Exercises in Probability Theory

Nikolai Chernov

All exercises (except Chapters 16 and 17) are taken from two books:
R. Durrett, *The Essentials of Probability*, Duxbury Press, 1994
S. Ghahramani, *Fundamentals of Probability*, Prentice Hall, 2000

1 Combinatorics

These problems are due on August 24

Exercise 1.1. In how many ways can we draw five cards from an ordinary deck of 52 cards (a) with replacement; (b) without replacement?

(a): 52^5 (b): $P_{52,5}$

Exercise 1.2. Suppose in a state, licence plates have three letters followed by three numbers, in a way that no letter or number is repeated in a single plate. Determine the number of possible licence plates for this state.

$26 \times 25 \times 24 \times 10 \times 9 \times 8 = 11,232,000$

Exercise 1.3. Eight people are divided into four pairs to play bridge. In how many ways can this be done?

$C_{8,2} \cdot C_{6,2} \cdot C_{4,2} \cdot C_{2,2} = 2,520.$

Exercise 1.4. A domino is an ordered pair (m, n) with $0 \le m \le n \le 6$. How many dominoes are in a set if there is only one of each?

$C_{7,2} + 7 = 28$

Exercise 1.5. A club with 50 members is going to form two committees, one with 8 members and the other with 7. How many ways can this be done if the committees must be disjoint?

$C_{50,8} \cdot C_{42,7} = 1.45 \times 10^{16}$

Exercise 1.6. Six students, three boys and three girls, lineup in a random order for a photograph. What is the probability that the boys and girls alternate?

$72/720 = 1/10$

Exercise 1.7. In a town of 50 people, one person tells a rumor to a second person, who tells a third, and so on. If at each step the recipient of the rumor is chosen at random, what is the probability the rumor will be told 8 times without being told to someone who knows it?

$P_{49,8}/49^8 = 0.547$

Exercise 1.8. A fair coin is tossed 10 times. What is the probability of (a) five Heads; (b) at least five Heads?

(a): $C_{10,5}/2^{10} = 0.246$ (b): $\sum_{k=5}^{10} C_{10,k}/2^{10} = 0.623$

Exercise 1.9. Suppose we roll a red die and a green die. What is the probability the number on the red die is larger ($>$) than the number on the green die?

$15/36$

Exercise 1.10. Four people are chosen randomly from 5 couples. What is the probability that two men and two women are selected?

$C_{5,2}^2/C_{10,4} = 10/21$

Exercise 1.11 (Bonus)**.** If five numbers are selected at random from the set $\{1, 2, \ldots, 20\}$, what is the probability that their minimum is larger than 5?

$C_{15,5}/C_{20,5} = 0.19$

Exercise 1.12 (Bonus)**.** The World Series is won by the first team to win four games. Suppose both teams are equally likely to win each game. What is the probability that the team that wins the first game will win the series? $((C_{6,3} + C_{6,4} + C_{6,5} + C_{6,6})/2^6)$

Exercise 1.13 (Bonus)**.** Show that the probability of an even number of Heads in n tosses of a fair coin is always $1/2$.

Exercise 1.14 (Bonus)**.** If n balls are randomly placed into n cells (so that more than one ball can be placed in a cell), what is the probability that each cell will be occupied?

$n!/n^n$

Exercise 1.15 (Graduate)**.** Read about Stirling's formula. Use it to approximate the probability of observing exactly n Heads in a game where a fair coin is flipped $2n$ times. Use that approximation to compute the probability for $n = 50$.

2 Probability Space

These problems are due on August 29

Exercise 2.1. Suppose we pick a letter at random from the word TENNESSEE. What is the sample space Ω and what probabilities should be assigned to the outcomes?

$\{T, E, N, S\}$, with probabilities 1/9. 4/9, 2/9, 2/9.

Exercise 2.2. In a group of students, 25% smoke cigarettes, 60% drink alcohol, and 15% do both. What fraction of students have at least one of these bad habits?

70%

Exercise 2.3. Suppose $P(A) = \frac{1}{3}$, $P(A^c \cap B^c) = \frac{1}{2}$, and $P(A \cap B) = \frac{1}{4}$. What is $P(B)$?

5/12

Exercise 2.4. Given two events A and B with $P(A) = 0.4$ and $P(B) = 0.7$. What are the maximum and minimum possible values for $P(A \cap B)$?

min: 0.1, max: 0.4

Exercise 2.5. John and Bob take turns in flipping a fair coin. The first one to get a Heads wins. John starts the game. What is the probability that he wins?

$1/2 + 1/8 + 1/32 + \cdots = 2/3$

Exercise 2.6. Use the inclusion-exclusion formula to compute the probability that a randomly chosen number between 0000 and 9999 contains at least one 1.

0.3439

Exercise 2.7 (Graduate)**.** Let A and B be two events. Show that if $P(A) = 1$ and $P(B) = 1$, then $P(A \cap B) = 1$. (Hint: use their complements A^c and B^c.)

Exercise 2.8 (Graduate)**.** Read about the continuity of probabilities. Suppose a number x is picked randomly from the interval $[0, 1]$. What is the probability that $x = 1/10$? What is the probability that $x = m/10$ for some $m = 0, \ldots, 10$? What is the probability that $x = m/100$ for some $m = 0, \ldots, 100$? What is the probability that $x = m/n$ for some $m \leq n$? What is the probability that x is rational?

3 Conditional Probability and Independence

These problems are due on September 7

Exercise 3.1. A friend flips two coins and tells you that at least one is Heads. Given this information, what is the probability that the first coin is Heads?

