S&DS 241 Lecture 17

Joint and marginal distributions

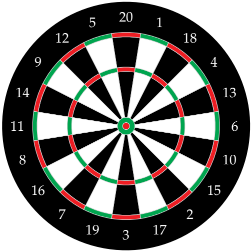
B-H: 7.1, 7.2

Motivating example

- So far we have been focusing on a single continuous random variable

Motivating example

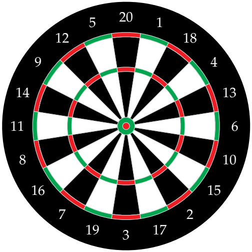
- So far we have been focusing on a single continuous random variable
- Let's consider the following example:



Suppose you are an amateur player and your dart lands uniformly at random on the board. What is the chance of hitting the bullseye? What is your average score?

Motivating example

- So far we have been focusing on a single continuous random variable
- Let's consider the following example:

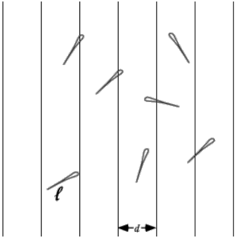


  Suppose you are an amateur player and your dart lands uniformly at random on the board. What is the chance of hitting the bullseye? What is your average score?

- Position of the dart (X, Y) is a pair of continuous random variables (RVs)

Buffon's needle

Suppose we have a floor made of parallel strips of wood, each of unit width, and we drop a needle of unit length onto the floor. What is the probability that the needle will touch a line between two strips?



- Source of randomness here: position and orientation of the needle — both are continuous random variables
- We will show that the answer is $\frac{2}{\pi}$.

Specific questions

- How to describe multiple continuous RVs?
- How to compute probability of events involving multiple RVs?
- How to make sense of conditioning and independence for continuous RVs?

Joint CDF & PDF

CDF

- The joint cumulative distribution function (CDF) of random variables X and Y is

$$F_{XY}(x, y) = P(X \leq x, Y \leq y)$$

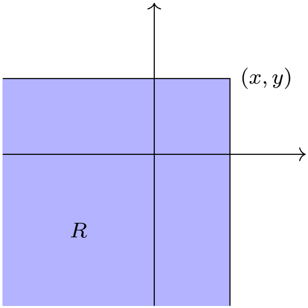
  (Reminder: comma means intersection i.e. AND)
- This defines a function $F_{XY} : \mathbb{R}^2 \to [0, 1]$, which is non-decreasing in each coordinate when the other is held fixed.

CDF

Equivalently,

$$F_{XY}(x, y) = P((X, Y) \in R)$$

where $R = (-\infty, x] \times (-\infty, y]$ is the following region



Recall: continuous random variable and PDF

A random variable X is continuous if its CDF can be expressed as an integral:

$$F_X(x) = \int_{-\infty}^{x} f_X(s)ds, \quad \text{for all } x$$

We call f_X the PDF of X, given by

$$f_X(x) = F_X'(x)$$

Continuous random variables and Joint PDF

A pair of random variables (X, Y) is continuous if its joint CDF can be expressed as a double integral, i.e., there exists a nonnegative function $f_{XY} : \mathbb{R}^2 \to \mathbb{R}$ such that

$$F_{XY}(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{XY}(s, t) ds dt, \quad \text{for all } x, y$$

We call f_{XY} the joint PDF of (X, Y), given by

$$f_{XY}(x, y) = \frac{\partial^2 F_{XY}(x, y)}{\partial x \partial y}.$$

Properties of continuous pairs of RVs

1. $F_{XY}(x, y)$ is continuous in x, y (no jumps)

Properties of continuous pairs of RVs

1. $F_{XY}(x, y)$ is continuous in x, y (no jumps)
2. For any rectangle $[a, b] \times [c, d] \subset \mathbb{R}^2$,

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

Proof: LHS $= F_{XY}(b, d) - F_{XY}(a, d) - (F_{XY}(b, c) - F_{XY}(a, c))$

Properties of continuous pairs of RVs

1. $F_{XY}(x, y)$ is continuous in x, y (no jumps)
2. For any rectangle $[a, b] \times [c, d] \subset \mathbb{R}^2$,

