July 29, 2023

Probability Theory: Homework problems

Homework 1.

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. Verify the odds of various events in the California State Lottery game MEGA Millions, as shown on
`https://www.calottery.com/draw-games/mega-millions#section-content-3-3`

4. (I) Seven different gifts are distributed among 10 (different) children. How many different outcomes are possible if every child can receive (a) at most one gift; (b) at most two gifts; (c) any number of gifts
(II) Answer the same questions if the gifts are identical (but children are still different).

5. (I) Twenty different gifts are to be distributed among 7 (different) children. How many outcomes are possible if every child is to receive (a) at least one gift (b) at least two gifts (c) any number of gifts.
(II) Answer the same questions if the gifts are identical (but children are still different).
(III) Now try to generalize problems 4 and 5.

6. Consider the two-dimensional Cartesian (standard rectangular x, y) coordinate system on the plane. You move around the points with integer coordinates in such a way that, at each step you can go either one unit up or one unit to the right (that is, from $(0, 0)$ you can go to either $(1, 0)$ or $(0, 1)$; from $(2, 3)$ you can go either to $(3, 3)$ or $(2, 4)$, etc.) Count the number of different paths from the point $(0, 0)$ to (a) the point $(4, 3)$; (b) the point $(4, 3)$, if you have to visit the point $(2, 2)$; (c) the point $(4, 4)$, if you are not allowed to go above the line $x = y$ (but you are allowed to hit the line, e.g. by visiting point the $(1, 1)$).

7. You have \$20K to invest, and have a choice of stocks, bonds, mutual funds, or a CD. Investments must be made in multiples of \$1K, and there are minimal amounts to be invested: \$2K in stocks, \$2K in bonds, \$3K in mutual funds, and \$4K in the CD. Count the number of choices in each situation: (a) You want to invest in all four; (b) You want to invest in at least three out of four.

8. 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: EF, $E \bigcup F$, FG, EF^c, EFG. For this problem, would you care whether the dice are fair?

Homework 2.

1. 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.

2. 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.

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

4. Alice and Bob are in a group of N people who are randomly arranged (a) in a row (b) in a circle. In each case, compute the probability that Alice and Bob are next to each other.

5. Two fair dice are rolled. For $i = 2, 3, \ldots, 12$, compute the probability that the first one shows 6 given that the sum is i. (Of course, it is zero if $i < 7$).

6. (I) Let A and B be two events such that $P(A) = 0.5$ and $P(B) = 0.6$. (a) Can A and B be mutually exclusive? (b) Assuming that A and B are independent, compute $P(A \bigcup B)$.

(II) 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? (b) Assuming that the events are independent, compute $P(A \bigcup B \bigcup C)$.

(III) 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 $P(A \bigcup B)$: 0.7 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 can draw the corresponding Venn diagram, and if you think that $P(A \bigcup B)$ cannot be 1, you can support your claim with suitable formulas.

7. True or false: if A and B are events such that $0 < P(A) < 1$ and $P(B|A) = P(B|A^c)$, then A and B are independent?

8. The probability that a student passes the first midterm is p, the probability that the student passes the second midterm is q, and the probability that the student passes at least one of the midterms is r. Identify the constraints on the possible values of p, q, r, and then compute the probability s, in terms of p, q, r, that the student passes both midterms. [$0 \le p, q, r \le 1, p + q \ge r$; by inclusion/exclusion, $p + q - s = r$, so $0 \le s = p + q - r \le 1$]

Homework 3.

1. In a certain community, 36% of all the families have a dog and 30% have a cat. Of those families with a dog, 22% also have a cat. Compute the probability that a randomly selected family (a) has both a dog and a cat; (b) has a dog GIVEN that it has a cat.

2. Three fair dice have different colors: red, blue, and yellow. These three dice are rolled and the face value of each is recorded as R, B, Y, respectively. Compute the probability that $B < Y < R$. You can proceed as follows: (a) compute the probability that no two dice show the same number; (b) compute the probability that $B < Y < R$ given that all the numbers are different; (c) solve the problem.

