JOINT DISTRIBUTIONS

ANNA KARLIN Most slides by Alex Tsun

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 bunch 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 self-driving car should slow down or come to a stop?

5.1 JOINT DISCRETE DISTRIBUTIONS



AGENDA

- CARTESIAN PRODUCTS OF SETS
- JOINT PMFS AND EXPECTATION
- MARGINAL PMFS

CARTESIAN PRODUCT OF SETS

Cartesian Product: Let A, B be sets. The Cartesian product of A and B is denoted

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

A small example:

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

Another example: The xy-plane (2D space) is denoted

 $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R} = \{(x, y) : x \in \mathbb{R}, y \in \mathbb{R}\}$

If A, B are finite sets, then $|A \times B| = |A| \cdot |B|$ by the product rule of counting.



EXAMPLE: RATINGS

for $x, y \in \Omega_{X,Y}$.

Let X be the value of the blue die, and Y the value of the red die. Specify

 $\Omega_X = \{1,2,3,4\} \qquad \Omega_Y = \{1,2,3,4\}$ $\Omega_{X,Y} = \Omega_X \times \Omega_Y$ Specify the joint PMF $p_{X,Y}(x,y) = P(X = x, Y = y)$

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

 $0 = \{1, 2, 3, 4\}$

$M_X = \{1, 2, 3, 4\}$ $M_Y = \{1, 2, 3, 4\}$	X\Y	1
$\Omega_{X,Y} = \Omega_X \times \Omega_Y$	_	1/16
Specify the joint PMF $p_{X,Y}(x, y) = P(X = x, Y = y)$ for $x, y \in \Omega_{X,Y}$.	2	1/16

 $0 = \{1, 2, 3, 4\}$

$p_{X,Y}(x,y) =$	(1/16,	$x, y \in \Omega_{X,Y}$
$p_{X,Y}(x,y) = $	0,	otherwise

X\Y	1	2	3	4
]	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



JOINT PMFS AND EXPECTATION

Joint PMFs: Let X, Y be discrete random variables. The joint PMF of X and Y is

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

The joint range is

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

Note that

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

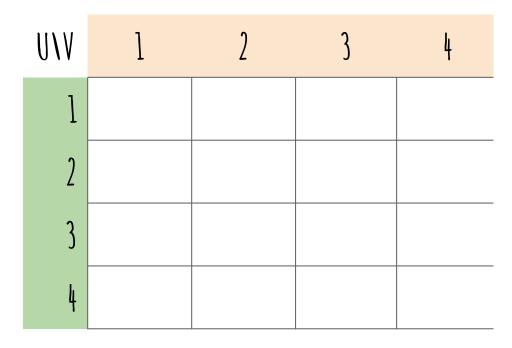
Suppose I roll two fair 4-sided die independently. Let X be the value of the blue die, and Y the value of the red die. Let $U = \min \{X, Y\}$ and $V = \max \{X, Y\}$.

 $\Omega_U = \{1, 2, 3, 4\} \qquad \qquad \Omega_V = \{1, 2, 3, 4\}$

 $\Omega_{U,V} = \{(u,v) \in \Omega_U \times \Omega_V : u \le v\} \neq \Omega_U \times \Omega_V$

Specify the joint PMF $p_{U,V}(u, v) = P(U = u, V = v)$ for $u, v \in \Omega_{U,V}$.





Suppose I roll two fair 4-sided die independently. Let X be the value of the blue die, and Y the value of the red die. Let $U = \min \{X, Y\}$ and $V = \max \{X, Y\}$.

 $\Omega_U = \{1,2,3,4\} \qquad \Omega_V = \{1,2,3,4\}$ $\Omega_{U,V} = \{(u,v) \in \Omega_U \times \Omega_V : u \le v\} \ne \Omega_U \times \Omega_V$

Specify the joint PMF $p_{U,V}(u, v) = P(U = u, V = v)$ for $u, v \in \Omega_{U,V}$.

	(2/16,	$u, v \in \Omega_U \times \Omega_V,$	v > u
$p_{U,V}(u,v) =$	1/16,	$u, v \in \Omega_U \times \Omega_V,$	v = u
	(0,	otherwise	

UNV	1	2	3	4
1	1/16	2/16	2/16	2/16
2	Q	1/16	2/16	2/16
3	Q	Q	1/16	2/16
4	Q	0	Û	1/16



Suppose I roll two fair 4-sided die independently. Let X be the value of the blue die, and Y the value of the red die. Let $U = \min \{X, Y\}$ and $V = \max \{X, Y\}$.