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

   Proof: LHS $= F_{XY}(b, d) - F_{XY}(a, d) - (F_{XY}(b, c) - F_{XY}(a, c))$

3. For region $A \subset \mathbb{R}^2$, e.g., rectangle or disk

$$P((X, Y) \in A) = \iint_A f_{XY}(x, y) dx dy$$

   Intuition: can approximate A by many rectangles

Properties of continuous pairs of RVs

1. $F_{XY}(x, y)$ is continuous in x, y (no jumps)
2. For any rectangle $[a, b] \times [c, d] \subset \mathbb{R}^2$,

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

   Proof: LHS $= F_{XY}(b, d) - F_{XY}(a, d) - (F_{XY}(b, c) - F_{XY}(a, c))$

3. For region $A \subset \mathbb{R}^2$, e.g., rectangle or disk

$$P((X, Y) \in A) = \iint_A f_{XY}(x, y) dx dy$$

   Intuition: can approximate A by many rectangles

4. For any $x, y \in \mathbb{R}$,

$$P(X = x, Y = y) = 0$$

   Moreover, $P((X, Y) \in A) = 0$ for any A with zero area. For example, $P(X = Y) = 0$.

Properties of Joint PDF

1. Non-negativity:

$$f_{XY}(x, y) \geq 0$$

2. Normalization:

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

since

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

Properties of Joint PDF

- Again, joint PDF is not a probability (can exceed 1); need to integrate to get probability.

For discrete RVs

Recall the notation $p_X(x) = P(X = x)$ denotes PMF of discrete X.
We can also define

- joint PMF

$$p_{XY}(x, y) = P(X = x, Y = y)$$

  which we have been implicitly using all along

For discrete RVs

Recall the notation $p_X(x) = P(X = x)$ denotes PMF of discrete X.
We can also define

- joint PMF

$$p_{XY}(x, y) = P(X = x, Y = y)$$

  which we have been implicitly using all along

- joint CDF

$$F_{XY}(x, y) = \sum_{a \le x} \sum_{b \le y} p_{XY}(a, b)$$

  (analogous to the connection between CDF and PMF in Lec 13)

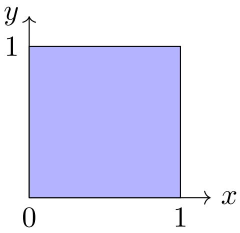
Running example

- Alice and Bob plan to go to lunch. Without any means to communicate to each other, they each arrive at a random time uniformly between noon and 1pm.

Running example

- Alice and Bob plan to go to lunch. Without any means to communicate to each other, they each arrive at a random time uniformly between noon and 1pm.
- Denote the arrival time of Alice and Bob by X and Y respectively. Then (X, Y) is continuous with joint PDF:

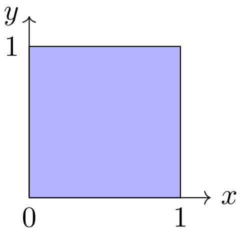
$$f_{XY}(x,y) = \begin{cases} 1 & 0 \le x \le 1, 0 \le y \le 1 \\ 0 & \text{else} \end{cases}$$



Running example

- Alice and Bob plan to go to lunch. Without any means to communicate to each other, they each arrive at a random time uniformly between noon and 1pm.
- Denote the arrival time of Alice and Bob by X and Y respectively. Then (X, Y) is continuous with joint PDF:

$$f_{XY}(x, y) = \begin{cases} 1 & 0 \leq x \leq 1, 0 \leq y \leq 1 \\ 0 & \text{else} \end{cases}$$



- This is in fact the uniform distribution over the unit square.

Lunch example (1)

What is the chance that Alice arrives earlier than Bob?

- Use joint PDF:

Lunch example (1)

What is the chance that Alice arrives earlier than Bob?

- Use joint PDF:

$$P(X < Y) = \iint f_{XY}(x, y) dx dy = \int_0^1 dx \int_x^1 dy = \int_0^1 dx(1-x) = \frac{1}{2}$$

Lunch example (1)

What is the chance that Alice arrives earlier than Bob?

