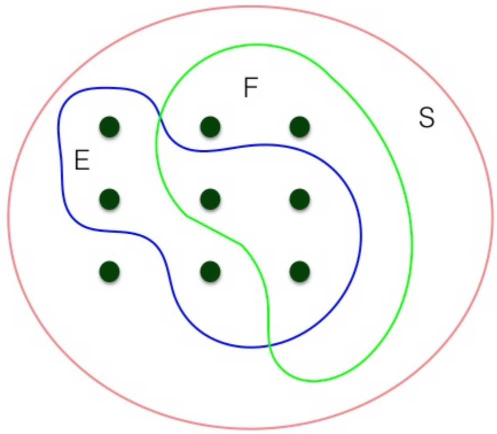

Figure 1: Probability space and events for 1(a) .

(b) Figure 2 shows two identical plots, however the scales on the vertical and horizontal axes are different, and in one instance the scale on the horizontal axis is not even shown. (Assume that the values for positive x tend to 0 as $x \to \infty$.) Tell which of these can, or cannot, be the density function for a random variable, and explain.

Solution. In the diagram on the left, there is a rectangle with base 2 and height 0.5 contained within the region under the graph. Since the area of the rectangle is 1, the area under the graph is strictly greater than 1, and

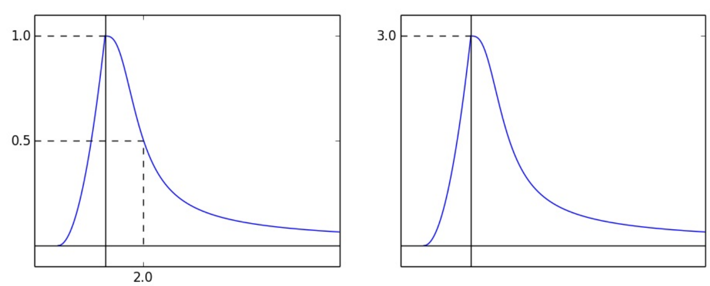

Figure 2: Candidate density functions for 1(b).

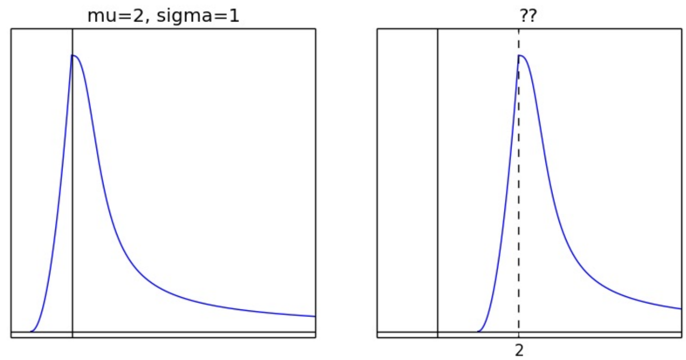

Figure 3: Density function and shifted version, for 1(c).

thus it can not be a density function. The graph on the right can be the graph of a density function. As long as the area under the graph is finite, we can adjust the vertical scale so that the area is 1.

(c) The same plot appears again on the left hand side of Figure 3. Assume that the scaling on the axes has been chosen so that this *is* the graph of the density function of a random variable, with expected value $\mu = 2$ and standard deviation $\sigma = 1$. On the right-hand side, the same plot appears, shifted two units to the right. Is this also the density function of a random variable? If so, what are its expected value and standard deviation?

Solution. Shifting the density two units to the right like this increases the mean by 2 and leaves the variance unchanged. So the right-hand figure has mean 4 and standard deviation 1.

(d) Figure 4 shows the left-hand density function from the previous problem, with a portion shaded in. Find a lower bound for this area; that is, a number guaranteed to be smaller than this area. Yes, of course, 0 is a lower bound, but you can do much better than that.