What is
$$p_U(u)$$
 for $u \in \Omega_U$?

$$p_U(u) = \begin{cases} u = 1\\ u = 2\\ u = 3\\ u = 4 \end{cases}$$

UVV	1	2	3	4
1	1/16	2/16	2/16	2/16
2	Û	1/16	2/16	2/16
3	Û	Û	1/16	2/16
4	0	Û	Û	1/16



Suppose I roll two fair 4-sided die independently. Let X be the value of the blue die, and Y the value of the red die. Let $U = \min \{X, Y\}$ and $V = \max \{X, Y\}$.

What is
$$p_U(u)$$
 for $u \in \Omega_U$?

$$p_U(u) = \begin{cases} 7/16, \ u = 1\\ 5/16, \ u = 2\\ 3/16, \ u = 3\\ 1/16, \ u = 4 \end{cases}$$

UVV	1	2	3	4
1	1/16	2/16	2/16	2/16
2	0	1/16	2/16	2/16
3	Û	Û	1/16	2/16
4	Û	0	0	1/16





Suppose I roll two fair 4-sided die independently. Let X be the value of the blue die, and Y the value of the red die. Let $U = \min \{X, Y\}$ and $V = \max \{X, Y\}$.

UNV	1	2	3	4
1	1/16	2/16	2/16	2/16
2	0	1/16	2/16	2/16
3	Q	Û	1/16	2/16
4	Q	0	0	1/16

JOINT PMFS AND EXPECTATION

Joint PMFs: Let X, Y be discrete random variables. The joint PMF of X and Y is

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

The joint range is

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

Note that

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

If $g: \mathbb{R}^2 \to \mathbb{R}$ is a function, then

$$E[g(X,Y)] = \sum_{x} \sum_{y} g(x,y) p_{X,Y}(x,y)$$

MARGINAL PMFS

<u>Marginal PMFs</u>: Let X, Y be discrete random variables. The marginal PMF of X is

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

MARGINAL PMFS

<u>Marginal PMFs</u>: Let X, Y be discrete random variables. The marginal PMF of X is

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

Similarly, the marginal PMF of Y is

$$p_Y(d) = \sum_{c \in \Omega_X} p_{X,Y}(c,d)$$

MARGINAL PMFS

Marginal PMFs: Let X, Y be discrete random variables. The marginal PMF of X is

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

Similarly, the marginal PMF of Y is

$$p_Y(d) = \sum_{c \in \Omega_X} p_{X,Y}(c,d)$$

(Extension) If Z is also a discrete random variable, then the marginal PMF of Z is

$$p_Z(z) = \sum_{x \in \Omega_X} \sum_{y \in \Omega_Y} p_{X,Y,Z}(x, y, z)$$

Independence (DRVs): Discrete random variables X, Y are independent, written $X \perp Y$, if for all $x \in \Omega_X$ and $y \in \Omega_Y$,

 $p_{X,Y}(x,y) = p_X(x)p_Y(y)$

Recall $\Omega_{X,Y} = \{(x, y): p_{X,Y}(x, y) > 0\} \subseteq \Omega_X \times \Omega_Y$. A necessary but not sufficient condition for independence is that $\Omega_{X,Y} = \Omega_X \times \Omega_Y$. That is, if $\Omega_{X,Y} \neq \Omega_X \times \Omega_Y$, then X and Y cannot be independent, but if $\Omega_{X,Y} = \Omega_X \times \Omega_Y$, then we have to check the condition.

This is because if 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 $p_Y(b) > 0$, violating independence.

