Recap Continuous Joint Distributions, Law of Total Probability (continuous) Law of Total Expectation

CSE 312 Spring 26

Lecture 22

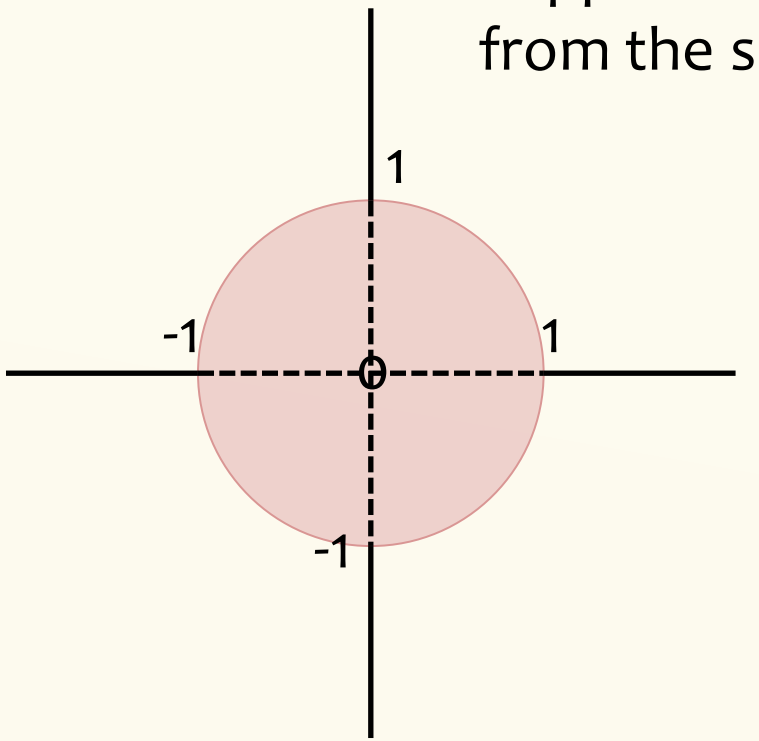
Agenda

- Recap – joint continuous distns
- Continuous law of total probability
- Law of total expectation

Reference Sheet

	Discrete	Continuous
Joint PMF/PDF	$p_{X,Y}(x,y) = P(X = x, Y = y)$	$f_{X,Y}(x,y) \neq P(X = x, Y = y)$
Joint CDF	$F_{X,Y}(x,y) = \sum_{t \leq x} \sum_{s \leq y} p_{X,Y}(t,s)$	$F_{X,Y}(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(t,s)\,ds\,dt$
Normalization	$\sum_{x} \sum_{y} p_{X,Y}(x,y) = 1$	$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx\,dy = 1$
Marginal PMF/PDF	$p_X(x) = \sum_{y} p_{X,Y}(x,y)$	$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy$
Expectation	$E[g(X,Y)] = \sum_{x} \sum_{y} g(x,y) p_{X,Y}(x,y)$	$E[g(X,Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) f_{X,Y}(x,y)\,dx\,dy$
Independence	$\forall x, y, p_{X,Y}(x,y) = p_X(x) p_Y(y)$	$\forall x, y, f_{X,Y}(x,y) = f_X(x) f_Y(y)$