$2/3$

Exercise 3.2. Suppose that a married man votes is 0.45, the probability that a married woman votes is 0.4, and the probability a woman votes given that her husband votes is 0.6. What is the probability that (a) both vote, (b) a man votes given that his wife votes?

(a) 0.27, (b) 0.675

Exercise 3.3. If 5% of men and 0.25% of women are color blind, what is the probability that a randomly selected person is color blind?

2.625%

Exercise 3.4 (Bonus)**.** How can 5 black and 5 white balls be put into two urns to maximize the probability a white ball is drawn when we draw a ball from a randomly chosen urn?

one urn must have a single white ball

Exercise 3.5. Roll two dice. Let $A =$"The first die is odd", $B =$"The second die is odd", and $C =$"The sum is odd". Show that these events are pairwise independent but not jointly independent.

$P(A \cap B \cap C) = 0$.

Exercise 3.6. Three couples that were invited to dinner will independently show up with probabilities 0.9, 0.8, and 0.75. Let N be the number of couples that show up. Calculate the probability that $N = 3$ and that of $N = 2$.

0.54 and 0.375

Exercise 3.7 (Bonus)**.** Show that if an event A is independent of itself, then either $P(A) = 0$ or $P(A) = 1$.

Exercise 3.8. How many times should a coin be tosses so that the probability of at least one head is at least 99%?

7 times

Exercise 3.9. On a multiple-choice exam with four choices for each question, a student either knows the answer to a question or marks it at random. Suppose the student knows answers to 60% of the exam questions. If she marks the answer to question 1 correctly, what is the probability that she knows the answer to that question?

6/7

Exercise 3.10 (Graduate)**.** Suppose $A_1, \ldots, A_n$ are independent. Show that

$$P(A_1 \cup \cdots \cup A_n) = 1 - \prod_{m=1}^{n} \big(1 - P(A_m)\big).$$

Exercise 3.11. In a certain city 30% of the people are Conservatives, 50% are Liberals, and 20% are Independents. In a given election, 2/3 of the Conservatives voted, 80% of the Liberals voted, and 50% of the Independents voted. If we pick a voter at random, what is the probability he/she is Liberal?

4/7

Exercise 3.12 (Graduate)**.** Show that if A and B are independent, then A and B^c are independent, A^c and B are independent, and A^c and B^c are independent.

Exercise 3.13 (Graduate)**.** Show that an event A is independent of *every* event B if $P(A) = 0$ or $P(A) = 1$.

4 Discrete Random Variables

These problems are due on September 14

Exercise 4.1. Suppose we roll two dice and let X be the minimum of the two numbers obtained. Determine the probability function of X and sketch its graph.

$P(X = 1) = 11/36,\ P(X = 2) = 9/36,\ P(X = 3) = 7/36,$
$P(X = 4) = 5/36,\ P(X = 5) = 3/36,\ P(X = 6) = 1/36$

Exercise 4.2. In successive rolls of a die, let X be the number of rolls until the first 6 appears. Determine the probability function of X.

$P(X = k) = 5^{k-1}/6^k$ for $k = 1, 2, \ldots$

Exercise 4.3. From the interval $(0, 1)$, five points are selected at random. What is the probability that at least two of them are less than $1/3$?

$\sum_{k=2}^{5} \binom{5}{k}(1/3)^k(2/3)^{5-k} = 131/243 \approx 0.539$

Exercise 4.4. A certain rare blood type can be found in only 0.05% of people. Use the Poisson approximation to compute the probability that at most two persons in a group of randomly selected 3000 people will have this rare blood type.

$\lambda = 1.5$, probability$=\sum_{k=0}^{2} \frac{1.5^k}{k!} e^{-1.5} = 3.625e^{-1.5} \approx 0.8088$

Exercise 4.5. An airline company sells 200 tickets for a plane with 198 seats, knowing that the probability a passenger will not show up for the flight is 0.01. Use the Poisson approximation to compute the probability they will have enough seats for all the passengers who show up.

$\lambda = 2$, probability$=1 - \sum_{k=0}^{1} \frac{2^k}{k!} e^{-2} = 1 - 3e^{-2} = 0.5946$

Exercise 4.6. Let X be a Poisson random variable with parameter $\lambda > 0$. Denote $P_k = \mathbb{P}(X = k)$ for $k = 0, 1, \ldots$. Compute P_{k-1}/P_k and show that this ratio is less than one if and only if $k < \lambda$. This shows that the most probable values of X are those near λ. In fact, the most probable value is the greatest integer less than or equal to λ.

Exercise 4.7 (Bonus)**.** Let X be a binomial random variable, $b(n,p)$. Denote $P_k = \mathbb{P}(X = k)$ for $k = 0, 1, \ldots, n$. Compute P_{k-1}/P_k and show that this ratio is less than one if and only if $k < np + p$. This shows that the most probable values of X are those near np.

Exercise 4.8 (Graduate)**.** Let X be a geometric random variable with parameter p. Denote $P_k = \mathbb{P}(X = k)$ for $k = 1, 2, \ldots$. Prove that

$$\mathbb{P}(X \geq m + n / X > m) = \mathbb{P}(X \geq n) \tag{1}$$

for any positive integers m and n. Explain what it means, in terms of trials till the first success. The property (1) is called the *memoryless property* of a discrete random variable (that takes values $1, 2, \ldots$).

