Joint Distributions

CSE 312 Spring 26

Lecture 19

Announcements

- The midterm exam will be on Thursday May 14th starting at 6 PM. We will be split between two classrooms.
  - **Bagley 131 - if your last name starts with the letters A-P**
  - **John 102 - if your last name starts with the letters Q-Z**
- You **must** bring a photo ID with you to the exam.
- Aim to be there by 5:50pm
- More info on webpage about exams

- Wednesday in class - review

Agenda

- CLT – wrap up, continuity correction
- Joint Distributions
  - Cartesian Products
  - Joint PMFs and Joint Range
  - Marginal Distribution
  - Analogues for continuous distributions
  - LOTUS for joint distns

Summary Central Limit Theorem

$X_1, \dots, X_n$ i.i.d., each with expectation μ and variance σ^2

Define $\boldsymbol{S_n = X_1 + \cdots + X_n}$ and $\boldsymbol{\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i}$. and $\boldsymbol{Y_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}}$

	S_n	$\bar{X}$	Y_n
mean	$n\mu$	μ	0
variance	$n\sigma^2$	$\frac{\sigma^2}{n}$	1
CLT:	$\approx \mathcal{N}(n\mu, n\sigma^2)$	$\to \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$	$\to \mathcal{N}(0,1)$

Outline of how CLT is used

- Write the event you are interested in, in terms of a sum of i.i.d. random variables.
- Normalize RV to have mean 0 and standard deviation 1.
- Write event in terms of Φ, the CDF of a $\mathcal{N}(0,1)$.
- Look up in table.

Using CLT to estimate Binomial probabilities

We flip n independent coins, heads with probability $p = 0.75$.

X = # heads $\quad \mu = \mathbb{E}(X) = 0.75n \quad \sigma^2 = \text{Var}(X) = p(1-p)n = 0.1875n$

$$\mathbb{P}(X \leq 0.7n)$$

n	exact	$\mathcal{N}(\mu, \sigma^2)$ approx
10	0.4744072	0.357500327
20	0.38282735	0.302788308
50	0.25191886	0.207108089
100	0.14954105	0.124106539
200	0.06247223	0.051235217
1000	0.00019359	0.000130365

Example – Naive Approximation

Fair coin flipped (independently) **40** times. Probability of **20** or **21** heads?

Exact. $$\mathbb{P}(X \in \{20,21\}) = \left[\binom{40}{20} + \binom{40}{21}\right]\left(\frac{1}{2}\right)^{40} \approx \boxed{0.2448}$$

Approx. $X = \#$ heads $\quad \mu = \mathbb{E}(X) = 0.5n = 20 \quad \sigma^2 = \mathrm{Var}(X) = 0.25n = 10$

$$\mathbb{P}(20 \leq X \leq 21) = \Phi\left(\frac{20-20}{\sqrt{10}} \leq \frac{X-20}{\sqrt{10}} \leq \frac{21-20}{\sqrt{10}}\right)$$

