Joint Distributions

CSE 312 26Su

Lecture 17

Announcements

Homework 5 due Wednesday

Quiz 5 on Friday

CLT joint dist.

- polling

Where Are We?

Last time:

- Polling

Today:

- Joint Distributions
  - Joint Support
  - Joint PMF / PDF
  - Joint CDF
  - Joint Expectation
- Conditional Distributions
  - Conditional Expectation
- Law of Total Expectation
- Covariance

Multiple Random Variables

Analyzing *Multiple* Random Variables

So far, we've pretty much only analyzed one random variable at a time

We've *worked* with multiple random variables (e.g., the sum of them) but haven't really analyzed their relationships except for independence

$Z = X + Y$

Today, we will talk about <u>analyzing the relations between multiple RVs</u>

$X \sim Bin(n, p)$

$P(X \leq k)$

Analyzing *Multiple* Random Variables

Examples:

- **Economics**: Analyzing the relationship between stock returns and volume helps in understanding market behavior and making investment decisions
- **Healthcare:** Understanding how different factors (like blood pressure and cholesterol levels) affects the probability of medical conditionals
- **Machine learning**: understanding how different features contribute to predicting a label and improving model accuracy (what features are most important)
- **Recommendation systems**: Understanding relationship between user preferences and item features to improve recommendation systems, or other factors like age and choices of products

Joint Distributions

We'll start with the **discrete** case

We have two **discrete** random variables X and Y

(that may or may not be independent)

We have two **discrete** random variables X and Y
(that may or may not be independent)

Support Ω_X

Joint **Support/Range** - $\Omega_{X,Y}$

$\Omega_{X,Y}$ is the set of all possible **pairs** of values X and Y can be together

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

x y

We have two $\boxed{\text{discrete}}$ random variables X and Y
(that may or may not be independent)

Joint **Support/Range** - $\Omega_{X,Y}$

$\Omega_{X,Y}$ is the set of all possible **pairs** of values X and Y can be together

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

Joint **PMF** *(probability mass function)* - $\mathrm{p}_{X,Y}(a,b)$

$\mathrm{p}_{X,Y}(a,b)$ defines probabilities of X and Y being certain values at once

> $p_{X,Y}(a,b) = \mathbb{P}(X = a \cap Y = b) = \mathbb{P}(X = a, Y = b)$
> *should be defined for **all** values of a and b*

Normalization Property: $\sum_{(a,b)\in\Omega_{X,Y}} p_{X,Y}(a,b) = 1$

$\sum_{x\in\Omega_X} p_X(x) = 1$

We have two **discrete** random variables X and Y
(that may or may not be independent)

Joint **Support/Range** - $\Omega_{X,Y}$

$\Omega_{X,Y}$ is the set of all possible **pairs** of values X and Y can be together

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

Joint **PMF** *(probability mass function)*- $\mathrm{p}_{X,Y}(a, b)$

$\mathrm{p}_{X,Y}(a, b)$ defines probabilities of X and Y being certain values at once

> $p_{X,Y}(a, b) = \mathbb{P}(X = a \cap Y = b) = \mathbb{P}(X = a, Y = b)$

*should be defined for **all** values of a and b*

Joint **CDF** *(cumulative distribution function)*- $\mathrm{F}_{X,Y}(a, b)$

> $F_{X,Y}(a, b) = \mathbb{P}(X \leq a \cap Y \leq b) = \mathbb{P}(X \leq a, Y \leq b)$

*should be defined for **all** values of a and b*

We have two **discrete** random variables X and Y
(that may or may not be independent)

(Joint) **Independence**

X and Y are independent if:

> $p_{X,Y}(a,b) = p_X(a) \cdot p_Y(b)$ for all $(a,b) \in \Omega_{X,Y}$ ✓
>
> $\Omega_{X,Y} = \Omega_X \times \Omega_Y$ ✓
>
> *if you find that $\Omega_{X,Y} \neq \Omega_X \times \Omega_Y$, that is enough to show that they aren't independent*

This is the same definition we've seen before for independence of random variables
$\mathbb{P}(X = a, Y = b) = \mathbb{P}(X = a) \cdot \mathbb{P}(X = b)$

Example: Weird Dice

Roll 2 fair 4-sided dice independently

Let X be the value of the 1st dice
Let Y be the value of the 2nd dice

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

The **joint support** is $\Omega_{X,Y} = \Omega_X \times \Omega_Y$
because all combinations are possible

What is the **joint PMF**?

$$p_{X,Y}(a,b) = \begin{cases} 1/16 & \text{if } (a,b) \in \Omega_{X,Y} \\ 0 & \text{o.w.} \end{cases}$$

$p_{X,Y}$	Y=1	Y=2	Y=3	Y=4
X=1	$\frac{1}{4} \cdot \frac{1}{4} = \frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$
X=2				
X=3				
X=4				

Example: Weird Dice

Roll 2 fair 4-sided dice independently

Let X be the value of the 1st dice
Let Y be the value of the 2nd dice

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

The **joint support** is $\Omega_{X,Y} = \Omega_X \times \Omega_Y$
because all combinations are possible

What is the **joint PMF**?