3. Suppose that 5% of cars and 0.25% of trucks are yellow. Compute the probability that a randomly selected yellow vehicle is a car if (a) the proportion of cars and trucks in the population is that same; (b) there are twice as many trucks as cars in the population.

4. English and American spellings are *rigour* and *rigor*, respectively. At a certain hotel, 40% of guests are from England and the rest are from America. A guest at the hotel writes the word (as either *rigour* or *rigor*), and a randomly selected letter from the spelling turns out to be a vowel. Compute the probability that the guest is from England.

5. Two people, A and B, are involved in a duel. The rules are as follows: shoot at each other once; if at least one is hit, the duel is over, if both miss, repeat (go to the next round), and so on. Denote by p_A and p_B the probabilities that A hits B and B hits A with one shot, and assume that hitting/missing is independent from round to round. Compute the probabilities of the following events: (a) the duel

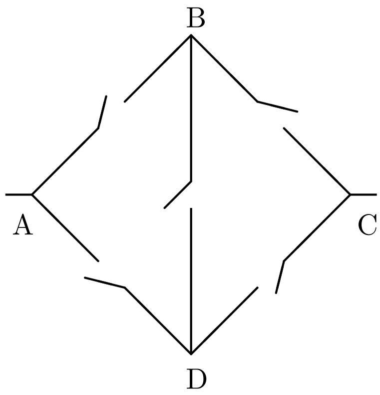

FIGURE 1. A random connection

ends and A is not hit; (b) the duel ends and both are hit; (c) the duel ends after round number n; (d) the duel ends after round number n GIVEN that A is not hit; (e) the duel ends after n rounds GIVEN that both are hit; (f) the duel goes on for ever.

6. On Figure 1, each of the five connections can be open or closed independently of other connections. The probability to have a specific connection closed is p.
(a) Compute the probability that there is a path of closed connections from A to C. [One possible solution: by inclusion-exclusion. To keep track of what you are doing, it might actually be easier to denote the probability of each connection by a different letter].
(b) Compute the conditional probability that the connection along the diagonal BD is closed given that there is a path of closed connections from A to C. [This question can have two different answers because you can consider some of the closed connections NOT considered in part (a)]
Whatever answer you get in both parts (a) and (b), check that the result is a function that is monotonically increasing from 0 when $p = 0$ to 1 when $p = 1$ (it might be easier to do it using a computer algebra system).

7. Thirty percent of drivers stopped by police are drunk. The sobriety test is 95% accurate on drunk drivers 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) How many times should the sobriety test be administered to be 99.999% sure that the drive is drunk? Assume that, when the sobriety test is administered several times, the test results are independent.

Homework 4.

1. Five USC students and five UCLA students are ranked according to their performance on a test. Assume that no two test scores are the same and all possible rankings are equally likely. Let X be the highest ranking of a USC student. Compute the distribution of X.

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

3. A fair coin is tossed n times. Let X be the difference between the number of heads and the number of tails. Compute the distribution of X when (a) $n = 3$; (b) $n = 4$; (c) $n \geq 2$ is arbitrary. How is the distribution of X different from the binomial distribution?

4. 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) Describe the distribution of X; (b) Compute $P(X > 0)$; (c) Compute the expected value of X; (d) Would you consider this a winning strategy? Explain your reasoning.

5. Two teams play a series of games until one of the teams wins n games. In every game, both teams have equal chances of winning and there are no draws. Compute the expected number of the games played when (a) $n = 2$; (b) $n = 3$. (To keep track of what you are doing, it can be easier to use different letters for the probabilities of win for the two teams).