$$\approx \Phi\left(0 \leq \frac{X-20}{\sqrt{10}} \leq 0.32\right)$$

$$= \Phi(0.32) - \Phi(0) \approx \boxed{0.1241}$$

😢

Example – Even Worse Approximation

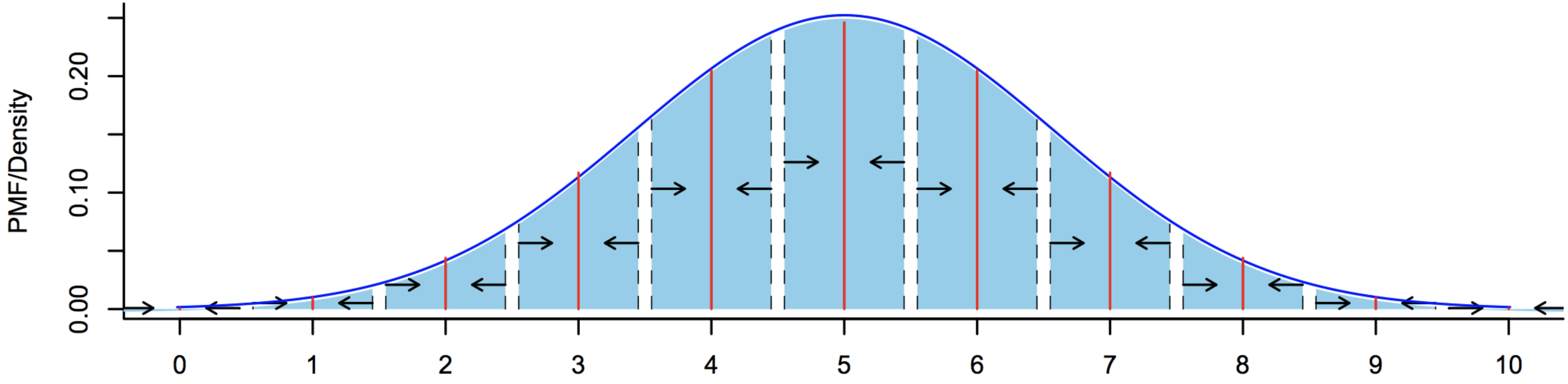
Fair coin flipped (independently) **40** times. Probability of **20** heads?

Exact. $\mathbb{P}(X = 20) = \binom{40}{20}\left(\frac{1}{2}\right)^{40} \approx \boxed{0.1254}$

Approx. $\mathbb{P}(20 \leq X \leq 20) = 0$ 😢

Solution – Continuity Correction

Probability estimate for i: Probability for all x that round to i!



To estimate probability that discrete RV lands in (integer) interval $\{a, \ldots, b\}$, compute probability continuous approximation lands in interval $[a - \frac{1}{2}, b + \frac{1}{2}]$

Example – Continuity Correction

Fair coin flipped (independently) **40** times. Probability of **20** or **21** heads?

Exact. $$\mathbb{P}(X \in \{20,21\}) = \left[\binom{40}{20} + \binom{40}{21}\right]\left(\frac{1}{2}\right)^{40} \approx \boxed{0.2448}$$

Approx. X = # heads $\quad \mu = \mathbb{E}(X) = 0.5n = 20 \quad \sigma^2 = \mathrm{Var}(X) = 0.25n = 10$

$$\mathbb{P}(19.5 \leq X \leq 21.5) = \Phi\left(\frac{19.5-20}{\sqrt{10}} \leq \frac{X-20}{\sqrt{10}} \leq \frac{21.5-20}{\sqrt{10}}\right)$$

$$\approx \Phi\left(-0.16 \leq \frac{X-20}{\sqrt{10}} \leq 0.47\right)$$

$$= \Phi(0.47) - \Phi(-0.16) \approx \boxed{0.2452}$$

Example – Continuity Correction

Fair coin flipped (independently) **40** times. Probability of **20** heads?

Exact. $\mathbb{P}(X = 20) = \binom{40}{20}\left(\frac{1}{2}\right)^{40} \approx \boxed{0.1254}$

Approx. $\mathbb{P}(19.5 \le X \le 20.5) = \Phi\left(\frac{19.5 - 20}{\sqrt{10}} \le \frac{X - 20}{\sqrt{10}} \le \frac{20.5 - 20}{\sqrt{10}}\right)$

$$\approx \Phi\left(-0.16 \le \frac{X - 20}{\sqrt{10}} \le 0.16\right)$$

$$= \Phi(0.16) - \Phi(-0.16) \approx \boxed{0.1272}$$

Agenda

- CLT and Polling
- Joint Distributions
  - Cartesian Products
  - Joint PMFs and Joint Range
  - Marginal Distribution
  - Analogues for continuous distributions
  - LOTUS for joint distns

Why joint distributions?