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

$p_{X,Y}$	Y=1	Y=2	Y=3	Y=4
X=1	1/16	1/16	1/16	1/16
X=2	1/16	1/16	1/16	1/16
X=3	1/16	1/16	1/16	1/16
X=4	1/16	1/16	1/16	1/16

Example: Weird Dice

X, Y – 2nd die roll
1st die roll

Roll 2 fair 4-sided dice independently

Let U be the min value $U = \min(X, Y)$
Let W be the max value $W = \max(X, Y)$

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

The **joint support** is:

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

$p_{U,W}$	*W*=1	*W*=2	*W*=3	*W*=4
U=1				
U=2				
U=3				
U=4				

Example: Weird Dice

Roll 2 fair 4-sided dice independently

Let U be the min value $U = \min(X, Y)$
Let W be the max value $W = \max(X, Y)$

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

The **joint support** is:
$\Omega_{U,W} = \{(u, w) \in \Omega_U \times \Omega_W; u \leq w\}$

$p_{U,W}$	W=1	W=2	W=3	W=4
U=1				
U=2				
U=3				
U=4				

Example: Weird Dice

e.g.
→ roll $\boxed{X=2, Y=4}$
then $U=2$, $W=4$

$\boxed{X=4, Y=2}$
then $U=2$, $W=4$

Roll 2 fair 4-sided dice independently

Let U be the min value $U = \min(X, Y)$
Let W be the max value $W = \max(X, Y)$

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

The **joint support** is:
$\Omega_{U,W} = \{(u, w) \in \Omega_U \times \Omega_W; \boxed{u \leq w}\}$

What is the **joint PMF?**

$p_{U,W}$	W=1	W=2	W=3	W=4
U=1	1/16	2/16	2/16	2/16
U=2	0	1/16	2/16	2/16
U=3	0	0	1/16	2/16
U=4	0	0	0	1/16

$$p_{U,W}(u, w) = \begin{cases} 1/16 & \text{if } (u,w) \in \Omega_{u,w} \text{ and } u = w \\ 2/16 & \text{if } (u,w) \in \Omega_{u,w} \text{ and } u < w \\ 0 & \text{o.w.} \end{cases}$$

Example: Weird Dice

Roll 2 fair 4-sided dice independently

Let U be the min value $U = \min(X, Y)$
Let W be the max value $W = \max(X, Y)$

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

The **joint support** is:
$\Omega_{U,W} = \{(u, w) \in \Omega_U \times \Omega_W; u \le w\}$

$p_{U,W}$	W=1	W=2	W=3	W=4
U=1	1/16	2/16	2/16	2/16
U=2	0	1/16	2/16	2/16
U=3	0	0	1/16	2/16
U=4	0	0	0	1/16

What is the **joint PMF?**

$$p_{U,W}(u, w) = \begin{cases} 1/16 & \text{if } (u, w) \in \Omega_{U,W} \text{ and } u = w \\ 2/16 & \text{if } (u, w) \in \Omega_{U,W} \text{ and } u < w \\ 0 & \text{otherwise} \end{cases}$$

Example: Weird Dice

Roll 2 fair 4-sided dice independently

Let U be the min value $U = \min(X, Y)$
Let W be the max value $W = \max(X, Y)$

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

The **joint support** is:
$\Omega_{U,W} = \{(u, w) \in \Omega_U \times \Omega_W ; u \leq w\}$

$p_{U,W}$	W=1	W=2	W=3	W=4
U=1	1/16	2/16	2/16	2/16
U=2	0	1/16	2/16	2/16
U=3	0	0	1/16	2/16
U=4	0	0	0	1/16

What is the **joint PMF?**

$$p_{U,W}(u, w) = \begin{cases} 1/16 & \text{if } (u, w) \in \Omega_{U,W} \text{ and } u = w \\ 2/16 & \text{if } (u, w) \in \Omega_{U,W} \text{ and } u < w \\ 0 & \text{otherwise} \end{cases}$$

U and W are **NOT independent** because $\Omega_{U,W} \neq \Omega_U \times \Omega_W$

Example: Weird Dice

Roll 2 fair 4-sided dice independently

Let U be the min value $U = \min(X, Y)$. Let W be the max value $W = \max(X, Y)$

Suppose I didn't know how to compute $p_U(u) = \mathbb{P}(U = u)$ *directly.*
Can we find it from the joint PMF $p_{U,W}(u, w)$*?*

We know $\mathbb{P}(U = 1 \cap W = w)$ for all w…

$\mathbb{P}(U = 1)$
$= P(U=1 \cap W=1) +$
$P(U=1 \cap W=2) +$
$P(U=1 \cap W=3) +$
$P(U=1 \cap W=4)$