6. A communication channel transmits a signal as sequence of digits 0 and 1. The probability of incorrect reception of each digit is p. To reduce the probability of error at reception, 0 is transmitted as 00000 (five zeroes) and 1, as 11111. Assume that the digits are received independently and the majority decoding is used. Compute the probability of receiving the signal incorrectly if the original signal is (a) 0; (b) 101. Evaluate the probabilities when $p = 0.2$.

7. 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. Use your favorite software package to evaluate the probabilities numerically.

Homework 5.

1. On a certain highway, there are, on average, 2.2 cars abandoned every week. Assuming Poisson distribution, compute the probability that (a) there will be no cars abandoned next week; (b) there will be at least 5 cars abandoned next month.

2. Suppose that the number of times a person catches cold in a year is a Poisson random variable with parameter $\lambda = 5$. A new drug is claimed to reduce this parameter λ to 3 for 75% of the population and has no effect on the rest of the population. Somebody takes the drug and gets two colds in a year. Compute the probability that the drug was beneficial for that person. (Note: this is a classical Bayes rule problem).

3. At time $t_0 = 0$, a fair coin is tossed and lands heads. Then, at a random time $T > 0$, the coin is tossed again. Given a $t > 0$, compute the probability that the coin shows heads at time t, if T is the moment of the first event in a Poisson process with parameter λ. How will the answer change if the coin is not fair?

4. 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. Decribe the probability distribution of X. (Note: the probability of success is $18/35$; start by verifying this).

5. Let X be a random variable with the cumulative distribution function (cdf) $F = F(x)$ and let α, β be real numbers with $\alpha \neq 0$. Determine the cdf of the random variable $\alpha X + \beta$. (Keep in mind that $F(x) = P(X \leq x)$ and you cannot assume that F is continuous.)

6. For $p \in (0, 1)$, let $x(p)$ be the smallest number of people so that there is a better than $100 \cdot p\%$ chance to have at least two born on the same day. Derive an approximate expression for $x(p)$ [it is $\sqrt{2n \ln(1/q)}$, $n = 365, q = 1 - p$] and sketch the graph of the function $x = x(p)$.

Homework 6.

1. 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, compute the value of C.
(b) Could f be a probability density function? If so, compute the value of C.

2. (I) 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).
(II) A stick is broken in three pieces at random. Compute the probability p that the pieces are sides of the triangle (that it, the sum of any two is bigger than the third). Consider the following three possibilities (a) selecting two iid uniforms on the stick [then $p = 1/4$, best seeing by thinking of the stick as the height of an equilateral triangle]; (b) breaking the stick at a uniform point in two pieces, then taking the longer piece and breaking it again [then $p = 2(\ln 2 - 0.5)$]; (c) breaking the stick at a uniform point in two pieces, then taking one of the pieces with probability proportional to the length of the piece and breaking it again [then $p = 1/4$].

3. 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) $P(X > 16)$. Either a table or statistical function on your calculator can be used to produce the numerical answers. For better results, reduce to standard normal before computing the numerical answer.

4. (a) Compute the standard deviation of a normal random variable X if $E(X) = 5$ and $P(X > 9) = 0.2$. (b) Compute the expected value of a normal random variable X if $\mathrm{var}(X) = 3$ and $P(X > 5) = 0.2$. (c) Determine the mean and the standard deviation of a normal random variable X in each of the two cases:

$$\begin{aligned} \text{Case 1}: \;\; P(X > 0) &= \frac{1 - 0.5763}{2} \quad \text{and} \quad P(X < 2) = \frac{1 + 0.8664}{2}; \\ \text{Case 2}: \;\; P(X < 0) &= \frac{1 - 0.5763}{2} \quad \text{and} \quad P(X < 2) = \frac{1 + 0.8664}{2}. \end{aligned}$$

Note. The numbers 0.5763 and 0.8664 come directly from a Z-table, and should lead you to the right Z-value. The Z-value corresponding to 0.5763 is ± 0.8; to figure out whether you take positive or negative value, draw a picture. (d) Create your own problem similar to (c). What are the constraints on the numbers you are using?

