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. Meeting for Lunch

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. They meet exactly when their arrival times differ by at most 15 minutes, which is one quarter of an hour. Thus the desired event is

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

Geometrically, the pair (X, Y) is uniformly distributed over the unit square. The event $|X-Y| > 1/4$ consists of two congruent right triangles, one above the line $y = x + 1/4$ and one below the line $y = x - 1/4$. Each triangle has side length $3/4$, so the total excluded area is

$$2 \cdot \frac{1}{2} \left(\frac{3}{4}\right)^2 = \frac{9}{16}.$$

Therefore the desired probability is

$$1 - \frac{9}{16} = \frac{7}{16}.$$

Final Answer

$$\frac{7}{16}$$

Problem 2. Two Breakpoints on a Stick

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 $[0, x]$.

(a) Find the joint PDF of X and Y. Be sure to specify the support.

(b) 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.

(c) 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 $X = x$.

(d) Find the marginal PDF of Y.

(e) Find the conditional PDF of X given $Y = y$.

Solution

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

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

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

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

Thus

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

It is zero outside this region.

Final Answer for Part (a)

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

(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.

Final Answer for Part (b)

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

(c) Using the definition of a conditional density,

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

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

Final Answer for Part (c)

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

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

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

Final Answer for Part (d)

$$f_Y(y) = \frac{1}{L}\log\left(\frac{L}{y}\right), \qquad 0 < y < L.$$

(e) Finally,

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

for $y < x < L$.

Final Answer for Part (e)

$$f_{X|Y}(x \mid y) = \frac{1}{x\log(L/y)}, \qquad y < x < L.$$

Problem 3. Broken Stick and Table Legs

Source: Blitzstein and Hwang, Introduction to Probability, Chapter 7, Problem 14.

(a) A stick is broken into three pieces by picking two points independently and uniformly along the stick, and breaking the stick at those two points. What is the probability that the three pieces can be assembled into a triangle?

Hint: A triangle can be formed from 3 line segments of lengths a, b, c if and only if $a, b, c \in (0, 1/2)$. The probability can be interpreted geometrically as proportional to an area in the plane, avoiding all calculus, but make sure for that approach that the distribution of the random point in the plane is Uniform over some region.

(b) Three legs are positioned uniformly and independently on the perimeter of a round table. What is the probability that the table will stand?

Solution

(a) Scale the stick to have length 1. Let the two breakpoints be independent uniform points on $(0, 1)$, and order them as $U < V$. The three piece lengths are

$$U, \qquad V - U, \qquad 1 - V.$$

A triangle can be formed exactly when no piece has length at least $1/2$. The ordered pair (U, V) is uniform over the triangle

$$0 < U < V < 1.$$

Inside this triangle, the forbidden regions are

$$U > \frac{1}{2}, \qquad V - U > \frac{1}{2}, \qquad 1 - V > \frac{1}{2}.$$

These three regions are congruent corner triangles whose total area is three quarters of the area of the full region. Therefore the remaining area is one quarter of the full region.

Final Answer for Part (a)

$$\frac{1}{4}$$

(b) The table stands exactly when the three points on the circumference are not all contained in some semicircle. Equivalently, the triangle formed by the three leg positions contains the center of the table.

For three independent uniform points on a circle, the probability that they all lie in some semicircle is $3/4$. Thus the complementary probability, namely the probability that the table stands, is

$$1 - \frac{3}{4} = \frac{1}{4}.$$

Final Answer for Part (b)

$$\frac{1}{4}$$

Problem 4. Rectangle Probability from a Joint CDF

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

The joint CDF is

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

We want the probability of the rectangle

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

Start with the probability of the lower-left rectangle ending at (a_2, b_2), namely $F(a_2, b_2)$. This includes too much. It includes points with $X \leq a_1$, whose probability is $F(a_1, b_2)$, and points with $Y \leq b_1$, whose probability is $F(a_2, b_1)$. After subtracting these two pieces, the corner $(X \leq a_1, Y \leq b_1)$ has been subtracted twice, so it must be added back once. Thus

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

This is the same expression as the one stated in the problem, only with the last two terms written in the usual order.

Final Answer

$$P\big((X, Y) \in [a_1, a_2] \times [b_1, b_2]\big) = F(a_2, b_2) - F(a_1, b_2) + F(a_1, b_1) - F(a_2, b_1).$$

Problem 5. Joint PDF I

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

Let X and Y have joint PDF

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

(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) The function is nonnegative on the unit square. 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.

Final Answer for Part (a)

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.

Final Answer for Part (b)

X and Y are not independent.

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

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

Similarly, for $0 < y < 1$,

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

Final Answer for Part (c)

$$f_X(x) = x + \frac{1}{2}, \qquad f_Y(y) = y + \frac{1}{2}, \qquad 0 < x, y < 1.$$

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

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

Final Answer for Part (d)

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

Problem 6. Joint PDF II

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

Let X and Y have joint PDF

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

(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 $c = 8$.

Final Answer for Part (a)

$$c = 8$$

(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.

Final Answer for Part (b)

X and Y are not independent.

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

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

For $0 < y < 1$,

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

Final Answer for Part (c)

$$f_X(x) = 4x(1-x^2), \qquad 0 < x < 1,$$
$$f_Y(y) = 4y^3, \qquad 0 < y < 1.$$

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

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

Final Answer for Part (d)

$$f_{Y|X}(y \mid x) = \frac{2y}{1 - x^2}, \qquad x < y < 1.$$

Problem 7. Random Quadratic

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

The polynomial has a real root exactly when its discriminant is nonnegative. Here that condition is

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

or equivalently

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

Since (A, B) is uniformly distributed on the unit square, the desired probability is the area of the region under the curve $a = b^2/4$, for $0 < b < 1$. Hence

$$P\left(A \leq \frac{B^2}{4}\right) = \int_0^1 \frac{b^2}{4}\, db = \frac{1}{4} \cdot \frac{1}{3} = \frac{1}{12}.$$

Final Answer

$$\frac{1}{12}$$

Problem 8. Distance Between Two Uniforms

Source: Blitzstein and Hwang, Introduction to Probability, Chapter 7, Problem 31.