- Use joint PDF:

$$P(X < Y) = \iint$$  $$f_{XY}(x,y)dxdy = \int_0^1 dx \int_x^1 dy = \int_0^1 dx(1-x) = \frac{1}{2}$$

- In hindsight, we could have done

$$P(X < Y) = \text{area}\,\Big($$  $$\Big) = \frac{1}{2}.$$

Lunch example (1)

What is the chance that Alice arrives earlier than Bob?

- Use joint PDF:

$$P(X<Y)=\iint_{}$$

$$f_{XY}(x,y)dxdy=\int_0^1 dx\int_x^1 dy=\int_0^1 dx(1-x)=\frac{1}{2}$$

- In hindsight, we could have done

$$P(X<Y)=\text{area}\Bigg($$

$$\Bigg)=\frac{1}{2}.$$

- In further hindsight, we could have used symmetry to conclude

$$P(X<Y)=\frac{1}{2}$$

Lunch example (2)

Suppose both Alice and Bob are willing to wait for each other for at most 20 minutes, that is, whoever arrives the first will leave if the other person does not show up in 20 minutes. What is the probability that the lunch takes place?

Lunch example (2)

Suppose both Alice and Bob are willing to wait for each other for at most 20 minutes, that is, whoever arrives the first will leave if the other person does not show up in 20 minutes. What is the probability that the lunch takes place?

- Note that
$$\{\text{lunch takes place}\} = \{|X - Y| \le 1/3\}$$

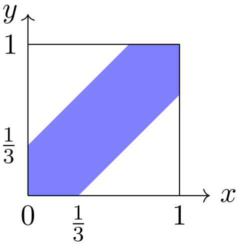
Lunch example (2)

Suppose both Alice and Bob are willing to wait for each other for at most 20 minutes, that is, whoever arrives the first will leave if the other person does not show up in 20 minutes. What is the probability that the lunch takes place?

- Note that

$$\{\text{lunch takes place}\} = \{|X - Y| \leq 1/3\}$$

that is, (X, Y) belongs to the region



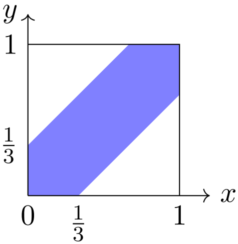
Lunch example (2)

Suppose both Alice and Bob are willing to wait for each other for at most 20 minutes, that is, whoever arrives the first will leave if the other person does not show up in 20 minutes. What is the probability that the lunch takes place?

- Note that

$$\{\text{lunch takes place}\} = \{|X - Y| \le 1/3\}$$

  that is, (X, Y) belongs to the region



- Thus

$$P(\text{lunch takes place}) = \text{area}\Big($$  $$\Big) = 5/9$$

LOTUS in 2-D

Let $Z = g(X, Y)$. As usual, two ways to find $E(Z)$:

1. Find CDF F_Z, then PDF f_Z, then $E(Z) = \int z f_Z(z) dz$.

LOTUS in 2-D

Let $Z = g(X, Y)$. As usual, two ways to find $E(Z)$:

1. Find CDF F_Z, then PDF f_Z, then $E(Z) = \int z f_Z(z) dz$.
2. Invoke LOTUS rule:

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

Lunch example (3)

Let Z denote the arrival time of the later one of the two (Alice and Bob). Find the pdf of Z and its mean.

Lunch example (3)

Let Z denote the arrival time of the later one of the two (Alice and Bob). Find the pdf of Z and its mean.

- Let $Z = \max(X, Y)$, taking values in $[0, 1]$

Lunch example (3)

Let Z denote the arrival time of the later one of the two (Alice and Bob). Find the pdf of Z and its mean.

- Let $Z = \max(X, Y)$, taking values in $[0, 1]$
- Find its CDF: for any $z \in [0, 1]$,

$$F_Z(z) = P(Z \leq z) = P(X \leq z, Y \leq z) = \int_0^z \int_0^z f_{XY}(x, y)dxdy = z^2$$

Lunch example (3)

Let Z denote the arrival time of the later one of the two (Alice and Bob). Find the pdf of Z and its mean.