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

6. The number of miles a certain car can drive before breaking is a random variable X. The car has been driven for 10000 miles. Compute the probability that the car will drive another 20000 if the distribution of X is (a) exponential, with average value 20000; (b) unform on $(0, 40000)$.

7. Let X be exponential random variable with mean 1. Determine the probability density function of $\ln X$.

8. Let X be uniform on $(0, 1)$. Determine the probability density function of e^X.

9. Let X be a continuous random variable, and assume that the cdf F_X of X is a strictly increasing function. Identify the distribution of the random variable $U = F_X(X)$. What, if anything, changes without the assumption that F_X is strictly increasing? What if X is not continuous?

Homework 7.

1. 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. Determine the joint distribution of (a) Z and Y; (b) X and Y; (c) Z and V.

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

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

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

3. Two people decide to meet at a certain location. The arrival time of person A is uniformly distributed between 12:15pm and 12:45pm. The arrival time of person B is uniformly distributed between 12pm and 1pm. The two people arrive independently of each other. (a) Compute the probability that the first to arrive will wait at most five minutes. (b) Compute the probability that person A arrives first. [Draw a picture].

4. Compute the probability that n points, selected randomly and independently on a circle, will be in the same semi-circle. (Suggestion: Fix a point P_i and denote by A_i the event that all points starting with P_i and going clockwise are in the same semi-circle. Argue that A_i and A_j are mutually exclusive and that the probability of each A_i is 2^{-n+1}.)

5. Two points X, Y are selected at random on the interval $[0,1]$ so that X is uniform on $(0,1/2)$, Y is uniform on $(1/2,1)$ and X, Y are independent. Compute $P(Y - X > 1/3)$.

6. Given the joint density $f = f(x,y)$ of two random variables X, Y, decide whether the random variables are independent:

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

7. Player A's bowling score in one game is approximately normal with mean 170 and standard deviation 20; player B's score in one game is approximately normal with mean 160 and standard deviation 15. Assuming that the scores are independent, compute the probability that, in one game, (a) Player A scores higher than player B; (b) The total score is over 350.

Homework 8.

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

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

Compute the conditional distribution of Y given $X = x$.

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

$$f_{XY}(x,y) = \begin{cases} Cy & \text{if } \ x^2 + y^2 \le 1, \ |x| \le 1, y \ge 0, \\ 0 & \text{otherwise.} \end{cases}$$

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

3. Three cars break down on a road of length L, randomly and independently of one another. Given a $d < L/2$, compute the probability that the distance between any two of the cars is at least d. (Keep in mind that there are six possible arrangements of the cars).

4. The random variables X, Y have the joint density

$$f(x,y) = \begin{cases} \frac{1}{\pi}, & x^2 + y^2 \le 1 \\ 0, & \text{otherwise.} \end{cases}$$

In other words, the vector (X,Y) is uniform in the unit disk. Compute the joint density of $\sqrt{X^2 + Y^2}$ and $\tan^{-1}(Y/X)$.

5. Let U, Z be independent random variables such that U is uniform on $(0, 2\pi)$ and Z is exponential with mean 1. (a) Show that $X = \sqrt{2Z}\cos U$ and $Y = \sqrt{2Z}\sin U$ are independent standard normal

random variables. (b) Confirm that $\mathbb{E}(Z|X=0)=1/2$ [because $X^2+Y^2=2Z^2$] but $\mathbb{E}(Z|\cos U=0)=1$ [because Z and U are independent] and comment on the result [$\mathbb{P}(X=0)=\mathbb{P}\big((Z=0)\cup(\cos U=0)\big)=0$; this is a version of the `Borel Paradox`].

6. Let X, Y be independent random variables, both uniform on $(0,1)$. Compute the joint density of the following random variables; (a) $X+Y,\ X/Y$; (b)$X,\ X/Y$, (c) $X+Y,\ X/(X+Y)$.