- Given all of its user's ratings for different movies, and any preferences you have expressed, Netflix wants to recommend a new movie for you.
- Given a large amount of medical data correlating symptoms and personal history with diseases, predict what is ailing a person with a particular medical history and set of symptoms.
- Given current traffic, pedestrian locations, weather, lights, etc. decide whether a self-driving car should slow down or come to a stop

Jointly distributed random variables

X	Y
Grade on exam	Amount of sleep the night before
Performance of Microsoft stock	Performance of Amazon stock
Grade of person A on exam in 312	Grade of person B on exam in 332
Number of job interviews to get a job	State of the economy

X_1 blood pressure
X_2 temperature
X_3 blood glucose
X_4 kidney function

Review Cartesian Product

Definition. Let A and B be sets. The **Cartesian product** of A and B is denoted

$$A \times B = \{(a, b) : a \in A, b \in B\}$$

Example.

$$\{1,2,3\} \times \{4, 5\} = \{(1, 4), (1,5), (2,4), (2,5), (3, 4), (3,5)\}$$

If A and B are finite sets, then $|A \times B| = |A| \cdot |B|$.

The sets don't need to be finite! You can have $\mathbb{R} \times \mathbb{R}$ (often denoted $\mathbb{R}^2$)

Joint PMFs and Joint Range

Definition. Let X and Y be discrete random variables. The **Joint PMF** of X and Y is

$$p_{X,Y}(a,b) = P(X = a, Y = b)$$

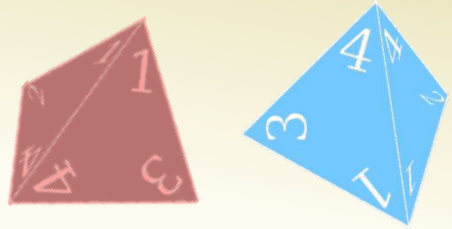
Definition. The **joint range** of $p_{X,Y}$ is

$$\Omega_{X,Y} = \{(c,d) : p_{X,Y}(c,d) > 0\} \subseteq \Omega_X \times \Omega_Y$$

Note that

$$\sum_{(s,t)\in\Omega_{X,Y}} p_{X,Y}(s,t) = 1$$

Example – Weird Dice



Suppose I roll two fair 4-sided die independently. Let X be the value of the first die, and Y be the value of the second die.

$\Omega_X = \{1,2,3,4\}$ and $\Omega_Y = \{1,2,3,4\}$

In this problem, the joint PMF is if

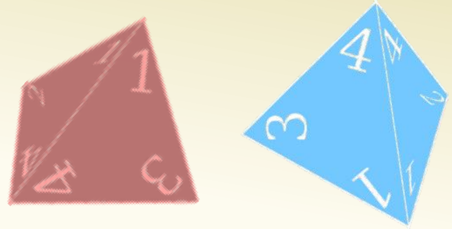
$$p_{X,Y}(x,y) = \begin{cases} 1/16 & \text{if } x,y \in \Omega_{X,Y} \\ 0 & \text{otherwise} \end{cases}$$

X\Y	1	2	3	4
1	1/16	1/16	1/16	1/16
2	1/16	1/16	1/16	1/16
3	1/16	1/16	1/16	1/16
4	1/16	1/16	1/16	1/16

and the joint range is (since all combinations have non-zero probability)

$\Omega_{X,Y} = \Omega_X \times \Omega_Y$

Example – Weirder Dice



Suppose I roll two fair 4-sided die independently. Let X be the value of the first die, and Y be the value of the second die. Let $U = \min(X, Y)$ and $W = \max(X, Y)$

$\Omega_U = \{1,2,3,4\}$ and $\Omega_W = \{1,2,3,4\}$

$$\Omega_{U,W} = \{(u, w) \in \Omega_U \times \Omega_W : u \leq w\} \neq \Omega_U \times \Omega_W$$

What is $p_{U,W}(1, 3) = P(U = 1, W = 3)$?