Example – Uniform distribution on a unit disk

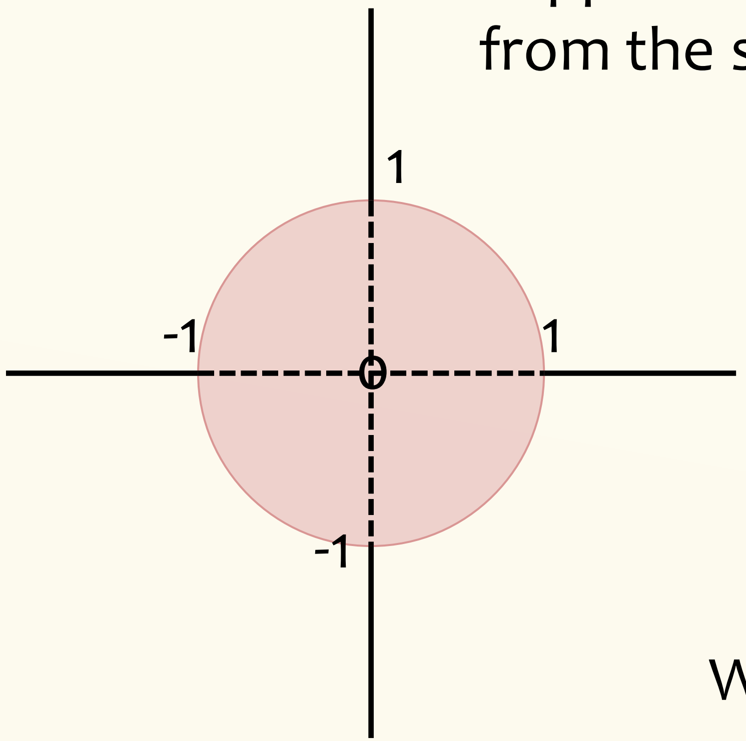


Suppose that a pair of random variables (X, Y) is chosen uniformly from the set of real points (x, y) such that $x^2 + y^2 \leq 1$

What is the joint density?

$$f_{X,Y}(x,y) = \begin{cases} c & \text{if } x^2 + y^2 \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Example – Uniform distribution on a unit disk



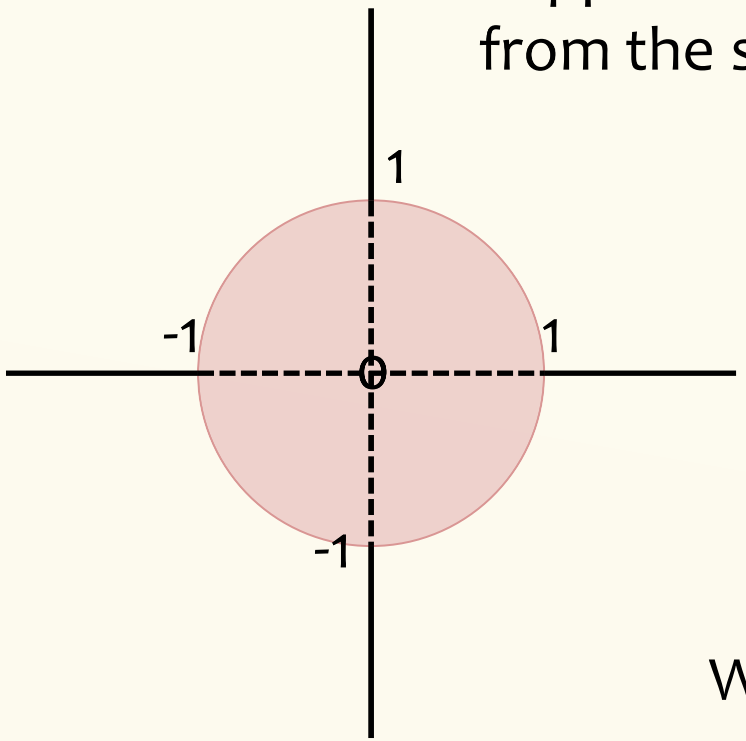
Suppose that a pair of random variabes (X, Y) is chosen uniformly from the set of real points (x, y) such that $x^2 + y^2 \leq 1$

This is a disk of radius 1 which has area π

$$f_{X,Y}(x,y) = \begin{cases} \dfrac{1}{\pi} & \text{if } x^2 + y^2 \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

What is $f_X(x)$?

Example – Uniform distribution on a unit disk



Suppose that a pair of random variables (X, Y) is chosen uniformly from the set of real points (x, y) such that $x^2 + y^2 \leq 1$

This is a disk of radius 1 which has area π

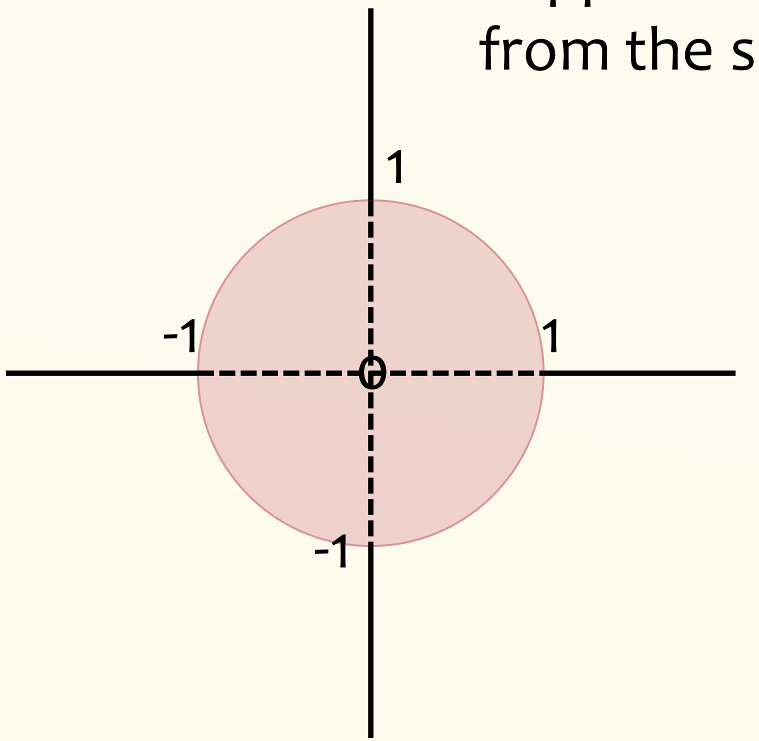
$$f_{X,Y}(x, y) = \begin{cases} \dfrac{1}{\pi} & \text{if } x^2 + y^2 \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