Then repeat the problem when X and Y are independent random variables having Gamma distribution Gamma(α,β) with the same scale $1/\beta$ (or rate β) parameter but possibly different shape parameters α.

7. Let X and Y be independent standard normal random variables. Define $U=X+Y$ and $V=X-Y$.
(a) Confirm that U and V are independent.
(b) Compute $E(X+2Y|U)$ and $E(X+2Y|V)$.

8. Consider a symmetric 2-by-2 matrix W such that the diagonal entries X, Y and the above-the-diagonal entry U are independent. In particular, the trace of W is $X+Y$ and the determinant of W is $XY-U^2$. Denote by L the distance between the eigenvalues of W. Assume that X and Y are standard Gaussian and U is Gaussian with mean zero and variance 1/2. Compute the pdf of L. [The answer is $(x/2)e^{-x^2/4}$, $x>0$, and there is no need for any Jacobians: after some computations, $L=\sqrt{(X-Y)^2+4U^2}$, which, in distribution, is the same as $\sqrt{R}$, where R is exponential random variable with mean 4.]

Homework 9.

1. Compute the expected winning in the following game. A fair die is rolled and a fair coin is tossed. If the coin lands heads, the winning amount is the twice the number on the die. If the coin lands tails, the winning amount is half the number on the die.

2. A fair die is rolled n times. Compute the expected value and the variance of the sum.

3. Two people, A and B, choose randomly and independently three objects out of 10. Compute the expected number of objects (a) chosen by both A and B; (b) chosen by neither A nor B; (c) chosen by exactly one person.

4. 1000 cards with numbers from 1 to 1000 randomly distributed among 1000 people. Assume there is no upper bound on age. Denote by X the number of people whose age is the same as the number on the card they got. (a) Compute the expected value of X [here, it does not matter whether there are two or more people have the same age]. (b) What is the range of possible values of the variance of X? [This part is hard; it is all about the distribution of the ages among the people. In particular, if everybody has the same age, then the variance will be zero].

5. The airplane has 101 seats. The first 100 people who showed up took seats at random, but the last one insisted on taking the assigned seat. Compute the expected value and the variance of the number of people who will have to move (a displaced person moves to the corresponding assigned seat; if the seat is occupied, whoever occupies it moves to the corresponding assigned seat, etc. As a result, any number of movements, from 0 to 100 is possible)

6. Let $X_1, X_2, \ldots$ be iid having moment generating function $M_X=M_X(t)$ defined for all $t\in(-\infty,+\infty)$. Let N be an integer-valued random variable (that is, N takes values $0,1,2,\ldots$) with probability generating function $G_N=G_N(z)$ defined for all $z\in(-\infty,+\infty)$. Assume that N is independent of all X_k and define $S=\sum_{k=1}^N X_k$. Confirm that the random variable S has the moment generating function $M_S=M_S(t)$ defined for all $t\in(-\infty,+\infty)$, and

$$M_S(t)=G_N\big(M_X(t)\big).$$

Then use the result to derive the formulas

$$E(S)=\mu_N\mu_X,\ \mathrm{Var}(S)=\sigma_N^2\mu_X^2+\mu_N\sigma_X^2,$$

where $\mu_N = E(N),\ \mu_X = E(X_1),\ \sigma_N^2 = \mathrm{Var}(N),\ \sigma_X^2 = \mathrm{Var}(X_1)$.

Homework 10.

1. (A) For a group of 100 people, compute the expected value and the variance of (a) the number of days in a 365-day year that are birthdays of exactly three people; (b) the number of distinct birthdays.

(B) As an abstract version of the previous problem, assume that m balls are thrown into n boxes so that the balls thrown independently of one another and each ball is equally likely to land in every box. For each $k = 0, \ldots, m$, compute the expected value of number of boxes with exactly k balls.

2. Compute the expected value and the variance of the number of rolls of a fair die before all sides appear at least once.

