

University of Washington
Department of Mathematics

Homework 7 Solutions

MATH 394: Probability I

Instructor: Arman Jahangiri

Submission: Single PDF on Gradescope
Deadline: 10:00 PM Pacific Time on the listed due date

This homework is worth **50 points**.

Across the quarter, there are **eight homework assignments**, worth **400 points total**. Homework assignments together account for **40%** of the final course grade. Further course policies can be seen in the **MATH 394 syllabus**.

Homework Policy

Submission and Deadline

Please:

- Submit homework as a single PDF on Gradescope by **10:00 PM (Pacific Time)** on the listed due date.
- Begin each problem on a new page, and ensure that pages are correctly tagged in Gradescope.
- Write solutions clearly and legibly.
- Typesetting solutions in LaTeX is encouraged, but not required.

Late Days

Each student is allotted **six late days** for the quarter. A late day extends a homework deadline by up to 24 hours without penalty. For example, submitting an assignment anytime between 10:01 PM on the due date and 10:00 PM the following day counts as one late day.

The following rules apply:

- No explanation is required to use late days.
- At most **two late days** may be used on any single homework assignment.
- Late days may not be used on the final homework assignment or on any assignment for which solutions have already been released.
- It is the student's responsibility to keep track of how many late days have been used during the quarter.

Once all late days have been exhausted, additional late submissions will incur a penalty of **10% per day**, up to a maximum deduction of **50%**. Assignments submitted more than five days late, or after solutions have been released, will not be accepted.

Submission Issues and Technical Difficulties

If a serious technical issue prevents a timely Gradescope submission, students may temporarily submit their assignment by email to the instructor at armanjg@uw.edu. In such cases:

- The submission must consist of a **single PDF file**. The email subject line must follow the format:

HW# Submission - FULL NAME - STUDENT ID.
- Once the Gradescope issue has been resolved, the student must upload the *exact same file* to Gradescope as soon as possible.
- The student should clearly indicate in the Gradescope submission that the assignment was previously submitted by email before the deadline.

Email submissions are intended only for genuine technical emergencies and should not be used as a substitute for timely Gradescope submission.

Problem 1.

Source: Blitzstein and Hwang, *Introduction to Probability*, Chapter 7, Problem 1.

Alice and Bob arrange to meet for lunch on a certain day at noon. However, neither is known for punctuality. They both arrive independently at uniformly distributed times between noon and 1 pm on that day. Each is willing to wait up to 15 minutes for the other to show up.

What is the probability they will meet for lunch that day?

Solution

Measure arrival times in hours after noon, and let Alice's and Bob's arrival times be

$$X, Y \sim \text{Unif}(0, 1),$$

independently.

Since X and Y are independent and each has density

$$f_X(x) = f_Y(y) = 1, \quad 0 \leq x, y \leq 1,$$

their joint density is

$$f_{X,Y}(x, y) = f_X(x)f_Y(y) = \begin{cases} 1, & 0 \leq x \leq 1, 0 \leq y \leq 1, \\ 0, & \text{otherwise.} \end{cases}$$

They meet exactly when their arrival times differ by at most 15 minutes, or $1/4$ of an hour. Thus

$$|X - Y| \leq \frac{1}{4}.$$

Therefore,

$$\mathbb{P}\left(|X - Y| \leq \frac{1}{4}\right) = \iint_{\{|x-y| \leq 1/4\}} f_{X,Y}(x, y) dy dx.$$

For a fixed x , the condition

$$|x - y| \leq \frac{1}{4}$$

is equivalent to

$$x - \frac{1}{4} \leq y \leq x + \frac{1}{4}.$$

Taking into account that $0 \leq y \leq 1$, we split the integral into three parts:

$$\begin{aligned} \mathbb{P}\left(|X - Y| \leq \frac{1}{4}\right) &= \int_0^{1/4} \int_0^{x+1/4} 1 dy dx \\ &\quad + \int_{1/4}^{3/4} \int_{x-1/4}^{x+1/4} 1 dy dx \\ &\quad + \int_{3/4}^1 \int_{x-1/4}^1 1 dy dx. \end{aligned}$$

Evaluating,

$$\begin{aligned} \mathbb{P}\left(|X - Y| \leq \frac{1}{4}\right) &= \int_0^{1/4} \left(x + \frac{1}{4}\right) dx + \int_{1/4}^{3/4} \frac{1}{2} dx + \int_{3/4}^1 \left(\frac{5}{4} - x\right) dx \\ &= \frac{3}{32} + \frac{1}{4} + \frac{3}{32} \\ &= \frac{7}{16}. \end{aligned}$$

Problem 2.