Solution. The shaded area is the probability that the random variable is within 2 standard deviations of the mean. The complementary probability is bounded above by Chebyshev's inequality:

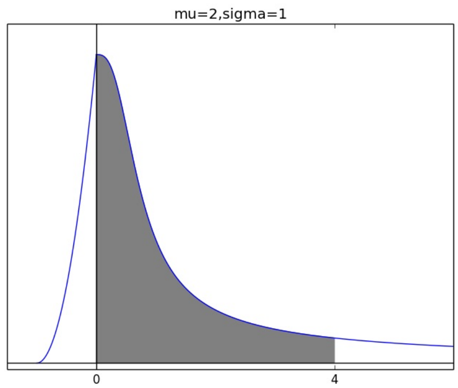

Figure 4: Shaded area, for 1(d).

$$P(|X-\mu|>k\sigma)<\frac{1}{k^2}=\frac{1}{4},$$

so the area of the shaded region is at least $\frac{3}{4}$.

2 Strange Dice

Your usual six-faced die is in the shop for repairs, so you decide to craft a crude substitute: You make one fair 3-faced die[1] with faces numbered 1,2,3; and a fair 4-faced die, with faces numbered 0,1,2,3. Thus the possible values for the sum of these two dice are 1,2,3,4,5,6, just as with a single standard die, although of course the probabilities are different. Let Z be the random variable giving the sum of the numbers showing on the two dice.

(a) What is $P(Z \geq 5)$? What is $P(Z = 6)$? **Solution.** Each outcome (i, j) has probability $\frac{1}{12}$. There are two such outcomes that make $Z = 5$, namely $(2, 3)$ and $(3, 2)$, and only one that makes $Z = 6$, namely $(3, 3)$. Thus $P(Z \geq 5) = \frac{3}{12} = \frac{1}{4}$, and $P(Z = 6) = \frac{1}{12}$.

(b) It is easy to check that $E(Z) = 3.5$, as for a single standard die. Find var(Z). (HINT: Work smart, not hard! You do not need to know the PMF of Z to compute this; that will only make the arithmetic more arduous.)

Solution. Working smart means using the additivity of variance. The variance for the first die is

$$\frac{1}{3} \cdot (1^2 + 2^2 + 3^2) - 2^2 = \frac{2}{3},$$

and for the second

$$\frac{1}{4} \cdot (0^2 + 1^2 + 2^2 + 3^2) - 1.5^2 = \frac{14}{4} - \frac{9}{4} = \frac{5}{4}.$$

Thus the variance of Z is

$$\frac{2}{3} + \frac{5}{4} = \frac{23}{12}.$$

(c) Roll the pair of strange dice twelve times in succession. What is the probability of getting a 6 at least 3 times? Your final answer should be an expression involving only sums, differences, products, quotients and powers

[1] What would a *three*-faced die look like? I have no idea! Don't think too hard about it.

of explicit numbers (for example, you should not include any factorials or binomial coefficients in your answer), but you do not need to evaluate this expression.

Solution. The number of 6's in twelve rolls has the binomial distribution. The expression for the complementary probability (the probability of getting fewer than three 6's) is simpler. So the answer is

$$1 - \sum_{j=0}^{2} \binom{12}{j} \left(\frac{1}{12}\right)^{j} \left(\frac{11}{12}\right)^{12-j} = 1 - \frac{1}{12^{12}} \cdot (11^{12} + 12 \cdot 11^{11} + 66 \cdot 11^{10}).$$

This is about 0.072.

(d) There is a simple and pretty good approximation to this probability, if you happen to have the values of the exponential function e^x handy. (You can leave the answer as an unevaluated expression involving the constant e, but the expression should be much simpler and relatively easy to evaluate by hand. In this regard it is helpful to know that $e \approx 2.72$.)

Solution. The number of 6's is approximately Poisson-distributed with $n = 12$ (number of rolls) and $p = \frac{1}{12}$ (probability of a 6). Thus $\lambda = 1$. The complementary probability (fewer than 3 6's) is thus approximately

$$e^{-1} \cdot (1 + 1 + \frac{1}{2}) = 2.5/e \approx 0.92,$$

so that the probability is approximately 0.080. (That's not a great approximation—the relative accuracy for the complementary probability is pretty good, but here the actual probability is off by one in the most significant digit. The issue is that $n = 12$ is a bit too small for the Poisson approximation to be very accurate. So this may not have been the best-designed problem.)

