Jingjia Chen, Michael Psenka and Tarang Srivastava

HW 8

Due: Thursday 8/11, 11:59 PM
Grace period until Friday 8/12, 11:59 PM

Sundry

Before you start writing your final homework submission, state briefly how you worked on it. Who else did you work with? List names and email addresses. (In case of homework party, you can just describe the group.)

1 Short Answer

(a) Let X be uniform on the interval $[0,2]$, and define $Y = 2X+1$. Find the PDF, CDF, expectation, and variance of Y.

(b) Let X and Y have joint distribution

$$f(x,y) = \begin{cases} cxy + \frac{1}{4} & x \in [1,2] \text{ and } y \in [0,2] \\ 0 & \text{otherwise.} \end{cases}$$

Find the constant c. Are X and Y independent?

(c) Let $X \sim \text{Exp}(3)$.

(i) Find probability that $X \in [0, 1]$.

(ii) Let $Y = \lfloor X \rfloor$. For each $k \in \mathbb{N}$, what is the probability that $Y = k$? Write the distribution of Y in terms of one of the famous distributions; provide that distribution's name and parameters.

(d) Let $X_i \sim \text{Exp}(\lambda_i)$ for $i = 1, \ldots, n$ be mutually independent. It is a (very nice) fact that $\min(X_1, \ldots, X_n) \sim \text{Exp}(\mu)$. Find μ.

2 It's Raining Fish

A hurricane just blew across the coast and flung a school of fish onto the road nearby the beach. The road starts at your house and is infinitely long. We will label a point on the road by its distance from your house (in miles). For each $n \in \mathbb{N}$, the number of fish that land on the segment of the road $[n, n+1]$ is independently Poisson(λ) and each fish that is flung into that segment of the road lands uniformly at random within the segment. Keep in mind that you can cite any result from lecture or discussion without proof.

(a) What is the distribution of the number of fish arriving in segment $[0,n]$ of the road, for some $n \in \mathbb{N}$?

(b) Let $[a,b]$ be an interval in $[0,1]$. What is the distribution of the number of fish that lands in the segment $[a,b]$ of the road?

(c) Let $[a,b]$ be any interval such that $a \geq 0$. What is the distribution of the number of fish that land in $[a,b]$?

(d) Suppose you take a stroll down the road. What is the distribution of the distance you walk (in miles) until you encounter the first fish?

(e) Suppose you encounter a fish at distance x. What is the distribution of the distance you walk until you encounter the next fish?

3 Waiting For the Bus

Edward and Jerry are waiting at the bus stop outside of Soda Hall.

Like many bus systems, buses arrive in periodic intervals. However, the Berkeley bus system is unreliable, so the length of these intervals are random, and follow Exponential distributions.

Edward is waiting for the 51B, which arrives according to an Exponential distribution with parameter λ. That is, if we let the random variable X_i correspond to the difference between the arrival time ith and $(i-1)$st bus (also known as the inter-arrival time) of the 51B, $X_i \sim \text{Expo}(\lambda)$.

Jerry is waiting for the 79, whose inter-arrival times also follows Exponential distributions with parameter μ. That is, if we let Y_i denote the inter-arrival time of the 79, $Y_i \sim \text{Expo}(\mu)$. Assume that all inter-arrival times are independent.

(a) What is the probability that Jerry's bus arrives before Edward's bus?

(b) After 20 minutes, the 79 arrives, and Jerry rides the bus. However, the 51B still hasn't arrived yet. Let D be the additional amount of time Edward needs to wait for the 51B to arrive. What is the distribution of D?

(c) Lavanya isn't picky, so she will wait until either the 51B or the 79 bus arrives. Find the distribution of Z, the amount of time Lavanya will wait before catching her bus.

(d) Khalil doesn't feel like riding the bus with Edward. He decides that he will wait for the second arrival of the 51B to ride the bus. Find the distribution of $T = X_1 + X_2$, the amount of time that Khalil will wait to ride the bus.

4 Conditioning on Exponentials

Let X_i be i.i.d. Expo(λ) random variables.

(a) Compute $\mathbb{E}[Y \mid Z]$, where $Y = \max\{X_1, X_2\}$ and $Z = \min\{X_1, X_2\}$.