What is $f_X(x)$?

$$f_X(x) = \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \frac{1}{\pi} \mathrm{d}y$$

$$= 2\sqrt{1 - x^2}/\pi$$

Example – Uniform distribution on a unit disk



Suppose that a pair of random variables (X, Y) is chosen uniformly from the set of real points (x, y) such that $x^2 + y^2 \leq 1$

This is a disk of radius 1 which has area π

$$f_{X,Y}(x,y) = \begin{cases} \dfrac{1}{\pi} & \text{if } x^2 + y^2 \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

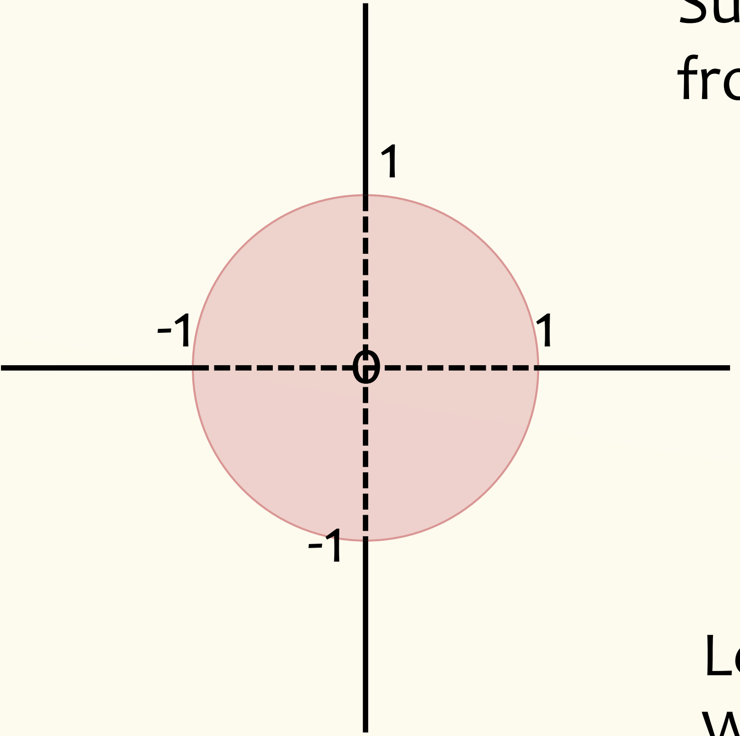
$$f_X(x) = \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \frac{1}{\pi} \mathrm{d}y = 2\sqrt{1-x^2}/\pi$$

$$f_Y(y) = \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} \frac{1}{\pi} \mathrm{d}x = 2\sqrt{1-y^2}/\pi$$

Are X and Y independent?

Is $f_{X,Y}(x,y) = f_X(x) \cdot f_Y(y)$ for all $x, y \in \mathbb{R}$?

Example – Uniform distribution on a unit disk



Suppose that a pair of random variables (X, Y) is chosen uniformly from the set of real points (x, y) such that $x^2 + y^2 \leq 1$

This is a disk of radius 1 which has area π

$$f_{X,Y}(x,y) = \begin{cases} \dfrac{1}{\pi} & \text{if } x^2 + y^2 \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Let D be the distance of the point (X, Y) from the origin.
What is $\mathrm{P}(\mathrm{D} \leq \mathrm{d})$? What is $f_{\mathrm{D}}(\mathrm{d})$? What is $\mathrm{E}(\mathrm{D})$?

Example – More joint densities

The joint density of X and Y is

$$f_{X,Y}(x,y) = \begin{cases} 2e^{-x}e^{-2y} & 0 \le \mathrm{x,y} \le \infty \\ 0 & \text{otherwise} \end{cases}$$

What is $\mathrm{P}(X > 1, Y < 1)$? What is $\mathrm{P}(X > \mathrm{Y})$? What is $f_X(x)$?