VARIANCE ADDS FOR INDEPENDENT RVS

If X, Y are independent random variables $X \perp Y$, then

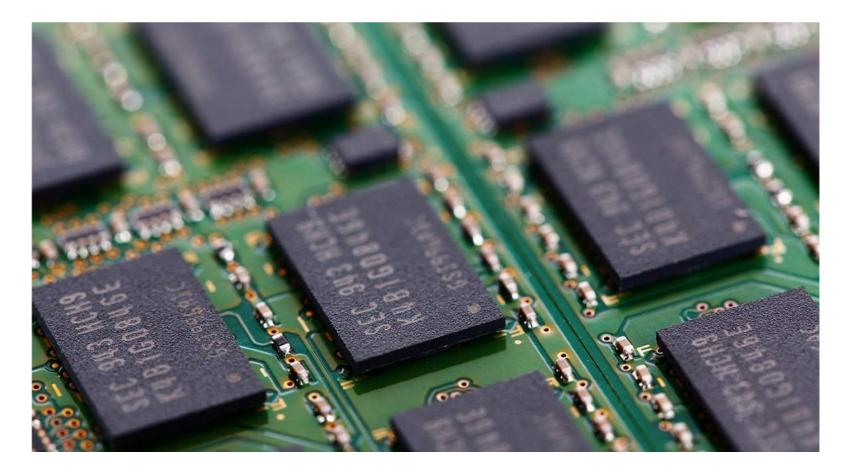
Var(X + Y) = Var(X) + Var(Y)

This property relies on the fact that they are independent, whereas linearity of expectation always holds, regardless. If $a, b, c \in \mathbb{R}$ are scalars, then

$$Var(aX + bY + c) = a^{2}Var(X) + b^{2}Var(Y)$$

If $X \perp Y$, then E[XY] = E[X]E[Y].

RANDOM PICTURE



5.2 JOINT CONTINUOUS DISTRIBUTIONS



AGENDA

- JOINT PDFS AND EXPECTATION
- MARGINAL PDFS
- INDEPENDENCE
- MULTIVARIATE: FROM DISCRETE TO CONTINUOUS

JOINT PDFS AND EXPECTATION

Joint PDFs: Let X, Y be continuous random variables. The joint PDF of X and Y is

 $f_{X,Y}(a,b)$

The joint range is

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

JOINT PDFS AND EXPECTATION

Joint PDFs: Let X, Y be continuous random variables. The joint PDF of X and Y is

 $f_{X,Y}(a,b)$

The joint range is

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

Note that

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}f_{X,Y}(u,v)dudv=1$$

JOINT PDFS AND EXPECTATION

Joint PDFs: Let X, Y be continuous random variables. The joint PDF of X and Y is

 $f_{X,Y}(a,b)$

The joint range is

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

Note that

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}f_{X,Y}(u,v)dudv=1$$

If $g: \mathbb{R}^2 \to \mathbb{R}$ is a function, then

$$E[g(X,Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(s,t) f_{X,Y}(s,t) ds dt$$

MARGINAL PDFS

Marginal PDFs: Let X, Y be continuous random variables. The marginal PDF of X is

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

MARGINAL PDFS

Marginal PDFs: Let X, Y be continuous random variables. The marginal PDF of X is

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

Similarly, the marginal PDF of Y is

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

MARGINAL PDFS

Marginal PDFs: Let X, Y be continuous random variables. The marginal PDF of X is

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

Similarly, the marginal PDF of Y is

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

(Extension) If Z is also a continuous random variable, then the marginal PDF of Z is

$$f_Z(z) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y,Z}(x,y,z) dx dy$$

Independence (CRVs): Continuous random variables X, Y are independent, written $X \perp Y$, if for all $x \in \Omega_X$ and $y \in \Omega_Y$,

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

Independence (CRVs): Continuous random variables X, Y are independent, written $X \perp Y$, if for all $x \in \Omega_X$ and $y \in \Omega_Y$,

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

Recall $\Omega_{X,Y} = \{(x, y): f_{X,Y}(x, y) > 0\} \subseteq \Omega_X \times \Omega_Y$. A necessary but not sufficient condition for independence is that $\Omega_{X,Y} = \Omega_X \times \Omega_Y$. That is, if $\Omega_{X,Y} \neq \Omega_X \times \Omega_Y$, then X and Y cannot be independent, but if $\Omega_{X,Y} = \Omega_X \times \Omega_Y$, then we have to check the condition.