5 Continuous Random Variables

These problems are due on September 21

Exercise 5.1. Let $F(x) = e^{-1/x}$ for $x > 0$ and $F(x) = 0$ for $x \leq 0$. Is $F(x)$ a distribution function? If so, find its density function.

Yes. The density is $f(x) = x^{-2}e^{-1/x}$ for $x > 0$.

Exercise 5.2. Suppose X has density function $f(x) = c(1-x^2)$ for $-1 < x < 1$ and $f(x) = 0$ elsewhere. Compute the value of c. Find the distribution function $F(x)$ of X. Sketch the graphs of $f(x)$ and $F(x)$. Compute the probabilities $\mathbb{P}(X > 0.5)$ and $\mathbb{P}(0 < X < 0.5)$.

$c = 3/4$. $F(x) = \frac{1}{2} + \frac{3}{4}x - \frac{1}{4}x^3$. Probabilities are 5/32 and 11/32.

Exercise 5.3 (Bonus)**.** Consider $f(x) = cx^{-1/2}$ for $x \geq 1$, and $f(x) = 0$ otherwise. Show that there is no value of c that makes f a density function.

Exercise 5.4. A point is selected at random from the interval $(0, 1)$; it then divides this interval into two segments. What is the probability that the longer segment is at least twice as long as the shorter segment?

2/3.

Exercise 5.5 (Graduate)**.** Read about distribution functions of discrete random variables. Now let X_n be a discrete uniform random variable that takes values $1, \ldots, n$. Consider the random variable $Y_n = \frac{1}{n} X_n$. Describe the distribution function $F_n(x)$ of the variable Y_n; sketch it for $n = 3$ and $n = 10$. Find the limit function, $F(x) = \lim_{n\to\infty} F_n(x)$. Is it a distribution function? For what random variable? If the limit function $F(x)$ is a distribution function of a random variable Y, then we say that Y_n converges to Y *in distribution.*

6 Exponential Random Variables

These problems are due on September 28

Exercise 6.1. Suppose the lifetime of a TV set is an exponential random variable X with half-life $t_{1/2} = 5$ (years). Find the parameter λ. Compute the probability $\mathbb{P}(X > 10)$ and the conditional probability $\mathbb{P}(X > 24/X > 14)$.

Exercise 6.2. Let X be an exponential random variable with parameter $\lambda > 0$. Find probabilities

$$\mathbb{P}\left(X > \frac{2}{\lambda}\right) \qquad \text{and} \qquad \mathbb{P}\left(\left|X - \frac{1}{\lambda}\right| < \frac{2}{\lambda}\right).$$

Exercise 6.3. Show that if X is exponential(λ), then the random variable $Y = \lambda X$ is exponential(1).

Exercise 6.4 (Graduate)**.** Let X_n be a geometric random variable with parameter $p = \lambda/n$. Compute

$$\mathbb{P}\left(\frac{X_n}{n} > x\right) \qquad \text{for} \quad x > 0$$

and show that as $n \to \infty$ this probability converges to $\mathbb{P}(Y > x)$, where Y is an exponential random variable with parameter λ. This shows that X_n/n is approximately an exponential random variable.

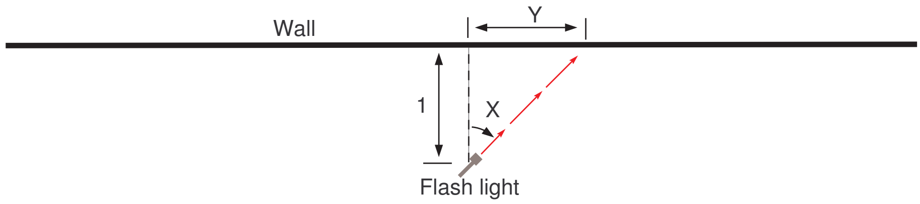

Figure 1: A drunk waving a flash light randomly at a wall.

7 Functions of Random Variables

These problems are due on October 3

Exercise 7.1. Suppose X is a uniform random variable on the interval $(0, 1)$. Find the distribution and density functions of $Y = X^n$ for $n \geq 2$.

Exercise 7.2. Suppose X is an exponential random variable with parameter $\lambda = 1$. Find the distribution and density functions of $Y = \ln(X)$. This is called the double exponential distribution.

Exercise 7.3 (Graduate)**.** A drunk standing one foot from a wall shines a flashlight at a random angle that is uniformly distributed between $-\pi/2$ and $\pi/2$, i.e. the angle X is $U(-\pi/2, \pi/2)$. Find the distribution and density functions of the place where the light hits the wall. Note: that place is given by the formula $Y = \tan(X)$, see illustration. The variable Y is said to have Cauchy distribution.

Exercise 7.4 (Bonus)**.** Let X be a uniform random variable on the interval $(-1, 1)$. Find the distribution and density functions of $Y = |X|$. Is the variable Y familiar? What is its type?

8 Normal Random Variables

These problems are due on October 7

Exercise 8.1. Suppose Z has a standard normal distribution. Use the table to compute
(a) $\mathbb{P}(-1.1 < Z < 2.95)$, 0.8627
(b) $\mathbb{P}(Z > -0.69)$, 0.7549
(c) $\mathbb{P}(-1.45 < Z < -0.28)$, 0.3162
(d) $\mathbb{P}(|Z| \le 2.5)$. 0.9876

Exercise 8.2. Suppose X has normal distribution $\mathcal{N}(-2, 9)$. Compute
(a) $\mathbb{P}(-3.2 < X < 1.3)$, 0.5197
(b) $\mathbb{P}(X > -1.19)$, 0.3936
(c) $\mathbb{P}(|X| \le 11)$. 0.9987

Exercise 8.3. Suppose the weight (in pounds) of a randomly selected woman has a normal distribution $\mathcal{N}(130, 400)$. Of the women who weigh above 140 pounds, what percent weigh over 170 pounds?