Let X and Y be i.i.d. $\text{Unif}(0, 1)$. Find the standard deviation of the distance between X and Y.

Solution

Let

$$D = |X - Y|.$$

We need $\text{SD}(D)$. First,

$$E(D^2) = E\big((X - Y)^2\big) = \text{Var}(X - Y).$$

Since X and Y are independent uniforms,

$$\text{Var}(X - Y) = \text{Var}(X) + \text{Var}(Y) = \frac{1}{12} + \frac{1}{12} = \frac{1}{6}.$$

Also, by the usual triangular-density calculation for the distance between two independent uniforms,

$$E(D) = \int_0^1 2d(1 - d)\, dd = \frac{1}{3}.$$

Therefore

$$\text{Var}(D) = E(D^2) - [E(D)]^2 = \frac{1}{6} - \frac{1}{9} = \frac{1}{18}.$$

Thus

$$\text{SD}(D) = \sqrt{\frac{1}{18}} = \frac{1}{3\sqrt{2}}.$$

Final Answer

$$\text{SD}(|X - Y|) = \frac{1}{3\sqrt{2}}.$$

Problem 9. Minimum and Maximum of Geometrics

Source: Blitzstein and Hwang, Introduction to Probability, Chapter 7, Problem 29.

Let X and Y be i.i.d. Geom(p), and let

$$L = \min(X, Y), \qquad M = \max(X, Y).$$

(a) Find the joint PMF of L and M. Are they independent?

(b) Find the marginal distribution of L in two ways: using the joint PMF, and using a story.

(c) Find $E(M)$.
Hint: A quick way is to use part (b) and the fact that $L + M = X + Y$.

(d) Find the joint PMF of L and $M - L$. Are they independent?

Solution

Let $q = 1 - p$, and use the convention that Geom(p) counts the number of failures before the first success, so

$$P(X = k) = q^k p, \qquad k = 0, 1, 2, \ldots.$$

(a) If $0 \le l < m$, then one of X, Y must equal l and the other must equal m, so

$$P(L = l, M = m) = 2p^2 q^{l+m}.$$

If $l = m$, then both variables must equal l, so

$$P(L = l, M = l) = p^2 q^{2l}.$$

The variables L and M are not independent. For example, the event $M < L$ has probability zero, so knowing L restricts possible values of M.

Final Answer for Part (a)

$$P(L = l, M = m) = \begin{cases} p^2 q^{2l}, & m = l, \\ 2p^2 q^{l+m}, & 0 \le l < m, \\ 0, & \text{otherwise.} \end{cases}$$

They are not independent.

(b) Using the joint PMF,

$$P(L = l) = p^2 q^{2l} + \sum_{m=l+1}^{\infty} 2p^2 q^{l+m}.$$

The geometric sum gives

$$P(L = l) = p^2 q^{2l} + 2p^2 q^l \frac{q^{l+1}}{1 - q} = p^2 q^{2l} + 2pq^{2l+1} = p(2 - p)q^{2l}.$$

The story proof is even cleaner. The event $L = l$ means that during the first l paired trials, both sequences fail, and at time l, at least one of them succeeds. The probability of both failing in a paired trial is q^2, and the probability that at least one succeeds is $1 - q^2 = p(2 - p)$. Hence

$$P(L = l) = (q^2)^l (1 - q^2) = p(2 - p)q^{2l}.$$

Final Answer for Part (b)

$$L \sim \text{Geom}\left(1-q^2\right), \qquad P(L=l) = p(2-p)q^{2l}.$$

(c) Since

$$L + M = X + Y,$$

we have

$$E(M) = E(X) + E(Y) - E(L).$$

For our convention, $E(X) = E(Y) = q/p$. Also, from part (b),

$$E(L) = \frac{q^2}{1-q^2} = \frac{q^2}{p(2-p)}.$$

Therefore

$$E(M) = \frac{2q}{p} - \frac{q^2}{p(2-p)}.$$

Final Answer for Part (c)

$$E(M) = \frac{2q}{p} - \frac{q^2}{p(2-p)}, \qquad q = 1-p.$$

(d) Let $D = M - L$. If $d = 0$, then $M = L = l$, so

$$P(L=l, D=0) = p^2 q^{2l}.$$

If $d > 0$, then (X, Y) can be $(l, l+d)$ or $(l+d, l)$, so

$$P(L=l, D=d) = 2p^2 q^{2l+d}.$$

The factorization shows that L and $D = M - L$ are independent. Indeed,

$$P(L=l) = p(2-p)q^{2l},$$

while

$$P(D=0) = \frac{p}{2-p}, \qquad P(D=d) = \frac{2p}{2-p} q^d, \quad d = 1, 2, \ldots,$$

and multiplying these marginals recovers the joint PMF above.

Final Answer for Part (d)

$$P(L=l, M-L=d) = \begin{cases} p^2 q^{2l}, & d = 0, \\ 2p^2 q^{2l+d}, & d = 1, 2, \ldots, \end{cases}$$

for $l = 0, 1, 2, \ldots$. The variables L and $M - L$ are independent.

Problem 10. Covariance and Independence

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 11. Multinomial Covariance

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

Let

$$(X_1, \ldots, X_k)$$

be Multinomial with parameters

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

Use indicator random variables to show that

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

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

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

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

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

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

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