Independence (CRVs): Continuous random variables X, Y are independent, written $X \perp Y$, if for all $x \in \Omega_X$ and $y \in \Omega_Y$,

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

Recall $\Omega_{X,Y} = \{(x, y): f_{X,Y}(x, y) > 0\} \subseteq \Omega_X \times \Omega_Y$. A necessary but not sufficient condition for independence is that $\Omega_{X,Y} = \Omega_X \times \Omega_Y$. That is, if $\Omega_{X,Y} \neq \Omega_X \times \Omega_Y$, then X and Y cannot be independent, but if $\Omega_{X,Y} = \Omega_X \times \Omega_Y$, then we have to check the condition.

This is because if there is some $(a, b) \in \Omega_X \times \Omega_Y$ but not in $\Omega_{X,Y}$, then $f_{X,Y}(a, b) = 0$ but $f_X(a) > 0$ and $f_Y(b) > 0$, violating independence.



Suppose (X, Y) are jointly and uniformly distributed on the circle of radius R centered at the origin (e.g., a dart throw). Set up but **DO NOT EVALUATE** any of your answers. Take care in setting up the limits of integration.



Suppose (X, Y) are jointly and uniformly distributed on the circle of radius R centered at the origin (e.g., a dart throw). Set up but **DO NOT EVALUATE** any of your answers. Take care in setting up the limits of integration.

a. Find and sketch the joint range $\Omega_{X,Y}$.



Suppose (X, Y) are jointly and uniformly distributed on the circle of radius R centered at the origin (e.g., a dart throw). Set up but **DO NOT EVALUATE** any of your answers. Take care in setting up the limits of integration.

a. Find and sketch the joint range $\Omega_{X,Y}$.

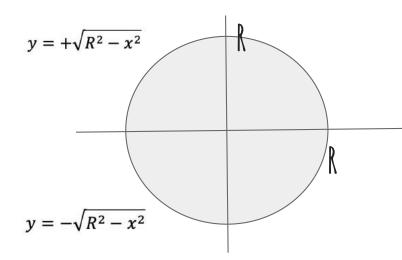
 $\Omega_{X,Y} = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \le R^2\}$ $y = +\sqrt{R^2 - x^2}$



Suppose (X, Y) are jointly and uniformly distributed on the circle of radius R centered at the origin (e.g., a dart throw). Set up but **DO NOT EVALUATE** any of your answers. Take care in setting up the limits of integration.

a. Find and sketch the joint range $\Omega_{X,Y}$.

 $\Omega_{X,Y} = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \le R^2\}$



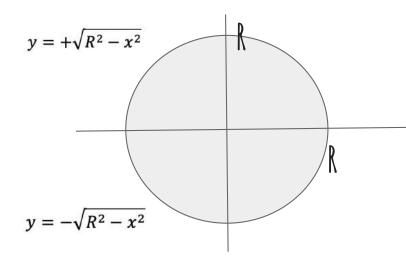
b. Write an expression for the joint PDF $f_{X,Y}(x, y)$ and carefully define it for all $x, y \in \mathbb{R}$.



Suppose (X, Y) are jointly and uniformly distributed on the circle of radius R centered at the origin (e.g., a dart throw). Set up but **DO NOT EVALUATE** any of your answers. Take care in setting up the limits of integration.

a. Find and sketch the joint range $\Omega_{X,Y}$.

 $\Omega_{X,Y} = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \le R^2\}$



b. Write an expression for the joint PDF $f_{X,Y}(x, y)$ and carefully define it for all $x, y \in \mathbb{R}$.

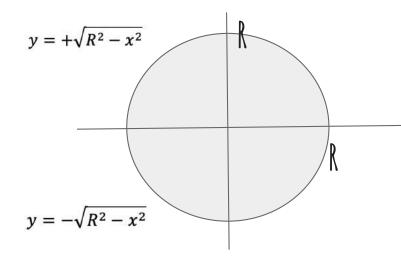
 $f_{X,Y}(x,y) = \begin{cases} \frac{1}{\pi R^2}, & x, y \in \Omega_{X,Y} \\ 0, & \text{otherwise} \end{cases}$



Suppose (X, Y) are jointly and uniformly distributed on the circle of radius R centered at the origin (e.g., a dart throw). Set up but **DO NOT EVALUATE** any of your answers. Take care in setting up the limits of integration.

a. Find and sketch the joint range $\Omega_{X,Y}$.