0.0739

Exercise 8.4 (Graduate)**.** The following inequality is obviously true for all $y > 0$:

$$(1 - 3y^{-4})\, e^{-y^2/2} < e^{-y^2/2} < (1 + y^{-2})\, e^{-y^2/2}.$$

Integrate it from x to ∞ and derive the following:

$$\frac{1}{x\sqrt{2\pi}}\left(1 - \frac{1}{x^2}\right) e^{-x^2/2} < 1 - \Phi(x) < \frac{1}{x\sqrt{2\pi}}\, e^{-x^2/2}.$$

Hint: use integration by parts to simplify the integrals

$$\int_x^\infty \frac{1}{y^2}\, e^{-y^2/2}\, dy \qquad \text{and} \qquad \int_x^\infty \frac{3}{y^4}\, e^{-y^2/2}\, dy.$$

Lastly, show that

$$1 - \Phi(x) \sim \frac{1}{x\sqrt{2\pi}}\, e^{-x^2/2}.$$

That is, as $x \to \infty$, the ratio of the two sides approaches 1.

Exercise 8.5 (Graduate)**.** Let Z be a standard normal random variable. Use the last formula from the previous exercise to show that for every $x > 0$

$$\lim_{t\to\infty} \mathbb{P}\Big(Z > t + \frac{x}{t} \,/\, Z > t\Big) = e^{-x}.$$

(Note: this is a conditional probability!)

9 Joint Distributions

These problems are due on October 17

Exercise 9.1. Suppose $\mathbb{P}(X = x, Y = y) = c(x + y)$ for $x, y = 0, 1, 2, 3$.
(a) What value of c will make this a probability function? $c = 1/48$
(b) What is $\mathbb{P}(X > Y)$? $3/8$
(c) What is $\mathbb{P}(X + Y = 3)$? $1/4$

Exercise 9.2. Suppose X and Y have joint density $f(x, y) = 2$ for $0 < y < x < 1$. Find $\mathbb{P}(X - Y > z)$, where $z > 0$ is a constant.

$(1 - z)^2$

Exercise 9.3. (Bonus) Two people agree to meet for a drink after work but they are impatient and each will only wait 15 minutes for the other person to show up. Suppose that they each arrive at independent random times uniformly distributed between 5 p.m. and 6 p.m. What is the probability they will meet?

$7/16 = 0.4375$

Exercise 9.4. (Bonus) Suppose n points are selected at random and independently inside a circle of radius R (each point is uniformly distributed inside the circle). Find the probability that the distance of the nearest point to the center is less than r, where $r < R$ is a constant.

$1 - [1 - (r^2/R^2)^n]$

Exercise 9.5. (Bonus) Suppose X and Y are uniform on $(0, 1)$ and independent. Find the density function of $X + Y$. (You will understand why the $X + Y$ is said to have a 'triangular' distribution.)

$f(x) = x$ for $0 < x < 1$ and $f(x) = 1 - x$ for $1 < x < 2$

Exercise 9.6. (Bonus) Let X_1, X_2, X_3, and X_4 be four independently selected random numbers from $(0, 1)$. Find $\mathbb{P}\big(\frac{1}{4} < X_{(3)} < \frac{1}{2}\big)$.

$k = 2$, $F(x) = 4x^3 - 3x^4$, probability is $67/256$

Exercise 9.7 (Graduate)**.** Let $X_1, \ldots, X_n$ be independently selected random numbers from $(0,1)$, and $Y_n = nX_{(1)}$. Prove that

$$\lim_{n\to\infty} \mathbb{P}\big(Y_n > x\big) = e^{-x},$$

thus Y_n is asymptotically exponential(1).

Exercise 9.8 (Graduate)**.** Suppose X_1 and X_2 are independent standard normal random variables. Find the distribution of $Y = (X_1^2 + X_2^2)^{1/2}$. This is the **Rayleigh distribution**.

Exercise 9.9 (Graduate)**.** Read about transformations of pairs of random variables. Now suppose X_1 and X_2 are independent uniform$(0,1)$ random variables. Consider two new random variables defined by

$$Z_1 = \sqrt{-2\ln X_1}\,\cos(2\pi X_2) \quad \text{and} \quad Z_2 = \sqrt{-2\ln X_1}\,\sin(2\pi X_2).$$

Show that Z_1 and Z_2 are two independent standard normal random variables, i.e. their joint density function is

$$f_{Z_1,Z_2}(z_1,z_2) = \frac{1}{2\pi}\,e^{-(z_1^2+z_2^2)/2}.$$

This gives a practical algorithm for simulating values of normal random variables.

Exercise 9.10 (Graduate)**.** Read about independent random variables. Determine if two random variables X and Y are independent given their joint density function $f_{X,Y}(x,y)$ below. If they are independent, determine the densities $f_X(x)$ and $f_Y(y)$.
(a) $f_{X,Y}(x,y) = \frac{1}{18}\,x^2y$ for $0 < x < 3$ and $0 < y < 2$ and 0 elsewhere.
(b) $f(x,y) = 8xy$ for $0 < x < y < 1$.