3. Let $X_1, X_2, \ldots$ be independent identically distributed continuous random variables. Define the random variable N as follows:

$$X_1 > X_2 > \ldots > X_{N-1} < X_N$$

Compute the expected value of N. (Hint: start by showing that $P(N \geq n) = 1/(n-1)!$).

4. Let $X_1, \ldots, X_4$ be random variables such that $EX_i = 0$, $EX_i^2 = 1$, $i = 1, \ldots, 4$; $E(X_iX_j) = 0$, $i \neq j$. Compute the correlation of (a) $X_1 + X_2$ and $X_1 + X_3$; (b) $X_1 + X_2$ and $X_3 + X_4$.

5. Consider a graph on n vertices. Define the degree D_i of the vertex i as the number of edges coming out of the vertex. Assume that between any two different vertices, an edge is preset with probability p, independent of all other edges. Compute the distribution of D_i and the correlation between D_i and D_j.

6. A fair die is rolled repeatedly. Denote by X the number of rolls necessary to get a 6 for the first time. Denote by Y the number of rolls necessary to obtain 5 for the first time. Compute (a) $E(X)$; (b) $E(X|Y = 1)$; (c) $E(X|Y = 5)$.

7. Consider a sequence of independent tosses of a fair coin with outcomes H and T.

(a) Compute the probability that HH will appear before HT [It is 1/2].

(b) Compute the expected number of tosses to get HH.

[The answer is 6. Indeed, if the number we need is x, then $x = (E_H+E_T)/2$, where E_C is the expected number of tosses to get HH if the first toss is C. Then $E_H = 1+(1+E_T)/2$, and $E_T = 1+(E_T+E_H)/2$.]

(c) Compute the expected number of tosses to get HT [The answer is 4.]

(d) Come up with an alternative (qualitative) explanation why the answer in part (b) is bigger than the answer in part (c).

8. Two people take turns tossing a fair coin. The winner is the one who completes the pattern HH.

(a) Would you rather go first or second? [Note: if you start, you cannot win on your first toss; if you go second, you can]

(b) Compute the probability of winning if you go first. [If p is the required probability, then $p = (h+t)/2$, where h is the probability of winning given that the first toss was H, and t is the probability of winning given that the first toss was T. Then $h = (1/2)(h+t)/2$ (if the first toss is H, the game will only continue if the second toss is T, and then it is like starting anew), and $t = (1/2)(h+t)/2 + (1/2)((1/2) + t/2)$ (if the first toss is T and the second is also T, then it is starting anew; if the first toss is T and the second is H, then, on the third toss, you either win with probability $1/2$ or get back to the "first toss is T" setting). Solving for h and t gives $t = 3/5$, $h = 1/5$, and $p = 2/5$.]

(c) How does the answer in (b) change if, instead of the pattern HH, the winning pattern is HT? [Now, using the same notations, $h = ((1/2) + h/2)/2$ or $h = 1/3$, and $t = (1/2)(h+t)/2 + (1/2)((1/2) + h/2)$ or $t = 5/9$, so that $p = (h+t)/2 = 4/9$.]

Homework 11.

1. The joint density of X and Y is

$$f(x,y)=\frac{e^{-x/y}e^{-y}}{y},\ 0<x<\infty,\ 0<y<\infty.$$

Compute $E(X^2|Y=y)$.

2. A prisoner is trapped in a cell containing three doors. The first door leads to a tunnel that would lead back to the cell after two days of travel. The second door leads to a tunnel that would lead back to the cell after 4 days of travel. The third door leads to a tunnel that would lead to freedom after one day of travel. The prisoner cannot label doors and will always select the doors randomly with probabilities $0.5, 0.3$ and 0.2, respectively. Compute the expected number of days before the prisoner reaches freedom.