 $\Omega_{X,Y} = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \le R^2\}$



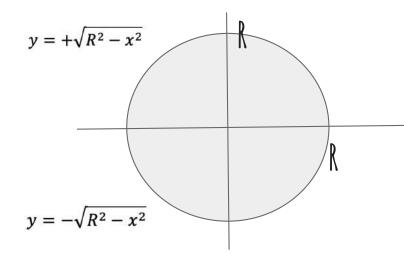
c. Find Ω_X and write an expression that we could evaluate to find $f_X(x)$.



Suppose (X, Y) are jointly and uniformly distributed on the circle of radius R centered at the origin (e.g., a dart throw). Set up but **DO NOT EVALUATE** any of your answers. Take care in setting up the limits of integration.

a. Find and sketch the joint range $\Omega_{X,Y}$.

 $\Omega_{X,Y} = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \le R^2\}$



c. Find Ω_X and write an expression that we could evaluate to find $f_X(x)$.

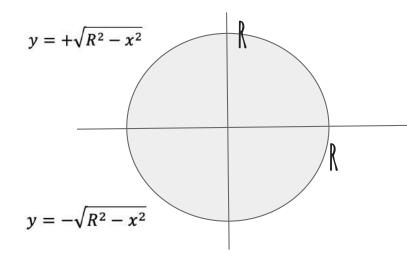
 $\Omega_X = [-R, R]$



Suppose (X, Y) are jointly and uniformly distributed on the circle of radius R centered at the origin (e.g., a dart throw). Set up but **DO NOT EVALUATE** any of your answers. Take care in setting up the limits of integration.

a. Find and sketch the joint range $\Omega_{X,Y}$.

 $\Omega_{X,Y} = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \le R^2\}$



c. Find Ω_X and write an expression that we could evaluate to find $f_X(x)$.

 $\Omega_X = [-R, R]$

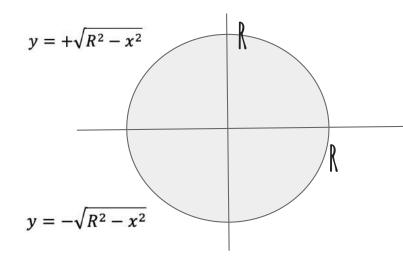
$$f_X(x) = \int_{-\sqrt{R^2 - x^2}}^{+\sqrt{R^2 - x^2}} f_{X,Y}(x, y) dy$$



Suppose (X, Y) are jointly and uniformly distributed on the circle of radius R centered at the origin (e.g., a dart throw). Set up but **DO NOT EVALUATE** any of your answers. Take care in setting up the limits of integration.

a. Find and sketch the joint range $\Omega_{X,Y}$.

 $\Omega_{X,Y} = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \le R^2\}$



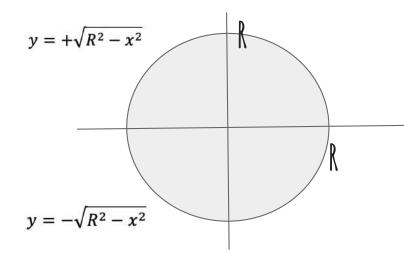
d. Let Z the distance from the center that the dart falls. Find Ω_Z and write an expression for E[Z].



Suppose (X, Y) are jointly and uniformly distributed on the circle of radius R centered at the origin (e.g., a dart throw). Set up but **DO NOT EVALUATE** any of your answers. Take care in setting up the limits of integration.

a. Find and sketch the joint range $\Omega_{X,Y}$.

 $\Omega_{X,Y} = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \le R^2\}$



d. Let Z the distance from the center that the dart falls. Find Ω_Z and write an expression for E[Z].

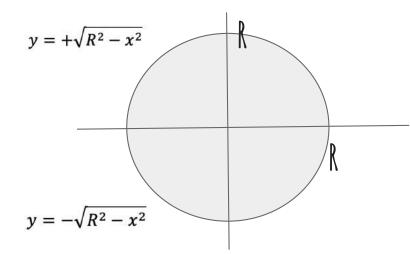
 $Z = \sqrt{X^2 + Y^2}$ $\Omega_Z = [0, R]$



Suppose (X, Y) are jointly and uniformly distributed on the circle of radius R centered at the origin (e.g., a dart throw). Set up but **DO NOT EVALUATE** any of your answers. Take care in setting up the limits of integration.

a. Find and sketch the joint range $\Omega_{X,Y}$.