10 Mean Value

These problems are due on October 21

Exercise 10.1. A man plays roulette and bets \$1 on black 19 times. He wins \$1 with probability 18/38 and loses \$1 with probability 20/38. What is his expected winnings?

Expected winning is -1, i.e., he expects to lose \$1.

Exercise 10.2. A player bets \$1 and if he wins he gets \$3. What must the probability of winning be for this to be a fair bet?

A fair bet is when the average net result is zero (no win or lose). The probability must be 1/4, as then the average win is $3 \cdot (1/4) - 1 \cdot (3/4) = 0$.

The problem admits a different interpretation so that the net win is $3-1 = 2$. Then the probability must be 2/3.

Exercise 10.3. Suppose X has density function $f(x) = 3x(2-x)/4$ for $0 < x < 2$, and $f(x) = 0$ elsewhere. Find $\mathbb{E}(X)$. Graph the density function, note its symmetry about the mean value.

Answer: $\int_0^2 \frac{3x^2(2-x)}{4}\, dx = 1$.

Exercise 10.4 (Bonus)**.** Two people agree to meet for a drink after work and each arrives independently at a time uniformly distributed between 5 p.m. and 6 p.m. What is the expected amount of time that the first person to arrive has to wait for the arrival of the second?

Let T denote the waiting time. Then its distribution function $F_T(x) = \mathbb{P}(T < x) = 1 - (1-x)^2$ (this is similar to Problem 9.3). Then the density function is $f_T(x) = 2(1-x)$ and the mean waiting time is $\int_0^1 2x(1-x)\, dx = 1 - \frac{2}{3} = \frac{1}{3}$ (or just 20 minutes).

Exercise 10.5. Suppose X is exponential(3), i.e. X is an exponential random variable with parameter $\lambda = 3$. Calculate $\mathbb{E}\big(e^X\big)$.

Answer: $\int_0^\infty e^x 3e^{-3x}\, dx = \int_0^\infty 3e^{-2x}\, dx = \frac{3}{2}$.

Exercise 10.6. Let $X_1, \ldots, X_n$ be independent and uniform random variables on $(0,1)$. Denote $V = \max\{X_1, \ldots, X_n\}$ and $W = \min\{X_1, \ldots, X_n\}$, as in the class notes. Find $\mathbb{E}(V)$ and $\mathbb{E}(W)$.

Answer: $\mathbb{E}(V) = \int_0^1 xnx^{n-1}\, dx = \frac{n}{n+1}$ and $\mathbb{E}(W) = \int_0^1 xn(1-x)^{n-1}\, dx = \frac{1}{n+1}$. The second integral can be taken by a change of variable $y = 1 - x$.

Exercise 10.7 (Graduate)**.** Show that $p(n) = \frac{1}{n(n+1)}$ for $n = 1, 2, \ldots$ is a probability function for some discrete random variable X. Find $\mathbb{E}(X)$.

Exercise 10.8 (Bonus)**.** Suppose X has the standard normal distribution. Compute $\mathbb{E}(|X|)$.

Exercise 10.9 (Graduate)**.** Suppose $X_1, \ldots, X_n > 0$ are independent and all have the same distribution. Let $m < n$. Find

$$\mathbb{E}\Big(\frac{X_1 + \cdots + X_m}{X_1 + \cdots + X_n}\Big).$$

Hint: since $X_1, \ldots, X_n$ are independent and have the same distribution, they are interchangeable, i.e., you can swap X_i and X_j for any $i \neq j$ without changing the above mean value.

Exercise 10.10 (Graduate)**.** Suppose X has the standard normal distribution. Use integration by parts to show that $\mathbb{E}(X^k) = (k-1)\,\mathbb{E}(X^{k-2})$. Derive that $\mathbb{E}(X^k) = 0$ for all odd $k \geq 1$. Compute $\mathbb{E}(X^4)$ and $\mathbb{E}(X^6)$. Derive a general formula for $\mathbb{E}(X^{2k})$.

Exercise 10.11 (Graduate)**.** Read about independent random variables. Show that if X and Y take only two values 0 and 1 and $\mathbb{E}(XY) = \mathbb{E}(X)\,\mathbb{E}(Y)$, then X and Y are independent.

11 Variance

These problems are due on October 26

Exercise 11.1. Suppose X has density function $f(x) = (r-1)\,x^{-r}$ for $x > 1$ and $f(x) = 0$ elsewhere; here $r > 1$ is a parameter. Find $\mathbb{E}(X)$ and $\mathsf{Var}(X)$. For which values of r do $\mathbb{E}(X)$ and $\mathsf{Var}(X)$ exist?

Answer: $\mathbb{E}(X) = \frac{r-1}{r-2}$ for $r > 2$ and $\mathbb{E}(X^2) = \frac{r-1}{r-3}$ for $r > 3$. Hence $\mathsf{Var}(X) = \frac{r-1}{r-3} - \frac{(r-1)^2}{(r-2)^2}$ for $r > 3$.

Exercise 11.2. Can we have a random variable X with $\mathbb{E}(X) = 3$ and $\mathbb{E}(X^2) = 8$?

Answer: No, because $\mathsf{Var}(X) = 8 - 3^3 = -1$ cannot be negative.

Exercise 11.3. Let X and Y be independent with $\mathsf{Var}(X) = 5$ and $\mathsf{Var}(Y) = 2$. Find $\mathsf{Var}(2X - 3Y + 11)$.

Answer: 38.