U\W	1	2	3	4
1				
2				
3				
4				

Example – Weirder Dice

Suppose I roll two fair 4-sided die independently. Let X be the value of the first die, and Y be the value of the second die. Let $U = \min(X, Y)$ and $W = \max(X, Y)$

$\Omega_U = \{1,2,3,4\}$ and $\Omega_W = \{1,2,3,4\}$

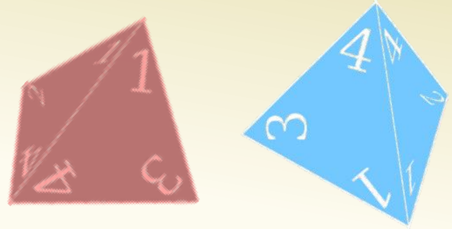
$$\Omega_{U,W} = \{(u, w) \in \Omega_U \times \Omega_W : u \leq w\} \neq \Omega_U \times \Omega_W$$

The joint PMF $p_{U,W}(u, w) = P(U = u, W = w)$ is

$$p_{U,W}(u, w) = \begin{cases} 2/16 & \text{if } (u, w) \in \Omega_U \times \Omega_W \text{ where } w > u \\ 1/16 & \text{if } (u, w) \in \Omega_U \times \Omega_W \text{ where } w = u \\ 0 & \text{otherwise} \end{cases}$$

U\W	1	2	3	4
1	1/16	2/16	2/16	2/16
2	0	1/16	2/16	2/16
3	0	0	1/16	2/16
4	0	0	0	1/16

Example – Weirder Dice

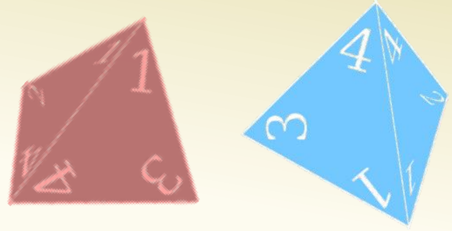


Suppose I roll two fair 4-sided die independently. Let X be the value of the first die, and Y be the value of the second die. Let $U = \min(X, Y)$ and $W = \max(X, Y)$

Suppose we didn't know how to compute $P(U = u)$ directly. Can we figure it out if we know $p_{U,W}(u, w)$?

U\W	1	2	3	4
1	1/16	2/16	2/16	2/16
2	0	1/16	2/16	2/16
3	0	0	1/16	2/16
4	0	0	0	1/16

Example – Weirder Dice



Suppose I roll two fair 4-sided die independently. Let X be the value of the first die, and Y be the value of the second die. Let $U = \min(X, Y)$ and $W = \max(X, Y)$

Suppose we didn't know how to compute $P(U = u)$ directly. Can we figure it out if we know $p_{U,W}(u, w)$?

Just apply LTP over the possible values of W:

$$p_U(1) = 7/16$$

$$p_U(2) = 5/16$$

$$p_U(3) = 3/16$$

$$p_U(4) = 1/16$$

U\W	1	2	3	4
1	1/16	2/16	2/16	2/16
2	0	1/16	2/16	2/16
3	0	0	1/16	2/16
4	0	0	0	1/16

Marginal PMF

Definition. Let X and Y be discrete random variables and $p_{X,Y}(a,b)$ their joint PMF. The **marginal PMF** of X

$$p_X(a) = \sum_{b \in \Omega_Y} p_{X,Y}(a,b)$$

Similarly, $p_Y(b) = \sum_{a \in \Omega_X} p_{X,Y}(a,b)$

Joint Expectation

Definition. Let X and Y be discrete random variables and $p_{X,Y}(a,b)$ their joint PMF. The **expectation** of some function $g(X,Y)$ is

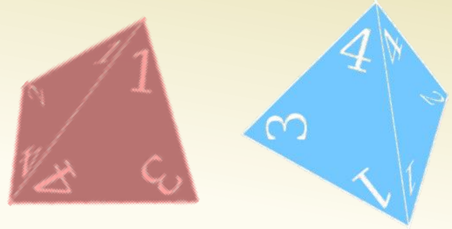
$$\mathbb{E}[g(X,Y)] = \sum_{a \in \Omega_X} \sum_{b \in \Omega_Y} g(a,b) \cdot p_{X,Y}(a,b)$$