$p_{U,W}$	W=1	W=2	W=3	W=4
U=1	1/16	2/16	2/16	2/16
U=2	0	1/16	2/16	2/16
U=3	0	0	1/16	2/16
U=4	0	0	0	1/16

Example: Weird Dice

Roll 2 fair 4-sided dice independently

Let U be the min value $U = \min(X, Y)$. Let W be the max value $W = \max(X, Y)$

Suppose I didn't know how to compute $p_U(u) = \mathbb{P}(U = u)$ *directly.*
Can we find it from the joint PMF $p_{U,W}(u, w)$*?*

We know $\mathbb{P}(U = 1 \cap W = w)$ for all w...

$$\mathbb{P}(U = 1)$$
$$= \mathbb{P}(U = 1 \cap W = 1) + \mathbb{P}(U = 1 \cap W = 2) + \mathbb{P}(U = 1 \cap W = 3) + \mathbb{P}(U = 1 \cap W = 4)$$
$$= \frac{1}{16} + \frac{2}{16} + \frac{2}{16} + \frac{2}{16} = \frac{7}{16}$$

Use LoTP because the events $W = 1$, $W = 2$, $W = 3$, $W = 4$ ***partition the sample space***

$p_{U,W}$	W=1	W=2	W=3	W=4
U=1	1/16	2/16	2/16	2/16
U=2	0	1/16	2/16	2/16
U=3	0	0	1/16	2/16
U=4	0	0	0	1/16

Example: Weird Dice

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

Roll 2 fair 4-sided dice independently

Let U be the min value $U = \min(X, Y)$. Let W be the max value $W = \max(X, Y)$

Suppose I didn't know how to compute $p_U(u) = \mathbb{P}(U = u)$ directly.
Can we find it from the joint PMF $p_{U,W}(u, w)$?

In general...

$\mathbb{P}(U = u)$
$= \mathbb{P}(U = u \cap W = 1) + \mathbb{P}(U = u \cap W = 2)$
$\quad + \mathbb{P}(U = u \cap W = 3) + \mathbb{P}(U = u \cap W = 4)$
$= \sum_{w \in \Omega_W} p_{U,W}(u, w)$

*Use __LoTP__ because the events $W = 1$, $W = 2$, $W = 3$, $W = 4$ **partition the sample space***

$p_{U,W}$	**W=1**	**W=2**	**W=3**	**W=4**
U=1	1/16	2/16	2/16	2/16
U=2	0	1/16	2/16	2/16
U=3	0	0	1/16	2/16
U=4	0	0	0	1/16

Example: Weird Dice

Roll 2 fair 4-sided dice independently

Let U be the min value $U = \min(X, Y)$. Let W be the max value $W = \max(X, Y)$

Suppose I didn't know how to compute $p_U(u) = \mathbb{P}(U = u)$ directly.
Can we find it from the joint PMF $p_{U,W}(u, w)$?

In general…

$\mathbb{P}(U = u) = \sum_{w \in \Omega_W} p_{U,W}(u, w)$

*Use LoTP because the events $W = 1$, $W = 2$, $W = 3$, $W = 4$ **partition the sample space***

$p_U(u)$ is the "marginal" PMF for U
(because we "marginalized" W)

$p_{U,W}$	W=1	W=2	W=3	W=4
U=1	1/16	2/16	2/16	2/16
U=2	0	1/16	2/16	2/16
U=3	0	0	1/16	2/16
U=4	0	0	0	1/16

Marginal Distribution

If you have the joint PMF for two random variables, you can find the PMF of one of them by **using the law of total probability**, **partitioning on the values of the other random variable:**

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

^ it's the same PMF we've talked about before, the "marginal" is just there to indicate is was derived from a joint PMF

Example: Weird Dice

Roll 2 fair 4-sided dice independently

Let U be the min value $U = \min(X, Y)$. Let W be the max value $W = \max(X, Y)$

Suppose I didn't know how to compute $p_U(u) = \mathbb{P}(U = u)$ *directly.*
Can we find it from the joint PMF $p_{U,W}(u, w)$*?*

$$p_U(u) = \begin{cases} \frac{7}{16} & \text{if } u = 1 \\ \frac{5}{16} & \text{if } u = 2 \\ \frac{3}{16} & \text{if } u = 3 \\ \frac{1}{16} & \text{if } u = 4 \\ 0 & \text{otherwise} \end{cases}$$

$p_{U,W}$	W=1	W=2	W=3	W=4
U=1	1/16	2/16	2/16	2/16
U=2	0	1/16	2/16	2/16
U=3	0	0	1/16	2/16
U=4	0	0	0	1/16

Joint **Expectation**

Joint Expectations

For a function $\boldsymbol{g(X,Y)}$, the expectation can be written in terms of the joint PMF.

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

LOTUS

Same ideas as before!

Examples of joint functions:

$g(X,Y) = X + Y$, $g(X,Y) = XY$, $g(X,Y) = X^Y$, $g(X,Y) = \binom{X}{Y}$, *etc.*

Expectation of a function of two RVs

What's $\mathbb{E}[UW]$ for U, W from the last slide?

08/03: typo corrected

Expectation of a function of two RVs

What's $\mathbb{E}[UW]$ for U, W from the last slide?