(e) You roll the pair of dice 1000 times in succession, and keep a running total of the sum. Estimate the probability that the resulting sum is between 3420 and 3550 inclusive. You can write your answer as an unevaluated expression using the symbol Φ for the cumulative standard normal distribution.

Solution. We let S_{1000} denote the sum of the one thousand rolls. S_{1000} has mean 3500 and standard deviation

$$\sqrt{23000/12} \approx 43.8.$$

We are asked for the probability

$$P(3420 \leq S_{1000} \leq 3550).$$

We normalize by subtracting 3500 and dividing by the standard deviation. (Incidentally, it is okay in solving this problem on an exam to just leave expressions like $\sqrt{23000/12}$ in your solution.) So the probability is

$$P(-80/43.8 \leq X \leq 50/43.8),$$

where $X = (S_{1000} - 3500)/43.8$. By the Central Limit Theorem, X has approximately the standard normal distribution, so the answer is about

$$\Phi(50/43.8) - \Phi(-80/43.8) \approx 0.84.$$

(f) You roll the pair of dice N times in succession and keep a running total S_N of the sum. What is

$$\lim_{N\to\infty} P(S_N < 3.5N + 50)?$$

Solution. The limit is $\frac{1}{2}$. If we normalize S_N by subtracting the mean and dividing by the standard deviation, we get $P(X < 50/(c\sqrt{N}))$, where $c = \sqrt{23/12}$. The value of c is actually immaterial here; the point is that the right-hand side of the inequality approaches 0 as N grows larger, so we get $P(X < 0) \longrightarrow \Phi(0) = \frac{1}{2}$.

3 Another Darts Game

You play a game in which you throw a dart at a circular board 1 unit in radius. Unfortunately, your eyeglasses broke and you can barely see the board, with the result that your shots that do hit the board are uniformly distributed. If you miss the board entirely, you get a do-over. (It is helpful to remember that the area of a circle of radius R is πR^2.)

(a) A circular red target of radius r is painted in the center of the board. If you hit the target, you win; if you miss, you lose. How big does r have to be for this game to be fair? (That is, to give you a $\frac{1}{2}$ probability of hitting the target.)

Solution. The area of the board is π, and of the red target πr^2, so the probability of hitting it is r^2. Thus the probability becomes $1/2$ when $r = 1/\sqrt{2} \approx 0.707$.

Let's change the game. In this version, there is no red target. Instead, for each throw, you receive $1 - d$ dollars, where d is the distance of your dart from the center. Let X be the random variable denoting this payout.

(b) What is $P(X \geq \frac{1}{4})$? What is $P(X = \frac{1}{4})$?

Solution. $X \geq \frac{1}{4}$ is equivalent to $d \leq \frac{3}{4}$, so the probability of this is $\left(\frac{3}{4}\right)^2 = \frac{9}{16}$. $P(X = \frac{1}{4}) = 0$, of course.

(c) Compute and plot both the cumulative distribution function and the density function of X.

$$P(X < x) = P(d \geq 1-x) = 1 - P(d \leq 1-x) = 1 - (1-x)^2 = 2x - x^2.$$

so this gives the value of the cdf in the interval $[0, 1]$. The cdf is 0 at negative values and 1 for values greater than 1. The density function is given by the derivative

$$f_X(x) = 2 - 2x$$

for $x \in [0, 1]$ and 0 outside this interval. The plots are shown below.

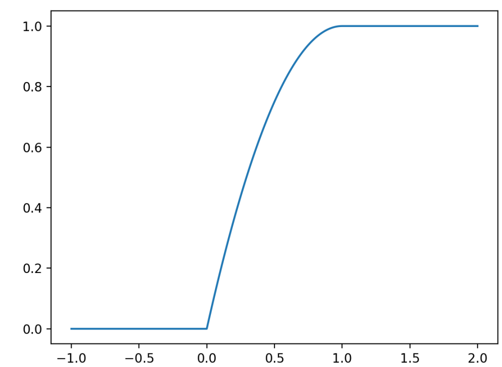

Figure 5: CDF for the darts problem.