Source: Blitzstein and Hwang, *Introduction to Probability*, Chapter 7, Problem 7.

A stick of length L , where L is a positive constant, is broken at a uniformly random point X . Given that $X = x$, another breakpoint Y is chosen uniformly on the interval $[0, x]$.

- Find the joint PDF of X and Y . Be sure to specify the support.
- We already know that the marginal distribution of X is $\text{Unif}(0, L)$. Check that marginalizing out Y from the joint PDF agrees that this is the marginal distribution of X .
- We already know that the conditional distribution of Y given $X = x$ is $\text{Unif}(0, x)$. Check that using the definition of conditional PDFs agrees that this is the conditional distribution of Y given $X = x$.
- Find the marginal PDF of Y .
- Find the conditional PDF of X given $Y = y$.

Solution

(a) Since $X \sim \text{Unif}(0, L)$, we have

$$f_X(x) = \frac{1}{L}, \quad 0 < x < L.$$

Given $X = x$, the random variable Y is uniform on $(0, x)$, so

$$f_{Y|X}(y | x) = \frac{1}{x}, \quad 0 < y < x.$$

Thus

$$f_{X,Y}(x, y) = f_X(x)f_{Y|X}(y | x) = \frac{1}{Lx}, \quad 0 < y < x < L.$$

It is zero outside this region.

(b) Marginalizing out Y , for $0 < x < L$, gives

$$f_X(x) = \int_0^x \frac{1}{Lx} dy = \frac{1}{Lx} \cdot x = \frac{1}{L}.$$

This is exactly the $\text{Unif}(0, L)$ density.

(c) Using the definition of a conditional density,

$$f_{Y|X}(y | x) = \frac{f_{X,Y}(x, y)}{f_X(x)} = \frac{1/(Lx)}{1/L} = \frac{1}{x}, \quad 0 < y < x.$$

This is the uniform density on $(0, x)$, as expected.

(d) For a fixed y , the possible values of x run from y to L . Hence, for $0 < y < L$,

$$f_Y(y) = \int_y^L \frac{1}{Lx} dx = \frac{1}{L} \log\left(\frac{L}{y}\right).$$

(e) Finally,

$$f_{X|Y}(x | y) = \frac{f_{X,Y}(x, y)}{f_Y(y)} = \frac{1/(Lx)}{(1/L) \log(L/y)} = \frac{1}{x \log(L/y)},$$

for $y < x < L$.

Problem 3.

Source: Blitzstein and Hwang, *Introduction to Probability*, Chapter 7, Problem 15.

Let X and Y be continuous random variables, with joint CDF $F(x, y)$. Show that the probability that (X, Y) falls into the rectangle $[a_1, a_2] \times [b_1, b_2]$ is

$$F(a_2, b_2) - F(a_1, b_2) + F(a_1, b_1) - F(a_2, b_1).$$

Solution

We have covered this fully in the lecture with a visual representation (Lecture 19).