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

$$\sum_{u \in \Omega_U} \sum_{w \in \Omega_W} uw \cdot p_{U,W}(u, w)$$

$$= 1 \cdot 1 \cdot \frac{1}{16} + 1 \cdot 2 \cdot \frac{2}{16} + 1 \cdot 3 \cdot \frac{2}{16} + 1 \cdot 4 \cdot \frac{2}{16} + 2 \cdot 2 \cdot \frac{1}{16} + 2 \cdot 3 \cdot \frac{2}{16} + 2 \cdot 4 \cdot \frac{2}{16} + 3 \cdot 3 \cdot \frac{1}{16} + 3 \cdot 4 \cdot \frac{2}{16} + 4 \cdot 4 \cdot \frac{1}{16}$$

$$= \frac{100}{16} = \frac{25}{4} = 6.25$$

08/03: typo corrected, missing term added, calculation corrected

Conditional Expectation

side tangent

Waaaaaay back when, we said conditioning on an event creates a new probability space, with all the laws holding.

So we can define things like "conditional expectations" which is the expectation of a random variable in that new probability space.

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

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

Conditional Expectations

All your favorite theorems are still true.

For example, linearity of expectation still holds

$$\underbrace{\mathbb{E}[(aX + bY + c)\,|E]} = a\underbrace{\mathbb{E}[X|E]}_{\text{condition on } E} + b\underbrace{\mathbb{E}[Y|E]} + c$$

Law of Total Expectation

Let $A_1, A_2, \ldots, A_k$ be a partition of the sample space, then

$$\boxed{\mathbb{E}[X]} = \sum_{i=1}^{n} \underbrace{\mathbb{E}[X|A_i]}_{\text{conditional exp.}} \mathbb{P}(A_i)$$

total exp.

total exp.

Let X, Y be discrete random variables, then

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

Similar in form to law of total probability, and the proof goes that way as well.

LTE

You will flip 2 (independent, fair coins). Call the number of heads X. Then (independently of the coin flips) draw an exponential random variable Y from the distribution $\mathrm{Exp}(X+1)$.

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

LTE

$\Omega_X = \{0, 1, 2\}$

You will flip 2 (independent, fair coins). Call the number of heads X. Then (independently of the coin flips) draw an exponential random variable Y from the distribution $\mathrm{Exp}(X+1)$.

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

Expect. of $X \sim \mathrm{Exp}(\lambda) = \frac{1}{\lambda}$

$\mathbb{E}[Y] =$

LTE formula

$= \mathbb{E}[Y|X=0]\mathbb{P}(X=0) + \mathbb{E}[Y|X=1]\mathbb{P}(X=1) + \mathbb{E}[Y|X=2]\mathbb{P}(X=2)$

$= \mathbb{E}[Y|X=0] \cdot \frac{1}{4} + \mathbb{E}[Y|X=1] \cdot \frac{1}{2} + \mathbb{E}[Y|X=2] \cdot \frac{1}{4}$

$= \frac{1}{0+1} \cdot \frac{1}{4} + \frac{1}{1+1} \cdot \frac{1}{2} + \frac{1}{2+1} \cdot \frac{1}{4} = \frac{7}{12}$

$Y|X=0 \sim \mathrm{Exp}(0+1)$ $\qquad$ $Y|X=1 \sim \mathrm{Exp}(1+1)$

Different dice

marginal pmf

Roll two fair dice independently. Let U be the minimum of the two rolls and W be ~~the~~ maximum

conditional pmf

$p_{U,W}$	W=1	W=2	W=3	W=4
U=1	1/16	2/16	2/16	2/16
U=2	0	1/16	2/16	2/16
U=3	0	0	1/16	2/16
U=4	0	0	0	1/16

What is $\mathbb{P}(U = 2|W = 3)$?

$$= \frac{\mathbb{P}(U=2 \cap W=3)}{\mathbb{P}(W=3)} = \frac{2/16}{5/16} = \frac{2}{5}$$

$$p_{U|W}(2|3) = \frac{2}{5}$$

$$p_{U|W}(2|3) = P(U=2|W=3)$$

Different dice

1-2 min

Find these values

$p_{W|U}(2|1) = P(W=2 | U=1)$

$= \frac{P(W=2 \cap U=1)}{P(U=1)}$

$\frac{2/16}{7/16}$

$p_{U|W}(1|2) =$

$\frac{P(U=1 \cap W=2)}{P(W=2)}$

$\frac{2/16}{3/16}$

$p_{U|W}(4|1) =$

$\frac{P(U=4 \cap W=1)}{P(W=1)} = 0$

$p_{U,W}$	$W=1$	$W=2$	$W=3$	$W=4$
$U=1$	1/16	2/16	2/16	2/16
$U=2$	0	1/16	2/16	2/16
$U=3$	0	0	1/16	2/16
$U=4$	0	0	0	1/16

Different dice

Find these values

$$p_{W|U}(2|1) = \frac{p_{W,U}(2,1)}{p_U(1)} = \frac{2/16}{7/16} = \frac{2}{7}$$

$$p_{U|W}(1|2) = \frac{p_{U,W}(1,2)}{p_W(2)} = \frac{2/16}{3/16} = \frac{2}{3}$$

$$p_{U|W}(4|1) = \frac{p_{U,W}(4,1)}{p_W(1)} = \frac{0}{1/16} = 0$$

$p_{U,W}$	W=1	W=2	W=3	W=4
U=1	1/16	2/16	2/16	2/16
U=2	0	1/16	2/16	2/16
U=3	0	0	1/16	2/16
U=4	0	0	0	1/16

Joint Distributions

For continuous joint distributions, everything will look very similar but we *replace summations with integrals*, and use a *density function instead of the PMF*