(b) Compute $\mathbb{E}[X_1 + X_2 \mid Z]$. [*Hint*: Use part (a).]

(c) Use part (b) to compute $\mathbb{E}[Z]$.

(d) Compute $\mathbb{E}[X_1 + X_2 \mid X_1 + X_2 + X_3]$.

5 Exponential Distributions: Lightbulbs

A brand new lightbulb has just been installed in our classroom, and you know the life span of a lightbulb is exponentially distributed with a mean of 50 days.

(a) Suppose an electrician is scheduled to check on the lightbulb in 30 days and replace it if it is broken. What is the probability that the electrician will find the bulb broken?

(b) Suppose the electrician finds the bulb broken and replaces it with a new one. What is the probability that the new bulb will last at least 30 days?

(c) Suppose the electrician finds the bulb in working condition and leaves. What is the probability that the bulb will last at least another 30 days?

6 Noisy Love

Suppose you have confessed to your love interest on Valentine's Day and you are waiting to hear back. Your love interest is trying to send you a binary message: "0" means that your love interest is not interested in you, while "1" means that your love interest reciprocates your feelings. Let X be your love interest's message for you. Your current best guess of X has $\mathbb{P}(X = 0) = 0.7$ and $\mathbb{P}(X = 1) = 0.3$. Unfortunately, your love interest sends you the message through a noisy channel, and instead of receiving the message X, you receive the message $Y = X + \varepsilon$, where ε is independent Gaussian noise with mean 0 and variance 0.49.

(a) First, you decide upon the following rule: if you observe $Y > 0.5$, then you will assume that your love interest loves you back, whereas if you observe $Y \leq 0.5$, then you will assume that your love interest is not interested in you. What is the probability that you are correct using this rule? (Express your answer in terms of the CDF of the standard Gaussian distribution $\Phi(z) = \mathbb{P}(\mathcal{N}(0,1) \leq z)$, and then evaluate your answer numerically.)

(b) Suppose you observe $Y = 0.6$. What is the probability that your love interest loves you back? [*Hint*: This problem requires conditioning on an event of probability 0, namely, the event $\{Y = 0.6\}$. To tackle this problem, think about conditioning on the event $\{Y \in [0.6, 0.6 + \delta]\}$, where $\delta > 0$ is small, so that $f_Y(0.6) \cdot \delta \approx \mathbb{P}(Y \in [0.6, 0.6 + \delta])$, and then apply Bayes Rule.]

(c) Suppose you observe $Y = y$. For what values is it more likely than not that your love interest loves you back? [*Hint*: As before, instead of considering $\{Y = y\}$, you can consider the event $\{Y \in [y, y+\delta]\}$ for small $\delta > 0$. So, when is $\mathbb{P}(X = 1 \mid Y \in [y, y+\delta]) \geq \mathbb{P}(X = 0 \mid Y \in [y, y+\delta])$?]

(d) Your new rule is to assume that your love interest loves you back if (based on the value of Y that you observe) it is more likely than not that your love interest loves you back. Under this new rule, what is the probability that you are correct?

7 Sum of Independent Gaussians

In this question, we will introduce an important property of the Gaussian distribution: the sum of independent Gaussians is also a Gaussian.

Let X and Y be independent standard Gaussian random variables. Recall that the density of the standard Gaussian is

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right).$$

(a) What is the joint density of X and Y?

(b) Observe that the joint density of X and Y, $f_{X,Y}(x,y)$, only depends on the quantity $x^2 + y^2$, which is the distance from the origin. In other words, the Gaussian is *rotationally symmetric*. Next, we will try to find the density of $X+Y$. To do this, draw a picture of the Cartesian plane and draw the region $x + y \leq c$, where c is a real number of your choice.

(c) Now, rotate your picture clockwise by $\pi/4$ so that the line $X+Y = c$ is now vertical. Redraw your figure. Let X' and Y' denote the random variables which correspond to the $\pi/4$ clockwise rotation of (X,Y) and express the new shaded region in terms of X' and Y'.