- Let $Z = \max(X, Y)$, taking values in $[0, 1]$
- Find its CDF: for any $z \in [0, 1]$,

$$F_Z(z) = P(Z \leq z) = P(X \leq z, Y \leq z) = \int_0^z \int_0^z f_{XY}(x, y) dx dy = z^2$$

- PDF of Z:

$$f_Z(z) = F'_Z(z) = \begin{cases} 2z & 0 \leq z \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Lunch example (3)

Let Z denote the arrival time of the later one of the two (Alice and Bob). Find the pdf of Z and its mean.

- Let $Z = \max(X, Y)$, taking values in $[0, 1]$
- Find its CDF: for any $z \in [0, 1]$,

$$F_Z(z) = P(Z \leq z) = P(X \leq z, Y \leq z) = \int_0^z \int_0^z f_{XY}(x, y)dxdy = z^2$$

- PDF of Z:

$$f_Z(z) = F_Z'(z) = \begin{cases} 2z & 0 \leq z \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

- Then $E(Z) = \int_0^1 2z^2 dz = \frac{2}{3} > E(X) = E(Y) = \frac{1}{2}$

Alternatively, using LOTUS

Since $Z = \max(X, Y)$,

$$
\begin{aligned}
E(Z) &= \iint \max(x,y) f_{XY}(x,y)dxdy = \int_0^1 \int_0^1 \max(x,y)dxdy \\
&= \int_0^1 \left(\int_0^y ydx \right) dy + \int_0^1 \left(\int_y^1 xdx \right) dy \\
&= \int_0^1 y^2 dy + \int_0^1 \frac{1-y^2}{2} dy = \frac{2}{3}
\end{aligned}
$$

Marginalization

Marginal PDF

- The joint distribution of X, Y contains more information than the (marginal) distribution of X or Y individually.
- Joint PDF determines the marginal PDFs, but not vice versa

How to recover the distribution of X from the joint distribution of X, Y?

Joint CDF $\rightarrow$ Marginal CDF

Recall:

$$F_{XY}(x, y) = P(X \leq x, Y \leq y)$$

Joint CDF → Marginal CDF

Recall:

$$\begin{aligned} & F_{XY}(x,y) = P(X \le x, Y \le y) \\ \xRightarrow{y=\infty}\ & F_{XY}(x,\infty) = P(X \le x, Y \le \infty) = \underbrace{P(X \le x)}_{F_X(x)} \end{aligned}$$

Joint CDF → Marginal CDF

Recall:

$$F_{XY}(x,y) = P(X \le x, Y \le y)$$
$$\xRightarrow{y=\infty} F_{XY}(x,\infty) = P(X \le x, Y \le \infty) = \underbrace{P(X \le x)}_{F_X(x)}$$

Summary:

$$\boxed{\begin{aligned} F_X(x) &= F_{XY}(x,\infty) \\ F_Y(y) &= F_{XY}(\infty,y) \end{aligned}}$$

Joint CDF → Marginal CDF

Recall:

$$
\begin{aligned}
& F_{XY}(x,y) = P(X \le x, Y \le y) \\
\xRightarrow{y=\infty}\ & F_{XY}(x,\infty) = P(X \le x, Y \le \infty) = \underbrace{P(X \le x)}_{F_X(x)}
\end{aligned}
$$

Summary:

$$
\boxed{\begin{aligned} F_X(x) &= F_{XY}(x,\infty) \\ F_Y(y) &= F_{XY}(\infty,y) \end{aligned}}
$$

Here $F_{XY}(x,\infty)$ is understood as $\lim_{y\to\infty} F_{XY}(x,y)$, $F_{XY}(\infty,y)$ as $\lim_{x\to\infty} F_{XY}(x,y)$.

Joint PDF → Marginal PDF

To find marginal PDF, let's express the marginal CDF as an integral:

$$F_X(x) = F_{XY}(x, \infty) = \int_{-\infty}^{x} \int_{-\infty}^{\infty} f_{XY}(s,t) ds dt$$

$$= \int_{-\infty}^{x} \underbrace{\left(\int_{-\infty}^{\infty} f_{XY}(s,t) dt \right)}_{f_X(s)} ds$$