Reference Sheet

	Discrete	Continuous
Joint PMF/PDF	$p_{X,Y}(x,y) = P(X = x, Y = y)$	$f_{X,Y}(x,y) \neq P(X = x, Y = y)$
Joint CDF	$F_{X,Y}(x,y) = \sum_{t \le x} \sum_{s \le y} p_{X,Y}(t,s)$	$F_{X,Y}(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(t,s)\,ds\,dt$
Normalization	$\sum_{x} \sum_{y} p_{X,Y}(x,y) = 1$	$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx\,dy = 1$
Marginal PMF/PDF	$p_X(x) = \sum_{y} p_{X,Y}(x,y)$	$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy$
Expectation	$E[g(X,Y)] = \sum_{x} \sum_{y} g(x,y) p_{X,Y}(x,y)$	$E[g(X,Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) f_{X,Y}(x,y)\,dx\,dy$
Independence	$\forall x, y, p_{X,Y}(x,y) = p_X(x) p_Y(y)$	$\forall x, y, f_{X,Y}(x,y) = f_X(x) f_Y(y)$

Agenda

- Recap – joint continuous distns
- Continuous law of total probability
- Law of total expectation

Law of total probability

Definition. Let A be an event and Y a discrete random variable. Then

$$P[A] = \sum_{y \in \Omega_Y} P(A|Y = y) p_Y(y)$$

Law of total probability

Definition. Let A be an event and Y a discrete random variable. Then

$$P[A] = \sum_{y \in \Omega_Y} P(A|Y = y) p_Y(y)$$

Definition. Let A be an event and Y a continuous random variable. Then

$$P[A] = \int_{-\infty}^{\infty} P(A|Y = y) f_Y(y) \mathrm{d}y$$

Example 1: use of law of total probability

We have a coin with unknown bias.

Specifically, the coin has probability P of heads where $P \sim \mathrm{Unif}(0,1)$.

What is $\boldsymbol{P}$(Next 10 flips are all heads)?

Definition. Let A be an event and Y a continuous random variable. Then

$$P[A] = \int_{-\infty}^{\infty} P(A|Y = y) f_Y(y) \mathrm{d}y$$

Example 1: use of law of total probability

We have a coin with unknown bias.

Specifically, the coin has probability P of heads where $P \sim \mathrm{Unif}(0,1)$.

What is $\boldsymbol{P}$(Next 10 flips are all heads)?

$$P[\text{Next 10 flips are all heads}]$$
$$= \int_{-\infty}^{\infty} P(\text{Next 10 flips are all heads} | \mathrm{P} = \mathrm{p}) f_{\mathrm{P}}(\mathrm{p}) \mathrm{dp}$$

Example 2: use of law of total probability

A certain worker is able to complete jobs efficiently until they are too tired, which happens after a number of hours T which is exponential, with parameter λ.

Jobs to be performed arrive according to a Poisson distribution with a rate of μ jobs per hour. Any job that arrives before time T (i.e., before the worker is tired) is successfully completed.

What is the probability that the worker completes k jobs before they are too tired?

Agenda

- Recap – joint continuous distns
- Continuous law of total probability
- Conditional expectation and Law of Total Expectation

Conditional Expectation and Law of Total Expectation

Suppose someone gave us $Y \sim \mathrm{Poi}(5)$ fair coins and we wanted to compute the expected number of heads X from flipping those coins.

Conditional Expectation

Definition. Let X be a discrete random variable then the **conditional expectation** of X given event A is

$$\mathbb{E}[X \mid A] = \sum_{x \in \Omega_X} x \cdot P(X = x \mid A)$$

Notes:

- Can be phrased as a "random variable version"

$$\mathbb{E}[X \mid Y = y]$$

- Linearity of expectation still applies here

$$\mathbb{E}[aX + bY + c \mid A] = a\,\mathbb{E}[X \mid A] + b\,\mathbb{E}[Y \mid A] + c$$

Law of Total Expectation

Law of Total Expectation (event version). Let X be a random variable and let events $A_1, \ldots, A_n$ partition the sample space. Then,

$$\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[X \mid A_i] \cdot P(A_i)$$

Law of Total Expectation (random variable version). Let X be a random variable and Y be a discrete random variable. Then,

$$\mathbb{E}[X] = \sum_{y \in \Omega_Y} \mathbb{E}[X \mid Y = y] \cdot P(Y = y)$$

Proof of Law of Total Expectation

Follows from Law of Total Probability and manipulating sums

$$\mathbb{E}[X] = \sum_{x \in \Omega_X} x \cdot P(X = x)$$

$$= \sum_{x \in \Omega_X} x \cdot \sum_{i=1}^{n} P(X = x \mid A_i) \cdot P(A_i)$$ (by LTP)

$$= \sum_{i=1}^{n} P(A_i) \sum_{x \in \Omega_X} x \cdot P(X = x \mid A_i)$$ (change order of sums)

$$= \sum_{i=1}^{n} P(A_i) \cdot \mathbb{E}[X \mid A_i]$$ (def of cond. expect.)