(d) By rotational symmetry of the Gaussian, (X',Y') has the same distribution as (X,Y). Argue that $X+Y$ has the same distribution as $\sqrt{2}Z$, where Z is a standard Gaussian. This proves the following important fact: *the sum of independent Gaussians is also a Gaussian*. Notice that $X \sim \mathcal{N}(0,1)$, $Y \sim \mathcal{N}(0,1)$ and $X+Y \sim \mathcal{N}(0,2)$. *In general, if X and Y are independent Gaussians, then $X+Y$ is a Gaussian with mean $\mu_X + \mu_Y$ and variance $\sigma_X^2 + \sigma_Y^2$.*

(e) Recall the CLT:

> If $\{X_i\}_{i\in\mathbb{N}}$ is a sequence of i.i.d. random variables with mean $\mu \in \mathbb{R}$ and variance $\sigma^2 < \infty$, then:
>
> $$\frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} \xrightarrow{\text{in distribution}} \mathcal{N}(0,1) \qquad \text{as } n \to \infty.$$

Prove that the CLT holds for the special case when the X_i are i.i.d. $\mathcal{N}(0,1)$.

8 Exponential LLSE

Let $X \sim U[0,a]$ and let $Y = e^X$. Compute $L[Y \mid X]$. What does $L[Y \mid X]$ approach as $a \to 0$?

9 LLSE

We have two bags of balls. The fractions of red balls and blue balls in bag A are $2/3$ and $1/3$ respectively. The fractions of red balls and blue balls in bag B are $1/2$ and $1/2$ respectively. Someone gives you one of the bags (unmarked) uniformly at random. You then draw 6 balls from that same bag with replacement. Let X_i be the indicator random variable that ball i is red. Now, let us define $X = \sum_{1 \le i \le 3} X_i$ and $Y = \sum_{4 \le i \le 6} X_i$.

(a) Compute $\mathbb{E}[X]$ and $\mathbb{E}[Y]$.

(b) Compute $\text{Var}(X)$.

(c) Compute $\text{cov}(X,Y)$. (*Hint*: Recall that covariance is bilinear.)

(d) Now, we are going to try and predict Y from a value of X. Compute $L(Y \mid X)$, the best linear estimator of Y given X. (*Hint*: Recall that

$$L(Y \mid X) = \mathbb{E}[Y] + \frac{\text{cov}(X,Y)}{\text{Var}(X)}\big(X - \mathbb{E}[X]\big).$$

)

10 Continuous LLSE

Suppose that X and Y are uniformly distributed on the shaded region in the figure below.

That is, X and Y have the joint distribution:

$$f_{X,Y}(x,y) = \begin{cases} 1/2, & 0 \le x \le 1, 0 \le y \le 1 \\ 1/2, & 1 \le x \le 2, 1 \le y \le 2 \end{cases}$$

(a) Do you expect X and Y to be positively correlated, negatively correlated, or neither?

(b) Compute the marginal distribution of X.

(c) Compute $L[Y \mid X]$, the best linear estimator of Y given X.

(d) What is $\mathbb{E}[Y \mid X]$?

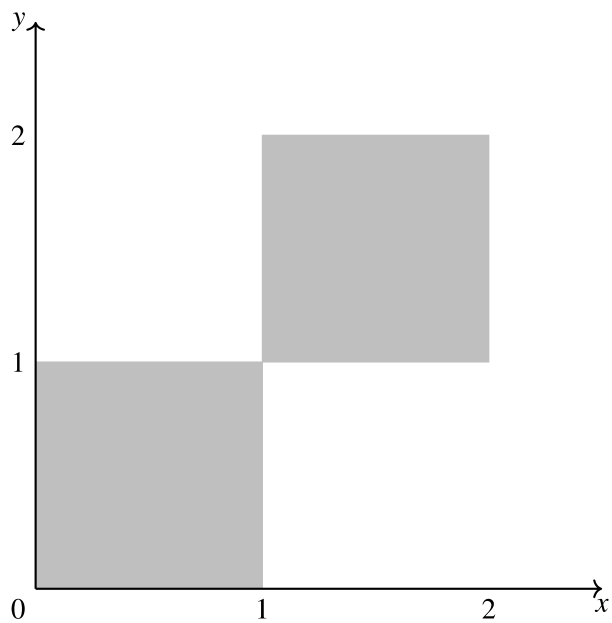

Figure 1: The joint density of (X,Y) is uniform over the shaded region.