Independence and joint distributions

Definition. Discrete random variables X and Y are **independent** iff

- $p_{X,Y}(x, y) = p_X(x) \cdot p_Y(y)$ for all $x \in \Omega_X, y \in \Omega_Y$

Example – Weirder Dice



Suppose I roll two fair 4-sided die independently. Let X be the value of the first die, and Y be the value of the second die. Let $U = \min(X, Y)$ and $W = \max(X, Y)$

$\Omega_U = \{1,2,3,4\}$ and $\Omega_W = \{1,2,3,4\}$

$$\Omega_{U,W} = \{(u, w) \in \Omega_U \times \Omega_W : u \leq w\} \neq \Omega_U \times \Omega_W$$

Are U and W independent?

U\W	1	2	3	4
1	1/16	2/16	2/16	2/16
2	0	1/16	2/16	2/16
3	0	0	1/16	2/16
4	0	0	0	1/16

A quick check for independence

The check: $\Omega_X \times \Omega_Y$ must equal $\Omega_{X,Y}$ for independence.

Suppose that there is some $(a, b) \in \Omega_X \times \Omega_Y$, but not in $\Omega_{X,Y}$.

Then: $p_{X,Y}(a, b) = 0$

But: $p_X(a) > 0$ and $p_Y(b) > 0$

But beware, the converse is not true: $\Omega_X \times \Omega_Y = \Omega_{X,Y}$ does not imply independence!

Continuous distributions on $\mathbb{R} \times \mathbb{R}$

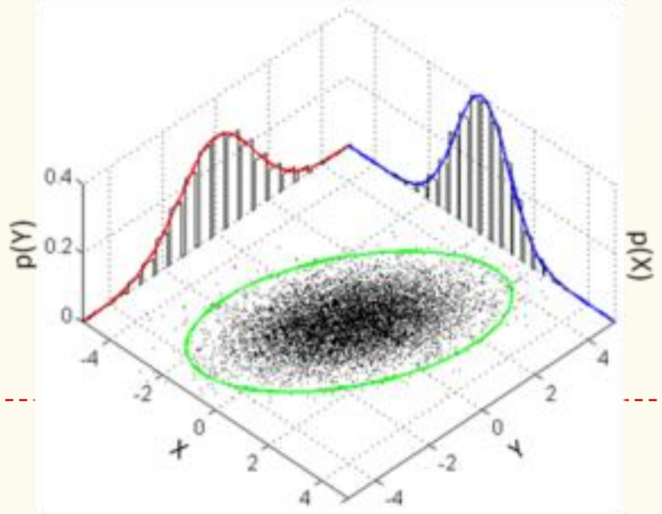
Definition. The **joint probability density function (PDF)** of continuous random variables X and Y is a function $f_{X,Y}$ defined on $\mathbb{R} \times \mathbb{R}$ such that

- $f_{X,Y}(x, y) \geq 0$ for all $x, y \in \mathbb{R}$
- $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x, y) \mathrm{d}x \, \mathrm{d}y = 1$

for $A \subseteq \mathbb{R} \times \mathbb{R}$ the probability that $(X, Y) \in A$ is $\iint_A f_{X,Y}(x, y) \, \mathrm{d}x\mathrm{d}y$