Example – Flipping a Random Number of Coins

Suppose someone gave us $Y \sim \mathrm{Poi}(5)$ fair coins and we wanted to compute the expected number of heads X from flipping those coins.

By the Law of Total Expectation

$$\mathbb{E}[X] = \sum_{i=0}^{\infty} \mathbb{E}[X \mid Y = i] \cdot P(Y = i) = \sum_{i=0}^{\infty} \frac{i}{2} \cdot P(Y = i)$$

$$= \frac{1}{2} \cdot \sum_{i=0}^{\infty} i \cdot P(Y = i)$$

$$= \frac{1}{2} \cdot \mathbb{E}[Y] = \frac{1}{2} \cdot 5 = 2.5$$

Example – Computer Failures (a familiar example)

Suppose your computer operates in a sequence of steps, and that at each step i your computer will fail with probability p (independently of other steps).

Let X be the number of steps it takes your computer to fail.

What is $\mathbb{E}[X]$?

Let Y be the indicator random variable for the event of failure in step 1

Then by LTE, $\mathbb{E}[X] = \mathbb{E}[X \mid Y = 1] \cdot P(Y = 1) + \mathbb{E}[X \mid Y = 0] \cdot P(Y = 0)$

$$= 1 \cdot p + \mathbb{E}[X \mid Y = 0] \cdot (1 - p)$$

$$= p + (1 + \mathbb{E}[X]) \cdot (1 - p)$$

since if $Y = 0$ experiment starting at step 2 looks like original experiment

Solving we get $\mathbb{E}[X] = 1/p$

Example -- Elevator rides

The number X of people who enter an elevator on the ground floor is a Poisson random variable with mean 10. If there are N floors above the ground floor, and if each person is equally likely to get off at any one of the N floors, independently of where others get off, compute the expected number of stops the elevator will make before discharging all the passengers.

Example -- Elevator rides

The number X of people who enter an elevator on the ground floor is a Poisson random variable with mean 10. If there are N floors above the ground floor, and if each person is equally likely to get off at any one of the N floors, independently of where others get off, compute the expected number of stops the elevator will make before discharging all the passengers.

$$\mathbb{E}[S] = \sum_{i \geq 0} \mathbb{E}[S \mid N = i] \cdot P(N = i)$$

$$= \sum_i \left(1 - \left(1 - \frac{1}{N}\right)^i \right) \cdot e^{-10} \cdot \frac{10^i}{i!}$$

Reference Sheet (with continuous RVs)

	Discrete	Continuous
Joint PMF/PDF	$p_{X,Y}(x,y) = P(X = x, Y = y)$	$f_{X,Y}(x,y) \neq P(X = x, Y = y)$
Joint CDF	$F_{X,Y}(x,y) = \sum_{t \le x} \sum_{s \le y} p_{X,Y}(t,s)$	$F_{X,Y}(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(t,s) ds dt$
Normalization	$\sum_{x} \sum_{y} p_{X,Y}(x,y) = 1$	$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x,y) dx dy = 1$
Marginal PMF/PDF	$p_X(x) = \sum_{y} p_{X,Y}(x,y)$	$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) dy$
Expectation	$E[g(X,Y)] = \sum_{x} \sum_{y} g(x,y) p_{X,Y}(x,y)$	$E[g(X,Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) f_{X,Y}(x,y) dx dy$
Conditional PMF/PDF	$p_{X\mid Y}(x \mid y) = \frac{p_{X,Y}(x,y)}{p_Y(y)}$	$f_{X\mid Y}(x \mid y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$
Conditional Expectation	$E[X \mid Y = y] = \sum_{x} x p_{X\mid Y}(x \mid y)$	$E[X \mid Y = y] = \int_{-\infty}^{\infty} x f_{X\mid Y}(x \mid y) dx$
Independence	$\forall x, y, p_{X,Y}(x,y) = p_X(x) p_Y(y)$	$\forall x, y, f_{X,Y}(x,y) = f_X(x) f_Y(y)$