Joint PDF → Marginal PDF

To find marginal PDF, let's express the marginal CDF as an integral:

$$F_X(x) = F_{XY}(x, \infty) = \int_{-\infty}^{x} \int_{-\infty}^{\infty} f_{XY}(s,t)dsdt$$
$$= \int_{-\infty}^{x} \underbrace{\left(\int_{-\infty}^{\infty} f_{XY}(s,t)dt \right)}_{f_X(s)} ds$$

Summary:

$$\boxed{\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} f_{XY}(x,y)dy \\ f_Y(y) &= \int_{-\infty}^{\infty} f_{XY}(x,y)dx \end{aligned}}$$

Joint PMF $\rightarrow$ Marginal PMF

Entirely analogously,

$$\boxed{\begin{aligned} p_X(x) &= \sum_y p_{XY}(x,y) \\ p_Y(y) &= \sum_x p_{XY}(x,y) \end{aligned}}$$

Example: Contingency table

Consider a randomly sampled individual.

- Let X be the indicator of being a smoker.
- Let Y be the indicator of developing lung cancer.

Suppose p_{XY} is given by:

	$Y = 1$	$Y = 0$	
$X = 1$	0.05	0.20	
$X = 0$	0.01	0.74	

Example: Contingency table

Consider a randomly sampled individual.

- Let X be the indicator of being a smoker.
- Let Y be the indicator of developing lung cancer.

Suppose p_{XY} is given by:

	$Y = 1$	$Y = 0$	
$X = 1$	0.05	0.20	
$X = 0$	0.01	0.74	
Sum	0.06	0.94	

Example: Contingency table

Consider a randomly sampled individual.

- Let X be the indicator of being a smoker.
- Let Y be the indicator of developing lung cancer.

Suppose p_{XY} is given by:

	$Y = 1$	$Y = 0$	Sum
$X = 1$	0.05	0.20	0.25
$X = 0$	0.01	0.74	0.75
Sum	0.06	0.94	1

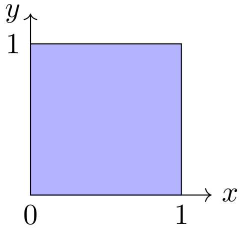
Summary

To marginalize the joint PDF/PMF to one variable, integrate/sum out the other variable.

Lunch example (4)

- Joint PDF of (X, Y), the arrival time of Alice and Bob:

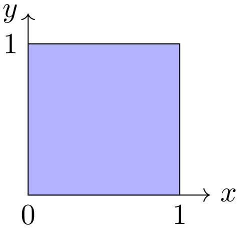
$$f_{XY}(x,y) = \begin{cases} 1 & 0 \le x \le 1, 0 \le y \le 1 \\ 0 & \text{else} \end{cases}$$



Lunch example (4)

- Joint PDF of (X, Y), the arrival time of Alice and Bob:

$$f_{XY}(x, y) = \begin{cases} 1 & 0 \le x \le 1, 0 \le y \le 1 \\ 0 & \text{else} \end{cases}$$



- Marginal PDF of X:

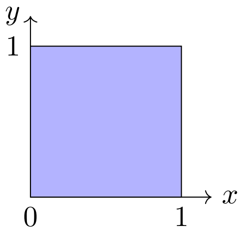
$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y)dy = \int_0^1 dy = 1,$$

if $x \in [0, 1]$; otherwise $f_X(x) = 0$.

Lunch example (4)

- Joint PDF of (X, Y), the arrival time of Alice and Bob:

$$f_{XY}(x,y) = \begin{cases} 1 & 0 \le x \le 1, 0 \le y \le 1 \\ 0 & \text{else} \end{cases}$$



- Marginal PDF of X:

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)dy = \int_0^1 dy = 1,$$

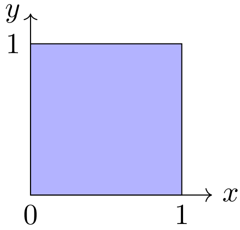
if $x \in [0,1]$; otherwise $f_X(x) = 0$.

- Marginal PDF of Y: same.

Lunch example (4)

- Joint PDF of (X, Y), the arrival time of Alice and Bob:

$$f_{XY}(x,y) = \begin{cases} 1 & 0 \le x \le 1, 0 \le y \le 1 \\ 0 & \text{else} \end{cases}$$



- Marginal PDF of X:

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)dy = \int_0^1 dy = 1,$$

if $x \in [0,1]$; otherwise $f_X(x) = 0$.