Recall that the joint CDF of X and Y is

$$F(x, y) = P(X \leq x, Y \leq y).$$

Since X and Y are continuous, let $f_{X,Y}$ denote their joint density. It is related to the joint CDF by

$$f_{X,Y}(x, y) = \frac{\partial^2}{\partial x \partial y} F(x, y).$$

The probability that (X, Y) falls inside the rectangle

$$[a_1, a_2] \times [b_1, b_2]$$

is therefore

$$P(a_1 \leq X \leq a_2, b_1 \leq Y \leq b_2) = \int_{a_1}^{a_2} \int_{b_1}^{b_2} f_{X,Y}(x, y) dy dx.$$

Substituting

$$f_{X,Y}(x, y) = \frac{\partial^2}{\partial x \partial y} F(x, y),$$

we obtain

$$P(a_1 \leq X \leq a_2, b_1 \leq Y \leq b_2) = \int_{a_1}^{a_2} \int_{b_1}^{b_2} \frac{\partial^2}{\partial x \partial y} F(x, y) dy dx.$$

First integrate with respect to y :

$$\begin{aligned} &= \int_{a_1}^{a_2} \left[\frac{\partial}{\partial x} F(x, y) \right]_{y=b_1}^{y=b_2} dx \\ &= \int_{a_1}^{a_2} \left(\frac{\partial}{\partial x} F(x, b_2) - \frac{\partial}{\partial x} F(x, b_1) \right) dx. \end{aligned}$$

Now integrate with respect to x :

$$\begin{aligned} &= [F(x, b_2)]_{a_1}^{a_2} - [F(x, b_1)]_{a_1}^{a_2} \\ &= (F(a_2, b_2) - F(a_1, b_2)) - (F(a_2, b_1) - F(a_1, b_1)). \end{aligned}$$

Hence,

$$P(a_1 \leq X \leq a_2, b_1 \leq Y \leq b_2) = F(a_2, b_2) - F(a_1, b_2) - F(a_2, b_1) + F(a_1, b_1).$$

Visual reasoning.

Recall that

$$F(a_2, b_2)$$

represents all probability mass in the region

$$X \leq a_2, \quad Y \leq b_2.$$

Think of this as one large lower-left rectangle extending up to

$$(a_2, b_2).$$

We only want the smaller rectangle

$$a_1 \leq X \leq a_2, \quad b_1 \leq Y \leq b_2.$$

Starting from

$$F(a_2, b_2),$$

we remove the part to the left of a_1 :

$$-F(a_1, b_2),$$

and remove the part below b_1 :

$$-F(a_2, b_1).$$

However, the lower-left corner

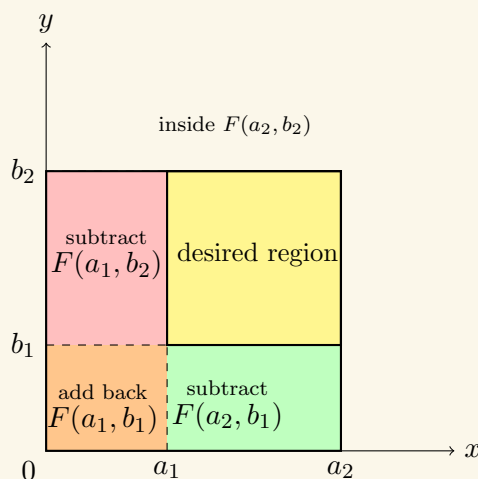
$$X \leq a_1, \quad Y \leq b_1$$

was contained in both of the pieces we subtracted, so it has been subtracted twice. We therefore add it back once:

$$+F(a_1, b_1).$$

Thus the desired rectangle is obtained by the two-dimensional inclusion-exclusion rule:

$$\text{desired rectangle} = \text{large rectangle} - \text{left rectangle} - \text{bottom rectangle} + \text{overlap}.$$



Problem 4.

Source: Blitzstein and Hwang, *Introduction to Probability*, Chapter 7, Problem 16.

Let X and Y have joint PDF

$$f_{X,Y}(x,y) = x + y, \quad 0 < x < 1, 0 < y < 1.$$

- (a) Check that this is a valid joint PDF.
- (b) Are X and Y independent?
- (c) Find the marginal PDFs of X and Y .
- (d) Find the conditional PDF of Y given $X = x$.

Solution

(a) First, note that if $x, y \in (0, 1)$, then $f_{X,Y}(x, y) = x + y \geq 0$. Otherwise, $f_{X,Y}(x, y) = 0 \geq 0$. So overall, $f_{X,Y}(x, y) \geq 0 \quad \forall x, y \in \mathbb{R}$.