The **(marginal) PDFs** f_X and f_Y are given by

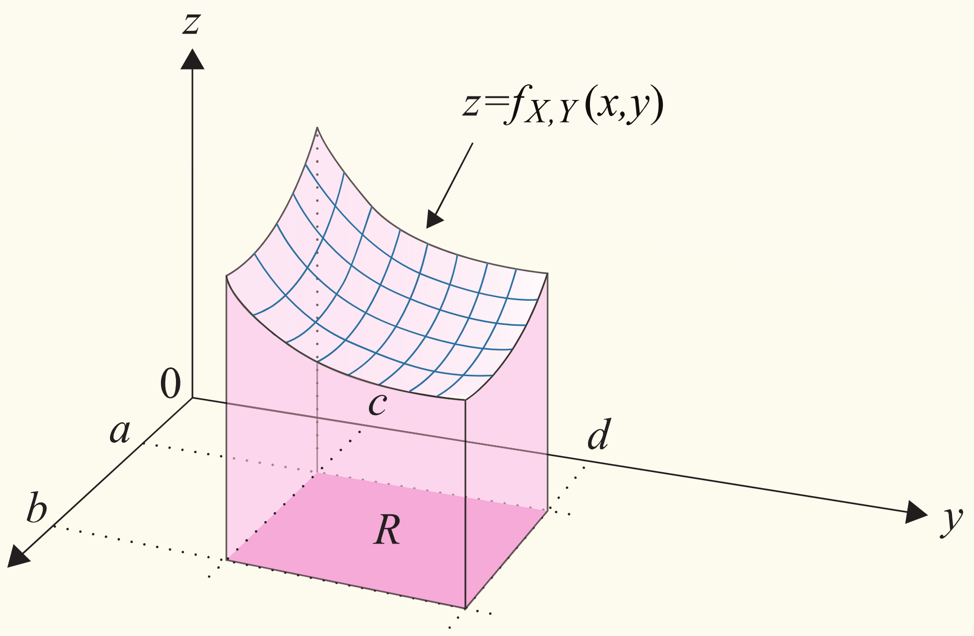
- $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) \, \mathrm{d}y$
- $f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) \, \mathrm{d}x$



Joint Densities

Volume under the curve equals:

$$\int_c^d \int_a^b f_{X,Y}(x,y)\,dx\,dy = \mathbb{P}(a \le X \le b \ \& \ c \le Y \le d)$$



Independence and joint distributions

Definition. Continuous random variables X and Y are **independent** iff

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

Example

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

Q: What is $\boldsymbol{E}[X]$?

Example

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

Q What is $\boldsymbol{E}[X]$?

A:

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy = \int_0^1 (x+y)dy = x + \frac{1}{2}$$

$$\boldsymbol{E}[X] = \int_{-\infty}^{\infty} f_X(x) \cdot x dx = \int_0^1 \left(x + \frac{1}{2}\right) \cdot x\, dx = \frac{7}{12}$$

Example

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

Are X and Y independent?

Example

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

Q: Are *X* and *Y* independent?

A: part (b)

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy = x + \frac{1}{2}$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dx = y + \frac{1}{2}$$

Clearly, $f_{X,Y}(x,y) \neq f_X(x) \cdot f_Y(y)$

Example

$$f_{X,Y}(x,y) = \begin{cases} 4xy & \text{if } 0 \leq x, y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Q: Are *X* and *Y* independent?

Example

$$f_{X,Y}(x,y) = \begin{cases} 4xy & \text{if } 0 \le x, y \le 1 \\ 0 & \text{otherwise} \end{cases}$$

Q: Are *X* and *Y* independent?

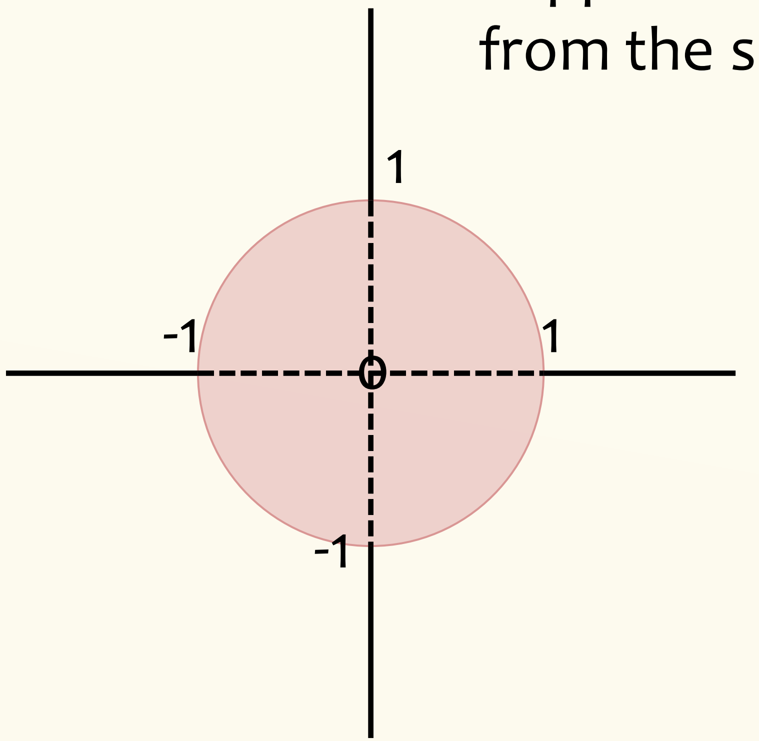
A:

$$f_X(x) = \int_0^1 4xy\,dy = 2x$$

$$f_Y(y) = \int_0^1 4xy\,dx = 2y$$

Clearly, $f_{X,Y}(x,y) = f_X(x) \cdot f_Y(y)$

Example – Uniform distribution on a unit disk

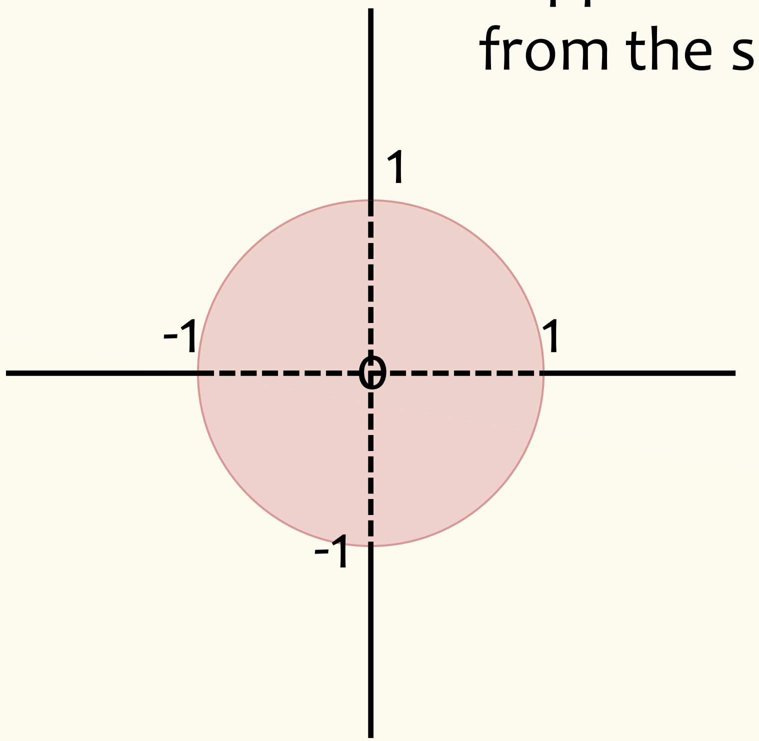


Suppose that a pair of random variables (X, Y) is chosen uniformly from the set of real points (x, y) such that $x^2 + y^2 \le 1$

What is the joint density?