Exercise 11.4. Suppose $X_1, \ldots, X_n$ are independent with $\mathbb{E}(X_i) = \mu$ and $\mathsf{Var}(X_i) = \sigma^2$ for all $i = 1, \ldots, n$, Find $\mathbb{E}\big[(X_1 + \ldots + X_n)^2\big]$.

Answer: Recall that for any random variable Y we have $\mathbb{E}(Y^2) = \mathsf{Var}(Y) + [\mathbb{E}(Y)]^2$. Now $\mathbb{E}\big[(X_1 + \ldots + X_n)^2\big] = \mathsf{Var}(X_1 + \ldots + X_n) + \big[\mathbb{E}(X_1 + \ldots + X_n)\big]^2 = (\sigma^2 + \cdots + \sigma^2) + [\mu + \cdots + \mu]^2 = n\sigma^2 + [n\mu]^2$.

Exercise 11.5 (Bonus)**.** Suppose X takes values in the interval $[-1, 1]$. What is the largest possible value of $\mathsf{Var}(X)$ and when is this attained? What is the answer when $[-1, 1]$ is replaced by $[a, b]$?

Answer: Largest variance is 1; it is attained when X takes values ± 1, each with probability 0.5.

Exercise 11.6 (Graduate)**.** Show that $\mathbb{E}(X - c)^2$ is minimized by taking $c = \mathbb{E}(X)$.

Exercise 11.7 (Graduate)**.** Suppose you want to collect a set of 5 (distinct) baseball cards. Assume that you buy one card at a time, and each time you get a randomly chosen card (from the 5 different cards available). Let X be the number of cards you have to buy before you collect all the 5. Describe X as a sum of geometric random variables. Find $\mathbb{E}(X)$ and $\mathsf{Var}(X)$.

12 Moment Generating Function

These problems are due on November 4

Exercise 12.1. Suppose X has a uniform distribution on the interval $(0, 1)$. Find the moment-generating function of X and use it to compute the moments of X.

Answer: $\mathbb{M}_X(t) = \frac{e^t - 1}{t}$. To compute moments, take derivatives and use l'Hopital's rule to find the limit as $t \to 0$. For example, $\mathbb{M}'(t) = \frac{te^t - e^t + 1}{t^2}$, and $\lim_{t \to 0} \mathbb{M}'(t) = 1/2$ by l'Hopital's rule.

Exercise 12.2. Is $\mathbb{M}(t) = \frac{e^t+e^{-t}}{6} + \frac{2}{3}$ a moment-generating function of a random variable? If yes, find the corresponding probability function. Also find the moments of that random variable.

Answer: yes. The probability function is $\mathbb{P}(X=0) = 2/3$, $\mathbb{P}(X=1) = 1/6$, and $\mathbb{P}(X=-1) = 1/6$. The moments are $\mathbb{E}(X^k) = 0$ for odd k's and $\mathbb{E}(X^k) = 1/3$ for even k's.

Exercise 12.3. Let X be $\mathcal{N}(1,2)$ and Y be $\mathcal{N}(4,7)$; assume that X and Y are independent. Find the probabilities of the following events: (a) $\mathbb{P}(X+Y>0)$, (b) $\mathbb{P}(X<Y)$, and (c) $\mathbb{P}(3X>20-4Y)$.

Answer: (a) $X+Y$ is $\mathcal{N}(5,9)$ and $\mathbb{P}(X+Y>0) = 0.9525$; (b) $X-Y$ is $\mathcal{N}(-3,9)$ and $\mathbb{P}(X-Y<0) = 0.8413$; (c) $3X+4Y$ is $\mathcal{N}(19,130)$ and $\mathbb{P}(3X>20-4Y) = 0.4641$.

Exercise 12.4 (Bonus)**.** Suppose that for a random variable X we have $\mathbb{E}(X^n) = 2^n$ for all $n = 1, 2, \ldots$. Calculate the moment-generating function and the probability function of X.

Answer: $\mathbb{M}_X(t) = e^{2t}$. The probability function is $\mathbb{P}(X=2) = 1$.

13 Covariance and Correlation

These problems are due on November 9

Exercise 13.1. Suppose X is a uniform random variable on the interval $(-1,1)$ and $Y = X^2$. Find $\mathsf{Cov}(X, X^2)$. Are X and Y independent?

Answer: $\mathsf{Cov}(X, X^2) = \mathbb{E}(X^3) - \mathbb{E}(X) \cdot \mathbb{E}(X^2) = 0 - 0 \times \frac{1}{3} = 0$, but these random variables are dependent.

Exercise 13.2. Show that for any random variables X and Y we have

$$\mathsf{Cov}(X+Y, X-Y) = \mathsf{Var}(X) - \mathsf{Var}(Y).$$

Notice that we are not assuming that X and Y are independent.

Solution: $\mathsf{Cov}(X+Y, X-Y) = \mathsf{Cov}(X,X) - \mathsf{Cov}(X,Y) + \mathsf{Cov}(Y,X) - \mathsf{Cov}(Y,Y) = \mathsf{Var}(X) - \mathsf{Var}(Y)$.

