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: 4-sided dice

independent tosses

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 $p_{X,Y}(x,y) = P(X = x, Y = y)$ for $x, y \in \Omega_{X,Y}$.

$Pr(X=2, Y=3) = p_{X,Y}(2,3)$
$= Pr(X=2)\,Pr(Y=3)$

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

Example: weird dice Again

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

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

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

Specify the joint PMF $p_{X,Y}(x,y) = P(X = x, Y = y)$ for $x, y \in \Omega_{X,Y}$.

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

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

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

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

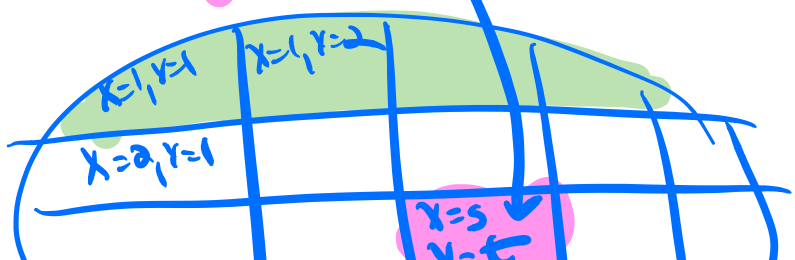
Note that

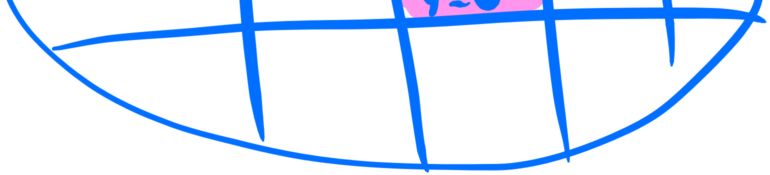
$$\sum_{(s,t) \in \Omega_{X,Y}} p_{X,Y}(s,t) = 1$$

(c,d)

Pr(X=c)=0

Pr(X=c, Y=d)=0





Example: Weird Dice Again

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 $V = \max\{X, Y\}$.

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

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

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

$P_{UV}(3,1) = 0$

$= Pr(\min = 3, \max = 1)$

U\V	1	2	3	4
1				
2				
3				
4				

$p_{U,V}(1,3) = P(U=1, V=3) = ?$

a) $\frac{1}{16}$

b) $\frac{2}{16}$

c) $\frac{1}{2}$

d) I don't know

$Pr(U=1, V=3) = Pr(\min(X,Y)=1, \max(X,Y)=3)$

$= Pr(X=1, Y=3 \cup X=3, Y=1)$

$= Pr(X=1, Y=3) + Pr(X=3, Y=1)$

$Pr(X=1)\,Pr(Y=3)$

$= \frac{1}{4}\cdot\frac{1}{4} + \frac{1}{4}\cdot\frac{1}{4} = \frac{2}{16}$

Example: Weird Dice Again

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 $V = \max\{X, Y\}$.

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

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

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

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

U\V	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

$\Pr(U=1)$

$\Updownarrow$

$p_U(1) = ?$

a) $\frac{1}{16}$
b) $\frac{2}{16}$
c) $\frac{5}{16}$
d) $\frac{7}{16}$

Example: weird dice again

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 $u \in \Omega_U$?

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

U\V	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: WEIRD DICE AGAIN

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 $u \in \Omega_U$?

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

marginal pmf of U

U\V	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

$$p_U(u) = \sum_{v \in \Omega_V} p_{UV}(u,v)$$

Example: Weird Dice Again

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 $V = \max\{X, Y\}$.

$$E(UV^2) = ?$$

$$= \sum_{u=1}^{4} \sum_{v=1}^{4} uv^2 P(U=u, V=v)$$

$$g(u,v) = uv^2$$

U\V	1	2	3	4
1	1 • 1/16	4 • 2/16	9 • 2/16	2/16
2	0	1/16	18 • 2/16	2/16
3	0	0	1/16	2/16
4	0	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