- Marginal PDF of Y: same.
- Both X and Y are distributed as Unif$(0,1)$.

Another (abstract) example

$$
\begin{aligned}
& E(X+Y) \\
\overset{\text{LOTUS}}{=} & \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x+y)f_{XY}(x,y)dxdy
\end{aligned}
$$

Another (abstract) example

$$\begin{aligned} & E(X+Y) \\ \overset{\text{LOTUS}}{=} & \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x+y)f_{XY}(x,y)dxdy \\ = & \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xf_{XY}(x,y)dxdy + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}yf_{XY}(x,y)dxdy \end{aligned}$$

Another (abstract) example

$$
\begin{aligned}
& E(X+Y) \\
\overset{\text{LOTUS}}{=} & \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x+y)f_{XY}(x,y)dxdy \\
= & \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xf_{XY}(x,y)dxdy + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}yf_{XY}(x,y)dxdy \\
= & \int_{-\infty}^{\infty}xdx\underbrace{\left(\int_{-\infty}^{\infty}f_{XY}(x,y)dy\right)}_{f_X(x)} + \int_{-\infty}^{\infty}ydy\underbrace{\left(\int_{-\infty}^{\infty}f_{XY}(x,y)dx\right)}_{f_Y(y)}
\end{aligned}
$$

Another (abstract) example

$$
\begin{aligned}
& E(X+Y) \\
\overset{\text{LOTUS}}{=} & \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x+y)f_{XY}(x,y)dxdy \\
= & \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xf_{XY}(x,y)dxdy + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}yf_{XY}(x,y)dxdy \\
= & \int_{-\infty}^{\infty}xdx\underbrace{\left(\int_{-\infty}^{\infty}f_{XY}(x,y)dy\right)}_{f_X(x)} + \int_{-\infty}^{\infty}ydy\underbrace{\left(\int_{-\infty}^{\infty}f_{XY}(x,y)dx\right)}_{f_Y(y)} \\
= & \underbrace{\int_{-\infty}^{\infty}xf_X(x)dx}_{E(X)} + \underbrace{\int_{-\infty}^{\infty}yf_Y(y)dy}_{E(Y)}
\end{aligned}
$$

Another (abstract) example

$$
\begin{aligned}
& E(X+Y) \\
\overset{\text{LOTUS}}{=} & \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x+y)f_{XY}(x,y)dxdy \\
= & \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xf_{XY}(x,y)dxdy+\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}yf_{XY}(x,y)dxdy \\
= & \int_{-\infty}^{\infty}xdx\underbrace{\left(\int_{-\infty}^{\infty}f_{XY}(x,y)dy\right)}_{f_X(x)}+\int_{-\infty}^{\infty}ydy\underbrace{\left(\int_{-\infty}^{\infty}f_{XY}(x,y)dx\right)}_{f_Y(y)} \\
= & \underbrace{\int_{-\infty}^{\infty}xf_X(x)dx}_{E(X)}+\underbrace{\int_{-\infty}^{\infty}yf_Y(y)dy}_{E(Y)}
\end{aligned}
$$

This justifies the linearity of expectation for continuous random variables.

Uniform distribution in 2D

2-dimensional uniform distribution

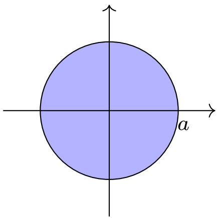
Let $A \subset \mathbb{R}^2$ be a region on the plane with finite area.

- We say (X, Y) is uniformly distributed over A if the joint PDF is

$$f_{XY}(x,y) = \begin{cases} \frac{1}{\text{area}(A)} & (x,y) \in A \\ 0 & \text{else} \end{cases}$$

- In the previous lunch example, the arrival times of Alice and Bob are uniform over the unit square $[0,1]^2$

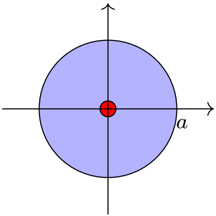
Example: Dart



Let (X, Y) denote the coordinate of the dart, which is uniformly distributed over the board (a disk of radius a). Then

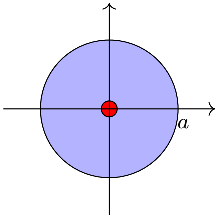
$$f_{XY}(x, y) = \begin{cases} \frac{1}{\pi a^2} & x^2 + y^2 \leq a^2 \\ 0 & \text{else} \end{cases}$$

Example: Dart



What's the probability of hitting the bullseye?