 $\Omega_{X,Y} = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \le R^2\}$



d. Let Z the distance from the center that the dart falls. Find Ω_Z and write an expression for E[Z].

$$Z = \sqrt{X^2 + Y^2}$$
$$\Omega_Z = [0, R]$$

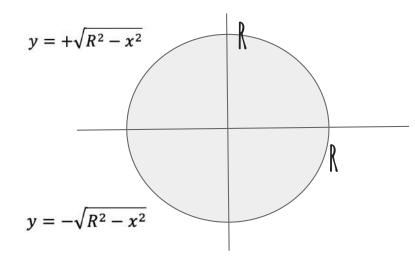
 $E[Z] = \int_{-R}^{R} \int_{-\sqrt{R^2 - x^2}}^{+\sqrt{R^2 - x^2}} \sqrt{x^2 + y^2} f_{X,Y}(x, y) dy dx$



Suppose (X, Y) are jointly and uniformly distributed on the circle of radius R centered at the origin (e.g., a dart throw). Set up but **DO NOT EVALUATE** any of your answers. Take care in setting up the limits of integration.

a. Find and sketch the joint range $\Omega_{X,Y}$.

 $\Omega_{X,Y} = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \le R^2\}$



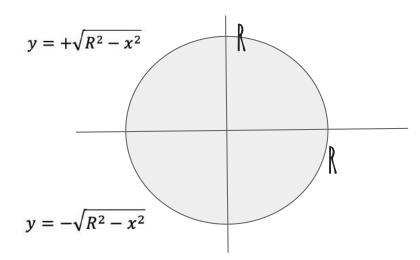
e. Are X and Y independent?



Suppose (X, Y) are jointly and uniformly distributed on the circle of radius R centered at the origin (e.g., a dart throw). Set up but **DO NOT EVALUATE** any of your answers. Take care in setting up the limits of integration.

a. Find and sketch the joint range $\Omega_{X,Y}$.

 $\Omega_{X,Y} = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \le R^2\}$



e. Are X and Y independent?

No, $\Omega_{X,Y} \neq \Omega_X \times \Omega_Y$ (not a rectangle).

 $0 = f_{X,Y}(0.99R, 0.99R) \neq f_X(0.99R) f_Y(0.99R) > 0$

RANDOM PICTURE



Consider a continuous joint distribution, (X, Y), where $X \in [0,1]$ is the proportion of the time until the midterm that you actually study for it, and $Y \in [0,1]$ is your percentage score on the exam. Set up but **DO NOT EVALUATE** any of your answers. Take care in setting up the limits of integration. The joint PDF is:

 $f_{X,Y}(x,y) = \begin{cases} ce^{-(y-x)}, & x,y \in [0,1] \text{ and } y \ge x \\ 0, & y < x \end{cases}$

Consider a continuous joint distribution, (X, Y), where $X \in [0,1]$ is the proportion of the time until the midterm that you actually study for it, and $Y \in [0,1]$ is your percentage score on the exam. Set up but **DO NOT EVALUATE** any of your answers. Take care in setting up the limits of integration. The joint PDF is:

 $f_{X,Y}(x,y) = \begin{cases} ce^{-(y-x)}, & x,y \in [0,1] \text{ and } y \ge x \\ 0, & y < x \end{cases}$

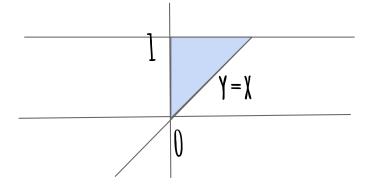
a. Sketch the joint range $\Omega_{X,Y}$, and interpret it in English.