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

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

$\Pr(X=a)$ $\qquad$ $\Pr(X=a, Y=b)$

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

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

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

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

$Z = z$

$\Pr(X = x, Y = y, Z = z)$

Independence

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

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

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

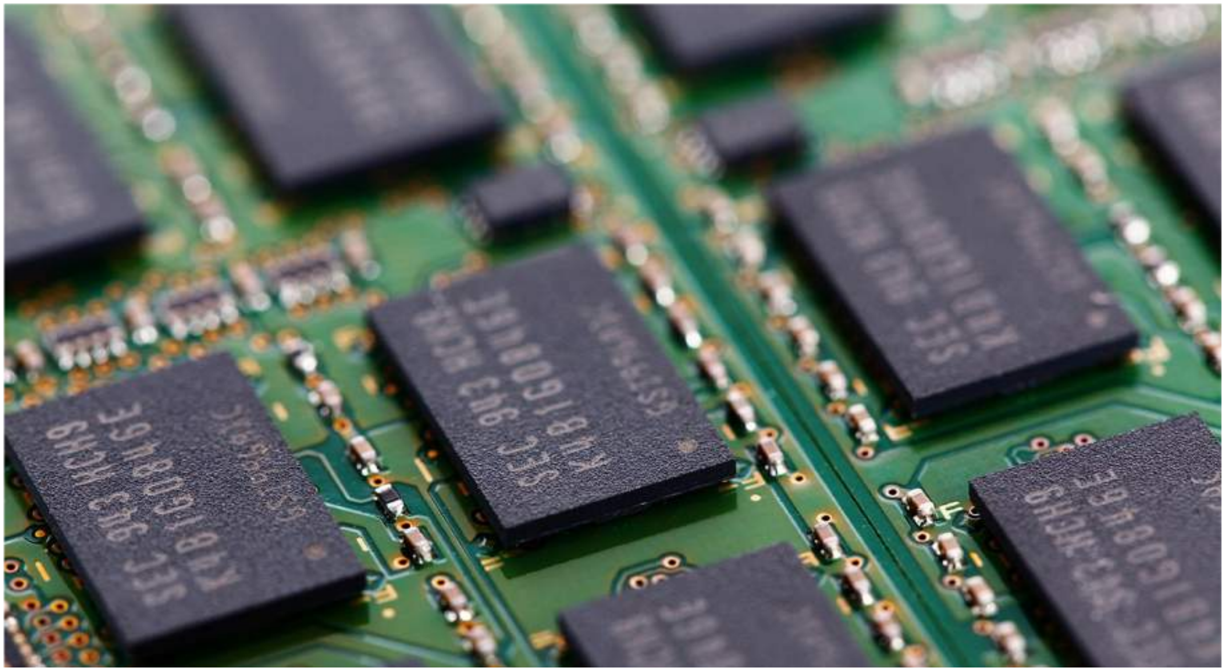
$$Var(X+Y) = Var(X) + Var(Y)$$

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

$f_X(x)\,dx \approx Pr(\text{being within } dx \text{ of } x)$

$Pr(x-\frac{\varepsilon}{2} \le X \le x+\frac{\varepsilon}{2}) \approx f_X(x)\,\varepsilon$

Joint PDFs and Expectation

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

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

$f_{XY}(a,b)\,dx\,dy \approx Pr(X \text{ is within } dx \text{ of } a,\ Y \text{ is within } dy \text{ of } b)$

The joint range is

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

Joint PDFs and Expectation

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

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

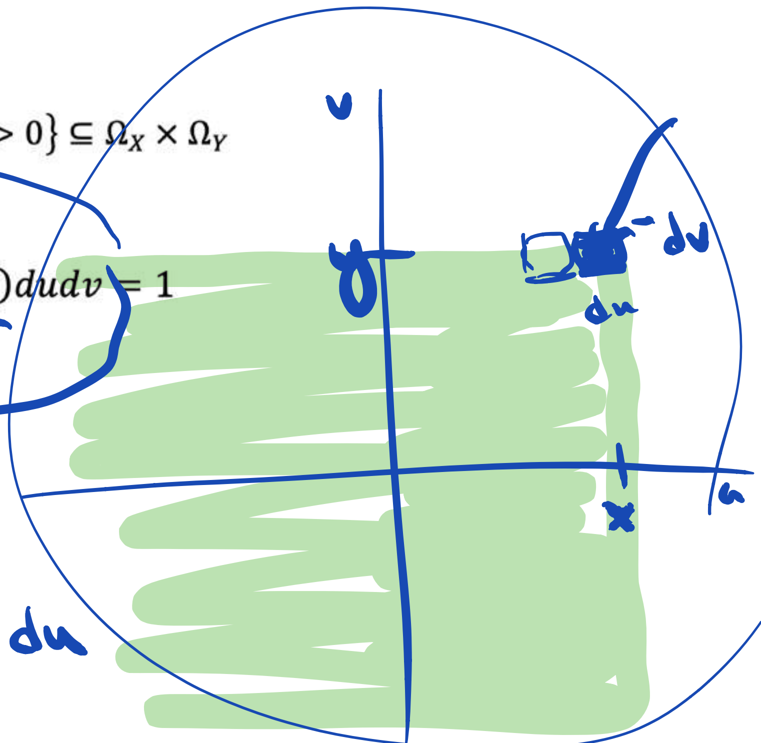
The joint range is

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

Note that

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

$$F_{X,Y}(x,y) = \Pr(X \le x, Y \le y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(u,v)\,dv\,du$$



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

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

Marginal PDFs

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

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

density at $X = x$

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

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

Independence

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

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

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.