Example: Dart

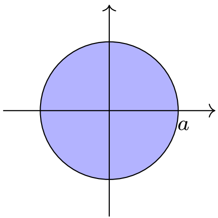


What's the probability of hitting the bullseye?

- For typical dartboard: diameter $= 17.75$", bullseye diameter $= 0.5$"

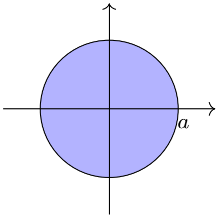
$$P(\text{bullseye}) = \iint_{\text{bullseye}} f_{XY}(x,y)dxdy = \frac{\text{area(bullseye)}}{\text{area(board)}}$$
$$= \frac{0.5^2}{17.75^2} \approx \frac{1}{1260}$$

Example: Dart



What's the distribution of the horizontal coordinate X?

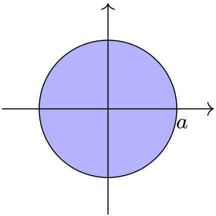
Example: Dart



What's the distribution of the horizontal coordinate X?

- X takes values in $[-a, a]$

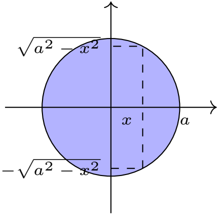
Example: Dart



What's the distribution of the horizontal coordinate X?

- X takes values in $[-a, a]$
- Uniform?

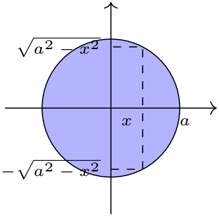
Example: Dart



Find the marginal PDF of X: for $x \in [-a, a]$

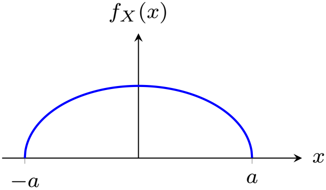
$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y)dy = \int_{-\sqrt{a^2-x^2}}^{\sqrt{a^2-x^2}} \frac{1}{\pi a^2} dy = \frac{2\sqrt{a^2 - x^2}}{\pi a^2}$$

Example: Dart



Find the marginal PDF of X: for $x \in [-a, a]$

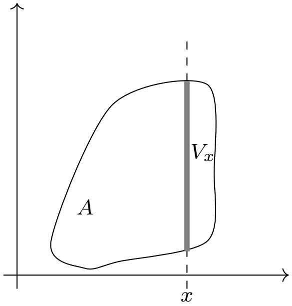
$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y)dy = \int_{-\sqrt{a^2-x^2}}^{\sqrt{a^2-x^2}} \frac{1}{\pi a^2} dy = \frac{2\sqrt{a^2 - x^2}}{\pi a^2}$$



Not uniform: more likely to be near the middle than ends. Why?

More generally: marginalize a uniform distribution

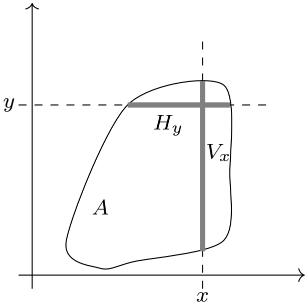
Let (X, Y) be uniformly distributed over a region A:



$$f_X(x) = \int_{V_x} f_{XY}(x, y)dy = \frac{\text{length}(V_x)}{\text{area}(A)}$$

More generally: marginalize a uniform distribution

Let (X, Y) be uniformly distributed over a region A:



$$f_X(x) = \int_{V_x} f_{XY}(x, y)dy = \frac{\text{length}(V_x)}{\text{area}(A)}$$

$$f_Y(y) = \int_{H_y} f_{XY}(x, y)dx = \frac{\text{length}(H_y)}{\text{area}(A)}$$

where H_y and V_x denote the horizontal and vertical segment intersecting A, respectively.