Its total integral is

$$\int_0^1 \int_0^1 (x + y) dy dx = \int_0^1 \left(x + \frac{1}{2} \right) dx = \frac{1}{2} + \frac{1}{2} = 1.$$

Thus it is a valid joint PDF.

(b) The variables are not independent. One way to see this is to compute the marginals and observe that their product does not equal the joint density.

(c) For $0 < x < 1$,

$$f_X(x) = \int_0^1 (x + y) dy = x + \frac{1}{2}.$$

Similarly, for $0 < y < 1$,

$$f_Y(y) = \int_0^1 (x + y) dx = y + \frac{1}{2}.$$

(d) For $0 < x < 1$, the conditional density is

$$f_{Y|X}(y | x) = \frac{f_{X,Y}(x, y)}{f_X(x)} = \frac{x + y}{x + 1/2}, \quad 0 < y < 1.$$

Problem 5.

Source: Blitzstein and Hwang, *Introduction to Probability*, Chapter 7, Problem 17.

Let X and Y have joint PDF

$$f_{X,Y}(x,y) = cxy, \quad 0 < x < y < 1.$$

- (a) Find c to make this a valid joint PDF.
- (b) Are X and Y independent?
- (c) Find the marginal PDFs of X and Y .
- (d) Find the conditional PDF of Y given $X = x$.

Solution

(a) We need

$$\int_0^1 \int_0^y cxy \, dx \, dy = 1.$$

The integral is

$$c \int_0^1 y \left(\int_0^y x \, dx \right) dy = c \int_0^1 y \cdot \frac{y^2}{2} dy = \frac{c}{2} \cdot \frac{1}{4} = \frac{c}{8}.$$

Thus $c = 8$.

(b) The variables are not independent. The support itself already shows this, since the condition $0 < X < Y < 1$ ties the possible values of X and Y together.

(c) With $c = 8$, for $0 < x < 1$,

$$f_X(x) = \int_x^1 8xy \, dy = 4x(1 - x^2).$$

For $0 < y < 1$,

$$f_Y(y) = \int_0^y 8xy \, dx = 4y^3.$$

(d) For fixed x , the conditional support is $x < y < 1$. Therefore

$$f_{Y|X}(y | x) = \frac{8xy}{4x(1 - x^2)} = \frac{2y}{1 - x^2}, \quad x < y < 1.$$

Problem 6.

Source: Blitzstein and Hwang, *Introduction to Probability*, Chapter 7, Problem 21.

Find the probability that the quadratic polynomial

$$Ax^2 + Bx + 1,$$

where the coefficients A and B are determined by drawing i.i.d. $\text{Unif}(0, 1)$ random variables, has at least one real root.

Hint: By the quadratic formula, the polynomial $ax^2 + bx + c$ has a real root if and only if $b^2 - 4ac \geq 0$.

Solution

A quadratic polynomial $Ax^2 + Bx + 1$ has at least one real root if and only if its discriminant is nonnegative. Therefore,

$$B^2 - 4A \geq 0.$$

Equivalently,

$$A \leq \frac{B^2}{4}.$$

Since

$$A, B \stackrel{\text{i.i.d.}}{\sim} \text{Unif}(0, 1),$$

their marginal densities are

$$f_A(a) = \begin{cases} 1, & 0 \leq a \leq 1, \\ 0, & \text{otherwise,} \end{cases} \quad f_B(b) = \begin{cases} 1, & 0 \leq b \leq 1, \\ 0, & \text{otherwise.} \end{cases}$$

Because A and B are independent, their joint density is

$$f_{A,B}(a, b) = f_A(a)f_B(b) = \begin{cases} 1, & 0 \leq a \leq 1, 0 \leq b \leq 1, \\ 0, & \text{otherwise.} \end{cases}$$

Hence the desired probability is

$$P(B^2 - 4A \geq 0) = P\left(A \leq \frac{B^2}{4}\right).$$

Using the joint density,

$$P\left(A \leq \frac{B^2}{4}\right) = \iint_{\{(a,b): 0 \leq b \leq 1, 0 \leq a \leq b^2/4\}} f_{A,B}(a, b) da db.$$

Since $f_{A,B}(a, b) = 1$ on the support,

$$\begin{aligned} P\left(A \leq \frac{B^2}{4}\right) &= \int_0^1 \int_0^{b^2/4} 1 da db \\ &= \int_0^1 \frac{b^2}{4} db \\ &= \frac{1}{4} \left[\frac{b^3}{3} \right]_0^1 \\ &= \frac{1}{12} \approx 0.0834 \end{aligned}$$

In other words, only 8.34% of such quadratic polynomials have at least one real root!