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

Example – Uniform distribution on a unit disk



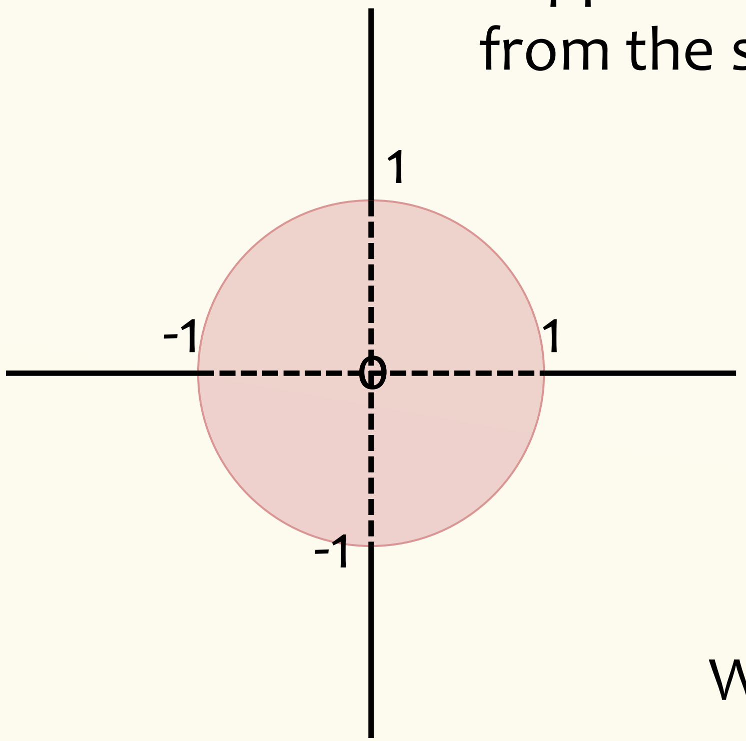
Suppose that a pair of random variables (X, Y) is chosen uniformly from the set of real points (x, y) such that $x^2 + y^2 \leq 1$

This is a disk of radius 1 which has area π

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

Are X and Y independent?

Example – Uniform distribution on a unit disk



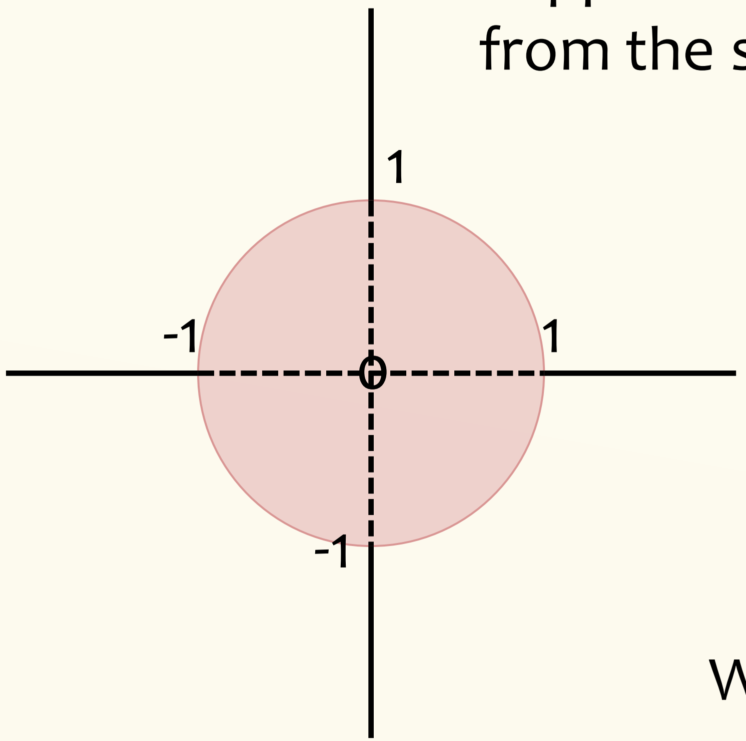
Suppose that a pair of random variabes (X, Y) is chosen uniformly from the set of real points (x, y) such that $x^2 + y^2 \leq 1$

This is a disk of radius 1 which has area π

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

What is $f_X(x)$?

Example – Uniform distribution on a unit disk



Suppose that a pair of random 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

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

What is $f_X(x)$?

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

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

LOTUS

Definition. Let X and Y be continuous random variables and $f_{X,Y}(x,y)$ their joint PMF. The **expectation** of some function $g(X,Y)$ is

$$\mathbb{E}[g(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y) \cdot f_{X,Y}(x,y)\mathrm{d}x\,\mathrm{d}y$$

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	$\forall x, y, p_{X,Y}(x,y) = p_X(x) p_Y(y)$	$\forall x, y, f_{X,Y}(x,y) = f_X(x) f_Y(y)$