pdf

We have two continuous random variables X and Y
(that may or may not be independent)

Joint **Support/Range** - $\Omega_{X,Y}$

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

Joint **PDF** - $f_{X,Y}(a,b)$

$f_{X,Y}(a,b)$ *defined for **all*** $(a,b) \in \mathbb{R} \times \mathbb{R}$

Joint **CDF** - $F_{X,Y}(a,b)$

$F_{X,Y}(a,b) = \mathbb{P}(X \le a, Y \le b)$
*defined for **all*** $(a,b) \in \mathbb{R} \times \mathbb{R}$

Normalization Property:

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

Joint **Independence**

> $f_{X,Y}(a,b) = f_X(a) \cdot f_X(b)$ for all $(a,b) \in \Omega_{X,Y}$
> $\Omega_{X,Y} = \Omega_X \times \Omega_Y$

Joint **Expectation**

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

Marginal **PDF** - $f_X(x)$, $f_Y(y)$

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

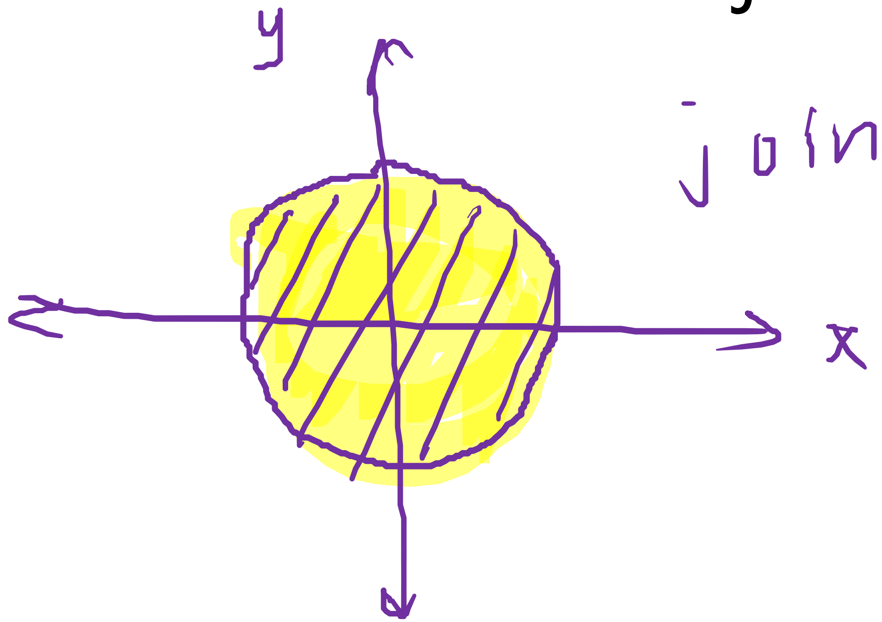
Notice we're integrating (summing) over what the other RV can be

Example: Darts

We throw a dart uniformly at random onto $\boxed{\text{a circle}}$ of radius r centered ~~around the version~~. X and Y are the x and y coordinates of the point the dart lands at.

origin

Let's find and sketch the **joint range** $\Omega_{X,Y}$



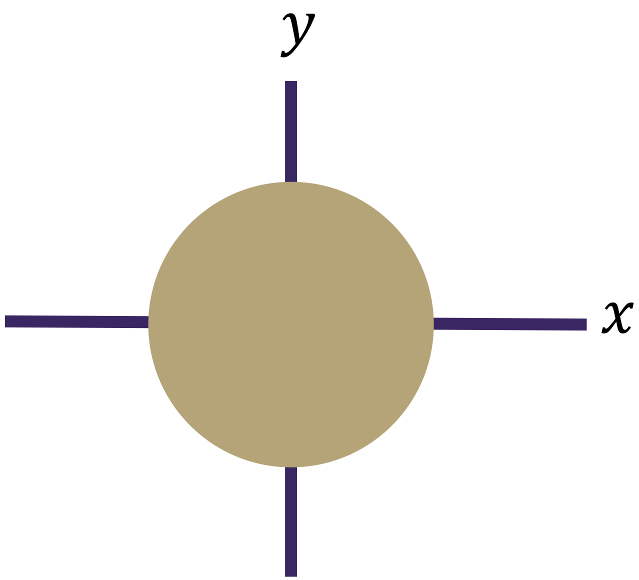
joint range

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

Example: Darts

We throw a dart uniformly at random onto a circle of radius r centered around the version. X and Y are the x and y coordinates of the point the dart lands at.