Exercise 13.3. Suppose that random variables X and Y are independent. Show that

$$\rho(X+Y, X-Y) = \frac{\mathsf{Var}(X) - \mathsf{Var}(Y)}{\mathsf{Var}(X) + \mathsf{Var}(Y)}.$$

Solution:

$$\begin{aligned}\rho(X+Y, X-Y) &= \frac{\mathsf{Cov}(X+Y, X-Y)}{\sqrt{\mathsf{Var}(X+Y)}\sqrt{\mathsf{Var}(X-Y)}} \\ &= \frac{\mathsf{Var}(X) - \mathsf{Var}(Y)}{\sqrt{\mathsf{Var}(X) + \mathsf{Var}(Y)}\sqrt{\mathsf{Var}(X) + \mathsf{Var}(Y)}}\end{aligned}$$

Exercise 13.4. In n independent Bernoulli trials, each with probability of success p, let X be the number of successes and Y the number of failures. Calculate $\mathbb{E}(XY)$ and $\mathsf{Cov}(X,Y)$.

Answer:

$$\begin{aligned}\mathbb{E}(XY) &= \mathbb{E}(X(n-X)) = n\mathbb{E}(X) - \mathbb{E}(X^2) = n\mathbb{E}(X) - \left[\mathsf{Var}(X) + [\mathbb{E}(X)]^2\right] \\ &= n^2p - \left[np(1-p) + (np)^2\right] = (n^2-n)(p-p^2)\end{aligned}$$

and

$$\begin{aligned}\mathsf{Cov}(X,Y) &= \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = (n^2-n)(p-p^2) - npn(1-p) \\ &= n(p^2-p) = -np(1-p)\end{aligned}$$

Exercise 13.5 (Bonus)**.** Suppose that X_1, X_2, X_3 are independent, have mean 0 and $\mathsf{Var}(X_i) = i$. Find $\rho(X_1 - X_2, X_2 + X_3)$.

Answer: $-2/\sqrt{15}$.

Exercise 13.6 (Bonus)**.** Let X and Y be random variables with $\mathbb{E}(X) = 2$, $\mathsf{Var}(X) = 1$, $\mathbb{E}(Y) = 3$, $\mathsf{Var}(Y) = 4$. What are the smallest and largest possible values of $\mathsf{Var}(X+Y)$?

Answer: the smallest is 1, the largest is 9. The key element is that $-1 \leq \rho_{X,Y} \leq 1$.

Exercise 13.7 (Graduate)**.** Suppose $X_1, \ldots, X_n$ are independent random variables with $\mathbb{E}(X_i) = \mu$ and $\mathsf{Var}(X_i) = \sigma^2$ for all $i = 1, \ldots, n$. Let $S_k = X_1 + \cdots + X_k$. Find $\rho(S_k, S_n)$.

14 Law of Large Numbers

These problems are due on November 11

Exercise 14.1. The average IQ score on a certain campus is 110. If the variance of these scores is 15, what can be said about the percentage of students with an IQ above 140?

Answer: $1/60$, or 1.67%.

Exercise 14.2 (Graduate)**.** Let $y \geq \sigma > 0$. Give an example of a random variable Y with $\mathbb{E}(Y) = 0$, $\mathsf{Var}(Y) = \sigma^2$, and $\mathbb{P}\big(|Y| \geq y\big) = \sigma^2/y^2$, the upper bound in Chebyshev's inequality.

Exercise 14.3 (Graduate)**.** Prove Bernstein's inequality: for any $t > 0$

$$\mathbb{P}(X > y) \leq e^{-ty}\,\mathbb{E}\big(e^{tX}\big).$$

Exercise 14.4 (Graduate)**.** Suppose $X_1, \dots, X_n$ are independent exponential(1) random variables, and let $S_n = X_1 + \cdots + X_n$. Apply Bernstein's inequality to estimate $\mathbb{P}(S_n > cn)$ with $c > 1$, then pick t to minimize the upper bound to show

$$\mathbb{P}(S_n > cn) \leq e^{-n(c-1-\ln c)}.$$

Next, verify that $c - 1 - \ln c > 0$ for every $c > 1$ (hint: let $f(c) = c - 1 - \ln c$; then check that $f(1) = 0$ and $f'(c) > 0$ for $c > 1$). Now we see that $\mathbb{P}(S_n > cn) \to 0$ exponentially fast as $n \to \infty$.

15 Central Limit Theorem

These problems are due on November 18

Exercise 15.1. A basketball player makes 80% of his free throws on the average. If the player attempts 40 free throws, what is the chance that he will make exactly 35? For an extra credit, find the exact answer (you can use the on-line calculator on the instructor's web page).

Solution: X is $b(40, 0.8)$; its normal approximation Y is $\mathcal{N}(32, 6.4)$. Probability is $\mathbb{P}(X = 35) = \mathbb{P}(34.5 < Y < 35.5) = \Phi\left(\frac{35.5-32}{\sqrt{6.4}}\right) - \Phi\left(\frac{34.5-32}{\sqrt{6.4}}\right) = \Phi(1.38) - \Phi(0.99) = 0.0773$. Exact answer: 0.0854.

Exercise 15.2. A basketball player makes 80% of his free throws on the average. The player attempts free throws repeatedly until he makes 25. What is the probability that at least 29 throws will be necessary? For an extra credit, find the exact answer (you can use the on-line calculator on the instructor's web page).

Solution: at least 29 throws are necessary means that 28 are not enough, so X is $b(28, 0.8)$; its normal approximation Y is $\mathcal{N}(22.4, 4.48)$. Probability is $\mathbb{P}(X \leq 24) = \mathbb{P}(Y < 24.5) = \Phi\left(\frac{24.5-22.4}{\sqrt{4.48}}\right) = \Phi(0.99) = 0.8389$. Exact answer: 0.8398.