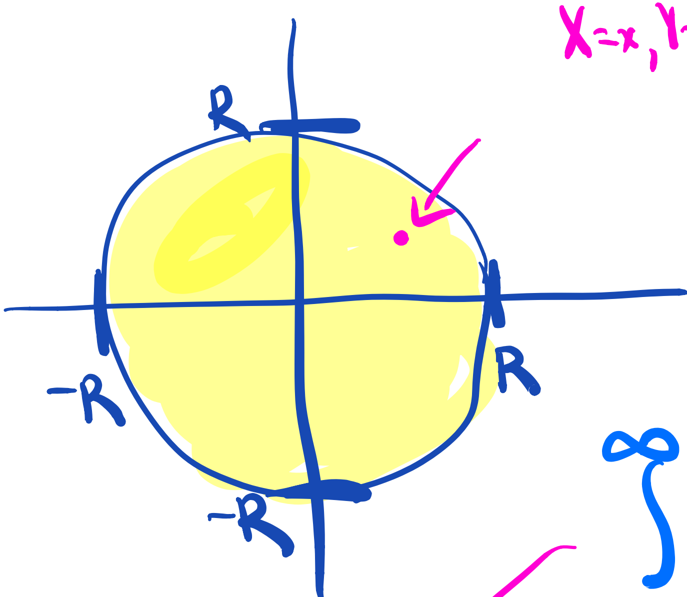
Joint PDFs (Example 1)

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 $\Omega_{X,Y}$.

$X = x, Y = y$

$\Omega_{X,Y} = \{(x,y) \mid x^2 + y^2 \leq R^2\}$



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

$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x,y)\, dy\, dx = 1$

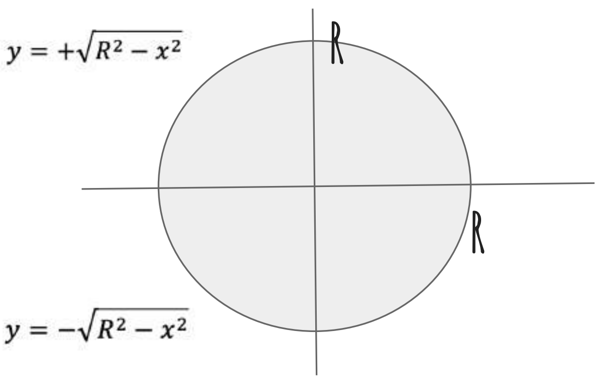
$$= c \iint_{x^2+y^2 \le R^2} dy\,dx = c\pi R^2 = 1 \Rightarrow c = \frac{1}{\pi R^2}$$

Joint PDFs (Example 1)

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 $\Omega_{X,Y}$.

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

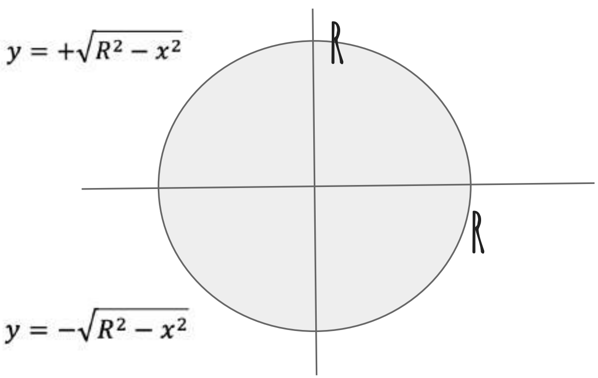


Joint PDFs (Example 1)

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 $\Omega_{X,Y}$.

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



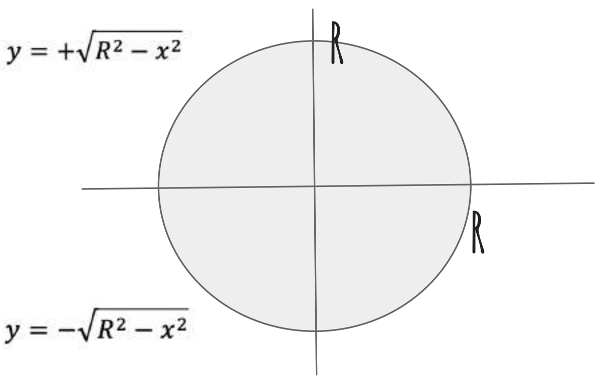
b. Write an expression for the joint PDF $f_{X,Y}(x, y)$ and carefully define it for all $x, y \in \mathbb{R}$.

Joint PDFs (Example 1)

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 $x, y \in \mathbb{R}$.

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