Problem 7.

Source: Blitzstein and Hwang, *Introduction to Probability*, Chapter 7, Problem 39–40.

- (a) Two fair, six-sided dice are rolled, one green and one orange, with outcomes X and Y for the green die and the orange die, respectively.
- (i) Compute the covariance of $X + Y$ and $X - Y$.
 - (ii) Are $X + Y$ and $X - Y$ independent?
- (b) Let X and Y be i.i.d. $\text{Unif}(0, 1)$.
- (i) Compute the covariance of $X + Y$ and $X - Y$.
 - (ii) Are $X + Y$ and $X - Y$ independent?

Solution

(a) Since the two dice are independent and identically distributed,

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

The cross-covariances are zero, and $\text{Var}(X) = \text{Var}(Y)$. Therefore

$$\text{Cov}(X + Y, X - Y) = \text{Var}(X) - \text{Var}(Y) = 0.$$

They are not independent. For instance, if $X + Y = 2$, then necessarily $X = Y = 1$, so $X - Y = 0$. Knowing the sum can restrict the difference.

Final Answer for Part (a)

$$\text{Cov}(X + Y, X - Y) = 0,$$

but $X + Y$ and $X - Y$ are not independent.

(b) The same covariance calculation applies for independent uniforms. We again get

$$\text{Cov}(X + Y, X - Y) = \text{Var}(X) - \text{Var}(Y) = 0.$$

They are not independent. The pair $(X + Y, X - Y)$ cannot take arbitrary values in a rectangle. For example, if $X + Y$ is very close to 0, then both X and Y must be close to 0, forcing $X - Y$ also to be close to 0.

Final Answer for Part (b)

$$\text{Cov}(X + Y, X - Y) = 0,$$

but $X + Y$ and $X - Y$ are not independent.

Problem 8.

Source: Blitzstein and Hwang, *Introduction to Probability*, Chapter 7, Problem 65.

Let

$$(X_1, \dots, X_k)$$

be Multinomial with parameters

$$n \quad \text{and} \quad (p_1, \dots, p_k).$$

Use indicator random variables to show that

$$\text{Cov}(X_i, X_j) = -np_i p_j, \quad i \neq j.$$

Solution

Think of the multinomial vector as coming from n independent trials, where each trial falls into exactly one of the k categories. For trial r , define

$$I_{r,i} = \mathbf{1}\{\text{trial } r \text{ falls in category } i\}.$$

Then

$$X_i = \sum_{r=1}^n I_{r,i}, \quad X_j = \sum_{s=1}^n I_{s,j}.$$

Therefore

$$\text{Cov}(X_i, X_j) = \sum_{r=1}^n \sum_{s=1}^n \text{Cov}(I_{r,i}, I_{s,j}).$$

When $r \neq s$, the indicators come from different independent trials, so their covariance is zero. When $r = s$, the two indicators cannot both be 1 because a single trial cannot fall into both category i and category j . Hence

$$E(I_{r,i} I_{r,j}) = 0,$$

while

$$E(I_{r,i}) = p_i, \quad E(I_{r,j}) = p_j.$$

Thus

$$\text{Cov}(I_{r,i}, I_{r,j}) = 0 - p_i p_j = -p_i p_j.$$

There are n such same-trial terms, so

$$\text{Cov}(X_i, X_j) = n(-p_i p_j) = -np_i p_j.$$

Final Answer

$$\text{Cov}(X_i, X_j) = -np_i p_j, \quad i \neq j.$$