Let's find and sketch the **joint range** $\Omega_{X,Y}$



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

08/03: circle equation typo corrected to r^2

Example: Darts

joint PDF

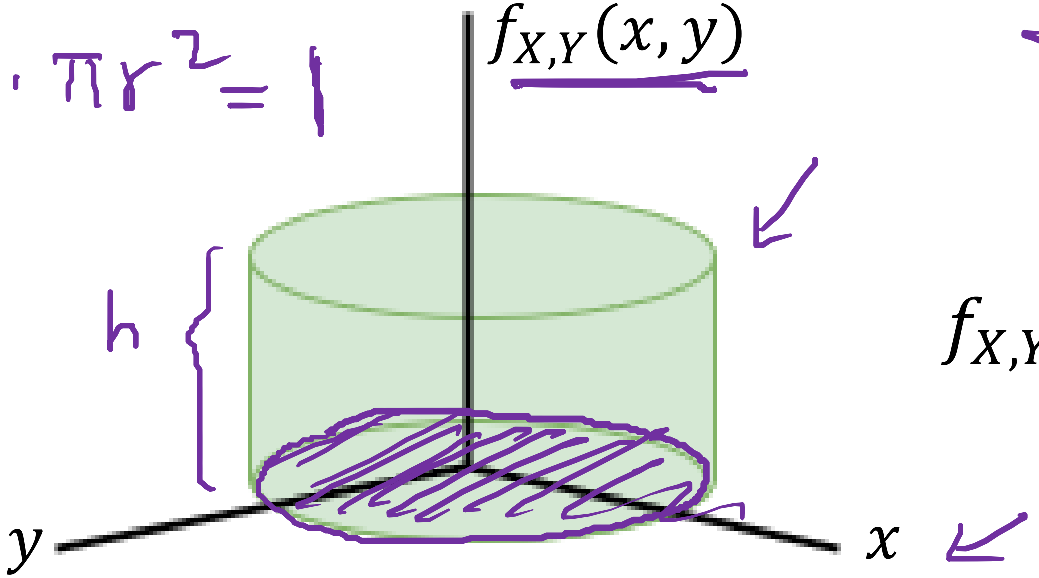
We throw a dart uniformly at random onto a circle of radius r centered around the version. X and Y are the x and y coordinates of the point the dart lands at.

normalization

integrate pdf over support = 1

Let's find and sketch the **joint PDF** $f_{X,Y}(x, y)$

$h \cdot \pi r^2 = 1$

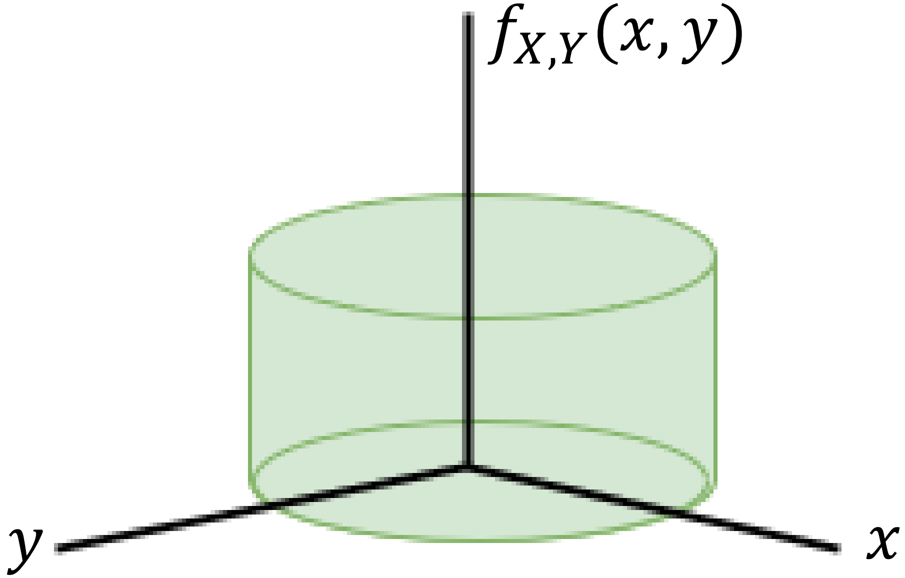


$$f_{X,Y}(x, y) = \begin{cases} \frac{1}{\pi r^2} & x^2 + y^2 \le r^2 \\ 0 & \text{o.w.} \end{cases}$$

Example: Darts

We throw a dart uniformly at random onto a circle of radius r centered around the version. X and Y are the x and y coordinates of the point the dart lands at.