3. The number of people entering the elevator on the ground floor is a Poisson random variable with mean 10. There are N floors above the ground floor, and everybody in the elevator is equally likely to exit on any of the N floors, independently of everybody else. Nobody enters the elevator above the ground floor. Compute the expected number of stops the elevator makes before everybody is out.

4. The expected number of accidents at an industrial facility is 5 per week. The average number of injured people in each accident is 2.5. Assuming all the independence you need, compute the expected number of injured workers per week.

5. The life time of the light bulb is characterized by the mean value μ and standard deviation σ. There are two types of light bulbs in a box, with the corresponding parameters μ_1, σ_1 and μ_2, σ_2. The proportion of type-1 bulbs in the box is p. A bulb is selected at random. Denote by X the life time of the this bulb. Compute the expected value and the variance of X.

6. The number of accident a person has in a year is a Poisson random variable with parameter λ. For 60% of the population, $\lambda = 2$; for the rest, $\lambda = 3$. Compute the probability that, in one year, a randomly selected person will have (a) no accidents; (b) three accidents; (c) three accidents given that there were no accidents the previous year. (The answers for (b) and (c) are different because this is not a standard Poisson process and so the numbers of accidents from year to year are not independent, although they become independent if you fix λ).

7. [One of Polya's urn models] An urn contains b black and w white balls. At each step, a ball is removed from the urn at random and then put back together with one more ball of the same color. Compute the probability p_n to get a black ball on step n, $n \ge 1$.

Homework 12.

1. Given independent identically distributed Poisson random variables $X_1, \dots, X_{20}$ with parameter 1, compute an approximation of $P(X_1 + \ldots + X_{20} > 15)$ using as many ways as possible and compare the results with the "exact" value (using your favorite software package, such as MatLab, or even your pocket calculator, if it can handle it).

Some options for computing an approximation: (a) Markov's inequality; (b) the central limit theorem for Poisson, with or without continuity correction; (c) central limit theorem for sum of exponentials or Poissons; (d) improved Chebyshev inequality; (e) Exponential Chebychev (Chernoff).

2. A fair die is rolled until the total sum exceeds 300. Compute approximately the probability that at least 80 rolls will be necessary.

3. Assume that the amount of weight, in units of 1000 pounds, a bridge can hold without collapsing, is normally distributed with mean 400 and standard deviation 40. Assume that the weight of a car, in the same units, is a random variable with mean 3 and standard deviation 0.3. How many cars should there be on the bridge for the probability of collapse to exceed 0.1?

4. The daily price Y_n of a certain stock is modeled by the relation

$$Y_n = Y_{n-1} + X_n, \ n \geq 1,$$

where X_k are independent identically distributed standard normal random variables. Suppose that the current stock price is \$100. Compute the probability that the price will exceed \$105 after 10 days.

5. Consider 100 independent tosses of a fair coin. Let p_1 be the probability to get exactly 50 heads, p_2, the probability to get at least 60 heads, and p_3, the probability to get at least 55 heads. Order the numbers p_k from smallest to largest using your intuition/common sense, and then verify your conclusion using the Central Limit Theorem. Will the ordering change if all the numbers are increased by a factor of 10, so that you toss the coin 1000 times and look at exactly 500 heads, at least 600 heads, and at least 550 heads?

6. Consider a lottery with n^2, tickets, of which only n tickets win prizes. Let p_n be the probability that, out of n randomly selected tickets, at least one wins a prize. Compute $\lim_{n\to\infty} p_n$. [the limit is $1 - e^{-1} \approx 0.632$]

7. Customers arrive at a bank according to a Poisson process. Two customers arrived during one hour. Compute the probability that (a) both arrived during the first 20 min; (b) at least one arrived during the first 20 min. Note that the rate λ of the process is not necessary to solve the problem.

8. Cars cross a certain point in the highway following a Poisson process, with 3 cars per minute on average. A dog runs straight across the road and will get injured on the encounter of two or more cars. Compute the probability that the dog crosses the highway unhurt if it takes the dog s seconds to cross the road.