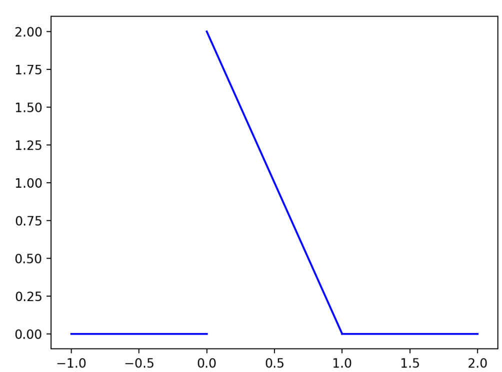

Figure 6: Density for the darts problem.

(d) How much should you pay to play this game? *I.e.,* What is the most that you should pay to throw a dart so that you at least break even over the long run?

Solution. This problem is asking for the expected winnings, which we find by integrating x times the density function. We get

$$E(X) = \int_0^1 (2x - 2x^2)dx = (x^2 - 2x^3/3)\Big|_0^1 = \frac{1}{3},$$

so you should not bet more than 33 cents on this game.

4 Leave a Hamilton, take a Hamilton

You know those little trays you often see next to the cash register in a convenience store—the ones that say 'Need a penny, take a penny; have a penny leave a penny'? Here we will perform an experiment testing people's generosity by playing this game with ten-dollar bills instead of pennies.

There is a jar on the cash counter, initially empty: At regular intervals a customer is let into the store and looks at the jar. If the jar is empty, 10% of the customers will leave a ten-dollar bill, 90% will do nothing. If the jar already contains a ten-dollar bill, 30% of the customers will take it, and 10% will leave a second ten-dollar bill, and the remaining customers will do nothing. The experiment ends once there are two ten-dollar bills in the jar.

(a) Represent this experiment as a Markov chain with three states 0,1,2, where the state is the number of ten-dollar bills in the jar. Draw both the state-transition diagram and show the transition matrix.

Solution. I'll leave off the transition diagram. Here is the transition matrix:

$$\begin{bmatrix} 0.9 & 0.1 & 0 \\ 0.3 & 0.6 & 0.1 \\ 0 & 0 & 0.1 \end{bmatrix}.$$

(b) What is the probability that the game ends after two steps? after 3 steps? (HINT: It is not necessary to multiply out the complete matrices to do this.)

Solution. The game ends in two steps if both of the first two customers leave a ten-dollar bill, which happens with probability $0.1 \times 0.1 = 0.01$. For the game to end in three steps, there are only two possible sequences: do nothing, leave a bill, leave a bill; leave a bill, do nothing, leave a bill. So the probability is $0.9 \times 0.1 \times 0.1 + 0.1 \times 0.6 \times 0.1 = 0.15$.

(c) What is the average number of customers that have to enter the store before the jar contains two ten-dollar bills?

Solution. The problem asks for the expected time to absorption. This is the sum of the entries in the first row of the inverse of $I - Q$, where Q is the matrix given by the first two rows and first two columns of the transition matrix for the chain. We have

$$I - Q = \begin{bmatrix} 0.1 & -0.1 \\ -0.3 & 0.4 \end{bmatrix}.$$

The first row of the inverse of this matrix is

$$0.4/D \quad 0.1/D,$$

where

$$D = 0.1 \times 0.4 - 0.1 \times 0.3 = 0.01.$$

So the entries in the row are 40 and 10. Thus it takes on average fifty customers to enter the store for the jar to collect two ten-dollar bills.

Helpful matrix algebra fact. You may need to calculate the inverse of a 2×2 matrix to solve part (c). For matrices of this size, there is a simple formula for the inverse:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \begin{bmatrix} d/D & -b/D \\ -c/D & a/D \end{bmatrix},$$

where $D = ad - bc$. (D is called the *determinant* of the original matrix.)