Let's find and sketch the **joint PDF** $f_{X,Y}(x,y)$



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

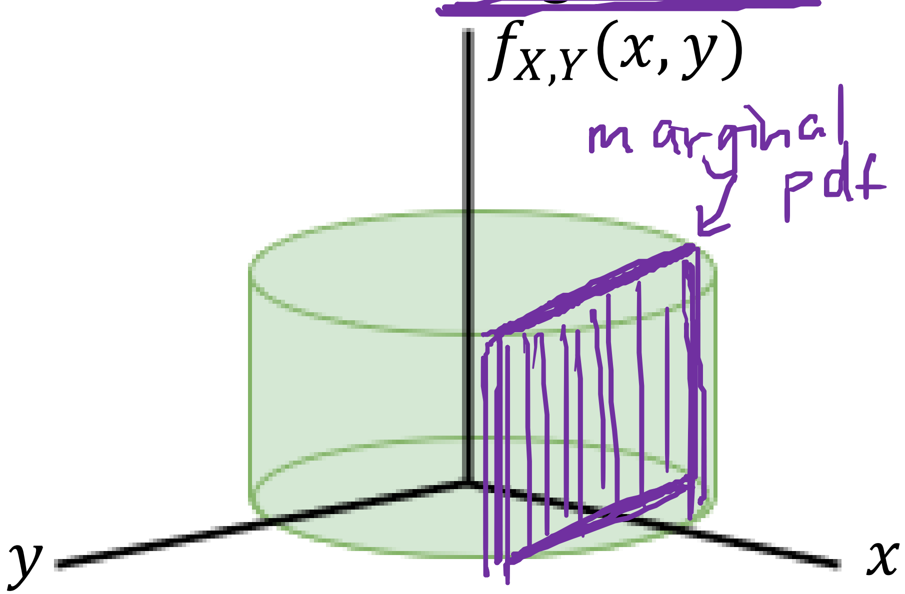
08/03: circle equation typo corrected to r^2

Example: Darts

We throw a dart uniformly at random onto a circle of radius r centered around the version. X and Y are the x and y coordinates of the point the dart lands at.

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

What is the marginal PDF for X and Y?



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

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

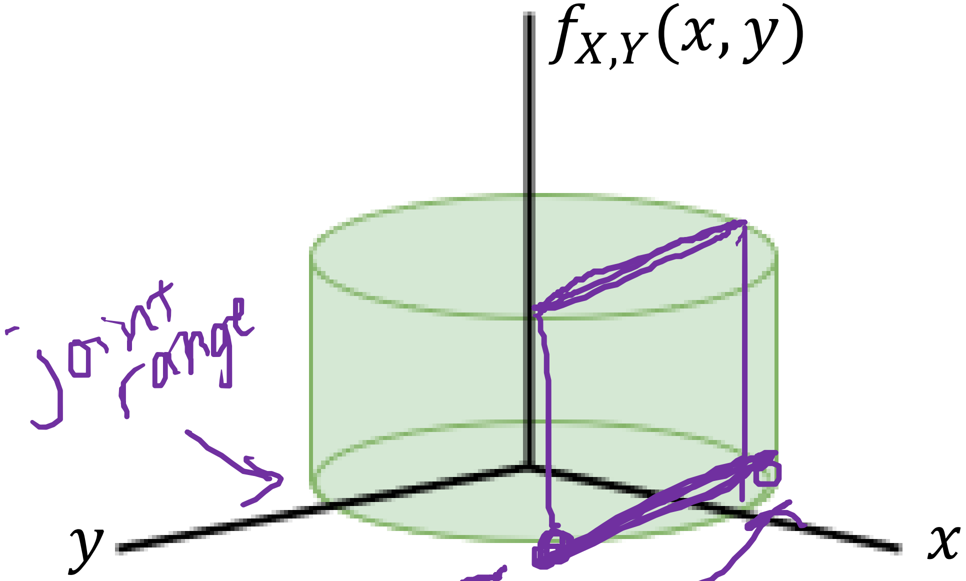
08/03: circle equation typo corrected to r^2

Example: Darts

We throw a dart uniformly at random onto a circle of radius r centered around the version. X and Y are the x and y coordinates of the point the dart lands at.

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

What is the marginal PDF for X and Y?



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

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

08/03: circle equation typo corrected to r^2

Example: Darts

We throw a dart uniformly at random onto a circle of radius r centered around the version. X and Y are the x and y coordinates of the point the dart lands at.