Consider a continuous joint distribution, (X, Y), where $X \in [0,1]$ is the proportion of the time until the midterm that you actually study for it, and $Y \in [0,1]$ is your percentage score on the exam. Set up but **DO NOT EVALUATE** any of your answers. Take care in setting up the limits of integration. The joint PDF is:

 $f_{X,Y}(x,y) = \begin{cases} ce^{-(y-x)}, & x,y \in [0,1] \text{ and } y \ge x \\ 0, & y < x \end{cases}$

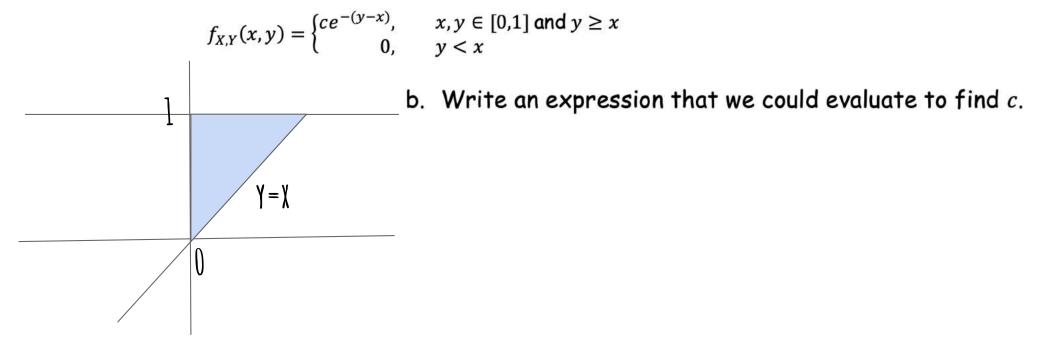
a. Sketch the joint range $\Omega_{X,Y}$, and interpret it in English.

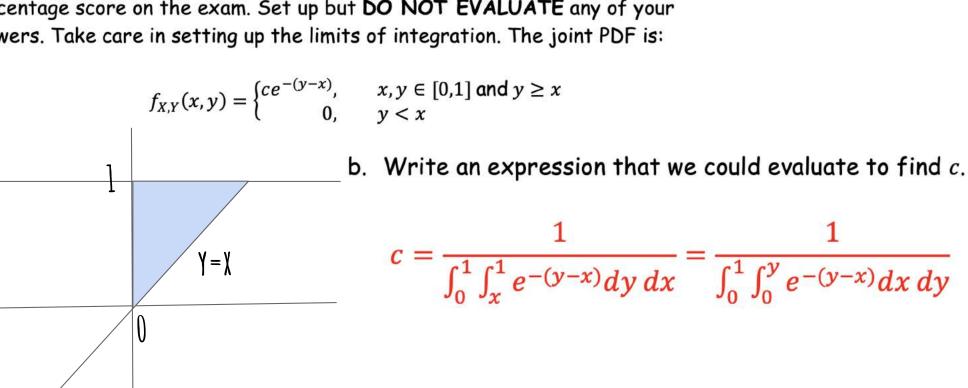
Your score is at least the proportion of the time you study.