Exercise 15.3. Suppose 1% of all screws made by a machine are defective. We are interested in the probability that a batch of 225 screws has at most one defective screw. Compute this probability by (a) Poisson approximation and (b) normal approximation. Compare your answers to the exact value, 0.34106, obtained by the on-line calculator on the instructor's web page. Which approximation works better when p is small and n is large?

Solution: X is $b(225, 0.01)$. (a) By Poisson approximation: $\lambda = 225 \times 0.01 = 2.25$, so the probability is $\mathbb{P}(X = 0) + \mathbb{P}(X = 1) \approx 2.25^0 e^{-2.25} + 2.25^1 e^{-2.25} = 0.3425$. (b) By normal approximation Y is $\mathcal{N}(2.25, 2.2275)$, so the probability is $\mathbb{P}(Y < 1.5) = \Phi\left(\frac{1.5-2.25}{\sqrt{2.2275}}\right) = \Phi(-0.50) = 0.3085$. The exact answer is 0.3411. The Poisson approximation is the better one.

Exercise 15.4. Suppose that, whenever invited to a party, the probability that a person attends with his or her guest is 1/3, attends alone is 1/3, and does not attend is 1/3. A company has invited all 300 of its employees and their guests to a Christmas party. What is the probability that at least 320 will attend?

Solution: for each person X_i takes three values: 0, 1, 2, each with probability 1/3. So $\mathbb{E}(X_i) = 1$ and $\mathsf{Var}(X_i) = 2/3$. Thus the normal approximation Y for the total S_{300} is $\mathcal{N}(300, 200)$. The probability is $\mathbb{P}(S_{300} \geq 320) \approx \mathbb{P}(Y \geq 319.5) = 1 - \Phi\left(\frac{319.5-300}{\sqrt{200}}\right) = 1 - \Phi(1.38) = 0.0838$.

Exercise 15.5. Suppose X is poisson(100). Use Central Limit Theorem to estimate $\mathbb{P}(85 \le X < 115)$.

Solution: the normal approximation Y is $\mathcal{N}(100, 100)$. The probability is $\mathbb{P}(85 \le X \le 114) \approx \mathbb{P}(84.5 < Y < 114.5) = \Phi\left(\frac{114.5-100}{10}\right) - \Phi\left(\frac{84.5-100}{\sqrt{10}}\right) = \Phi(1.45) - \Phi(-1.55) = 0.8659$.

Exercise 15.6. A die is rolled until the sum of the numbers obtained is larger than 200. What is the probability that you can do this in 66 rolls or fewer?

Solution: The event is that in 66 rolls the sum exceeds 200, i.e., is ≥ 201. The sum has mean $66 \times 3.5 = 231$ and variance $66 \times 2.92 = 192.72$. The normal approximation Y is $\mathcal{N}(231, 192.72)$. The probability is $\mathbb{P}(Y > 200.5) = 1 - \Phi\left(\frac{200.5-231}{\sqrt{192.72}}\right) = 1 - \Phi(-2.20) = 0.9861$.

Exercise 15.7. Suppose the weight of a certain brand of bolt has mean of 1 gram and a standard deviation of 0.06 gram. Use the central limit theorem to estimate the probability that 100 of these bolts weigh more than 101 grams.

Answer: 0.0475

Exercise 15.8 (Graduate)**.** Let $X_1, X_2, \ldots$ be a sequence of independent standard normal random variables. Let $S_n = X_1^2 + \cdots + X_n^2$; this variable is said to have χ^2-distribution with n degrees of freedom. Find the mean value and variance of S_n. Describe the normal approximation to S_n. Use it to estimate $\mathbb{P}\big(S_n \le n + \sqrt{2n}\big)$.

16 Random Walks (Gambler's Ruin)

These problems are due on December 2

Exercise 16.1. An investor has a stock that each week goes up \$1 with probability 52% or down \$1 with probability 48%. She bought the stock when it cost \$50 and will sell it when it reaches \$70 or falls to \$40.

(a) What is the probability that she will end up with \$70?

(b) Find the mean number of weeks the investor keeps the stock.

Exercise 16.2. An investor has a stock that each week goes up \$1 with probability 50% or down \$1 with probability 50%. She bought the stock when it cost \$50 and will sell it when it reaches \$80 or falls to \$30.

(a) What is the probability that she will end up with \$80?

(b) Find the mean number of weeks the investor keeps the stock.

Exercise 16.3. A casino owner wins \$1 with probability 0.55 and loses \$1 with probability 0.45. If he starts with \$60, what is the probability that he will ever hit \$0?

17 Poisson Process

These problems are optional, they will be graded for extra credit. They are due on December 9, after the last class meeting. Please put them in the plastic bin on the wall near the instructor's office. The corrected and graded problems can be picked up from the instructor's office or from the bin on December 12, 13, or in the final exam on December 14.

Exercise 17.1. Assume that fires in a city occur at a rate 14 per week.

(a) If there was no fire yesterday, what is the probability the there will be no fire today?

(b) What is the distribution of the waiting time till the next fire? Write down formulas for the distribution function, give its mean value and variance. State clearly what time units you are using.

(c) Let X be the number of fires in a month (30 days). What is the type of the random variable X and what is its parameter?

Exercise 17.2. Mushrooms grow in a forest randomly, with density 0.5 per square yard.

(a) If you find a mushroom, what is the chance that at least one more will be within one yard from it? What is the chance that there is exactly one mushroom within the distance one yard from the point you stay?

(Bonus) Let X be the distance to the nearest mushroom from the points you stay. Find the distribution function and the density function of X. (Do not just copy the formula from the class notes, provide calculation.)