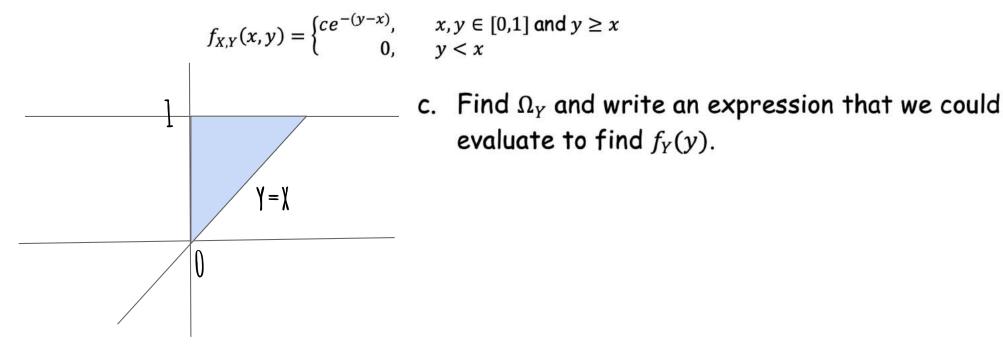


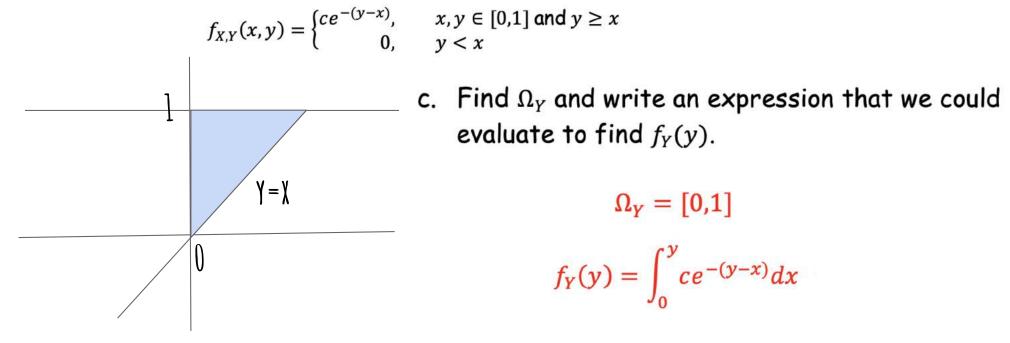






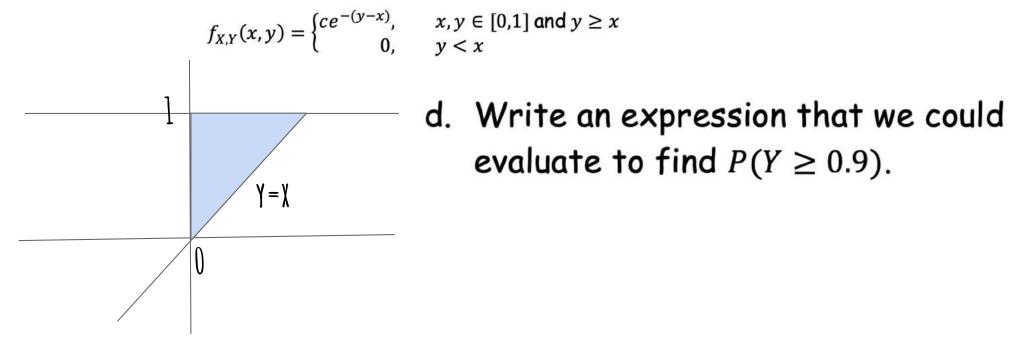


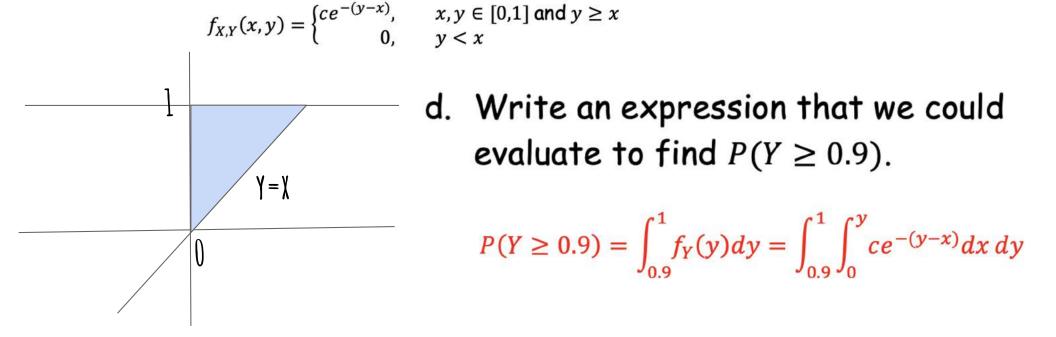








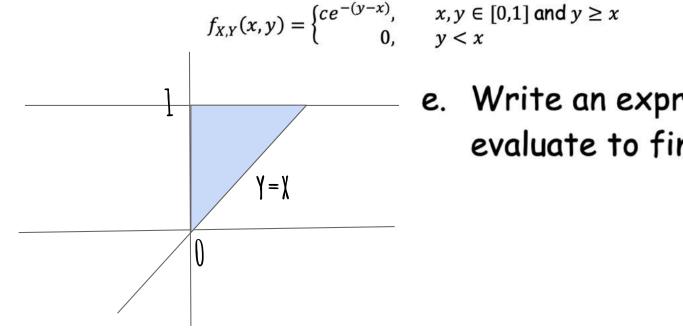






Consider a continuous joint distribution, (X, Y), where $X \in [0,1]$ is the proportion of the time until the midterm that you actually study for it, and $Y \in [0,1]$ is your percentage score on the exam. Set up but **DO NOT EVALUATE** any of your answers. Take care in setting up the limits of integration. The joint PDF is:





e. Write an expression that we could evaluate to find E[Y].