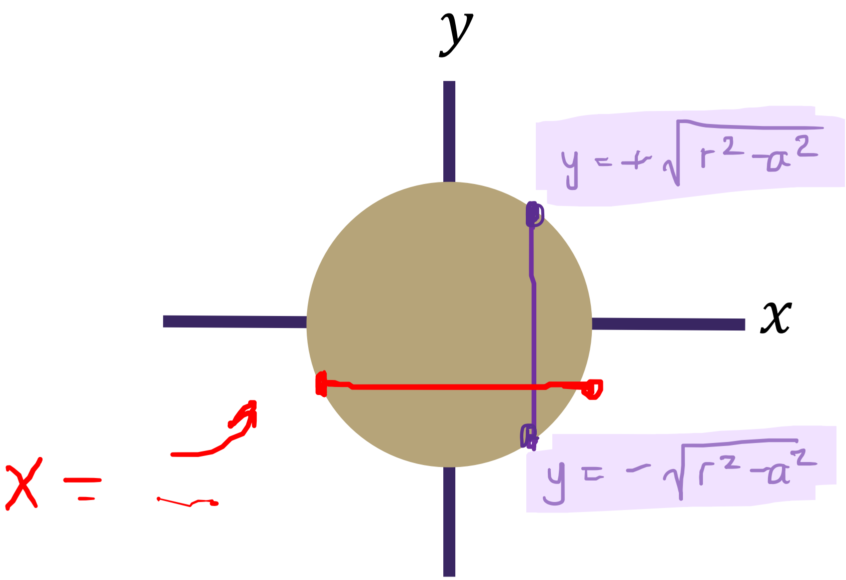
fix $x = a$

$x^2 + y^2 = r^2$

$a^2 + y^2 = r^2 \rightarrow y^2 = r^2 - a^2$

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

What is the marginal PDF for X and Y?



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

$$f_Y(b) = \int_{-\infty}^{\infty} f_{X,Y}(x,b)\,dx = \int_{-\sqrt{r^2-b^2}}^{\sqrt{r^2-b^2}} \frac{1}{\pi r^2}\,dx$$

08/03: circle equation typo corrected to r^2

Example: Darts

$\Omega_X = [-1, 1]$

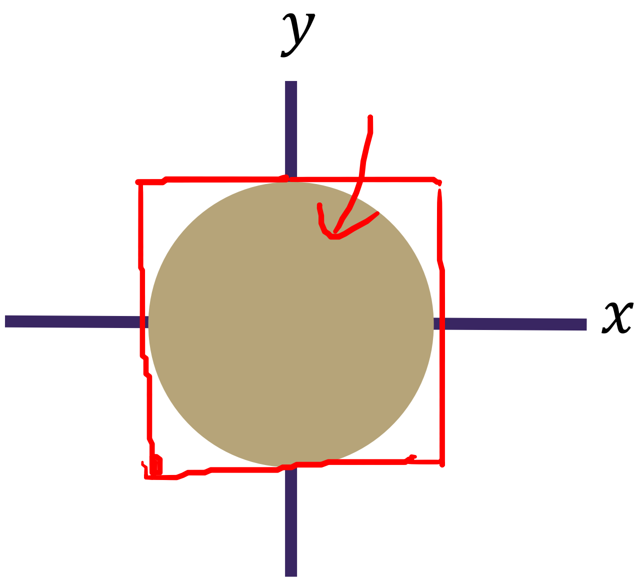
$\Omega_Y = [-1, 1]$

We throw a dart uniformly at random onto a circle of radius r centered around the version. X and Y are the x and y coordinates of the point the dart lands at.

$X =$ ___

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

What is the marginal PDF for X and Y?



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

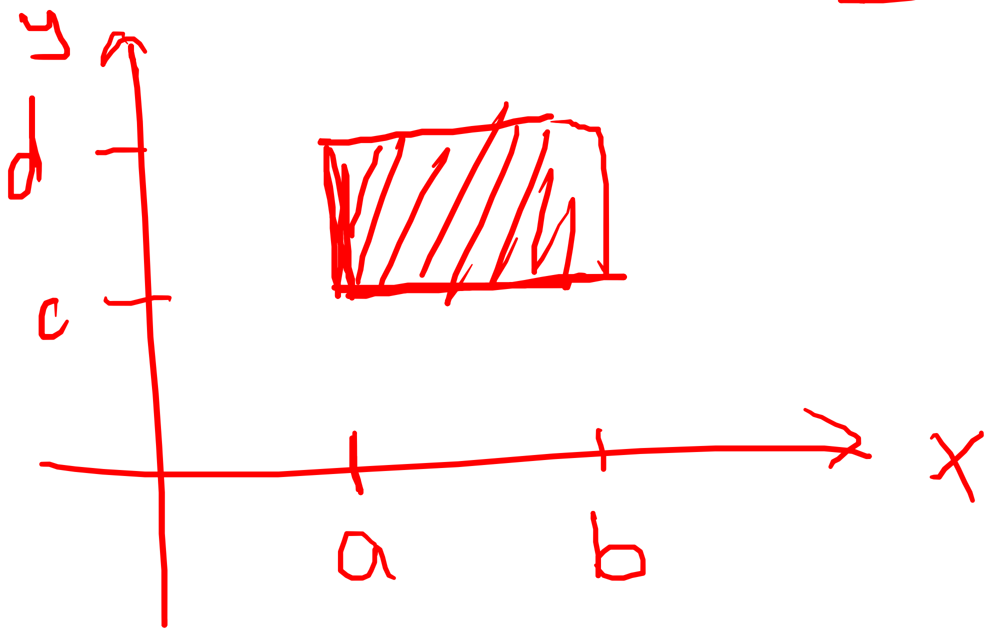
$$f_Y(b) = \int_{-\infty}^{\infty} f_{X,Y}(x,b)\,dx = \int_{-\sqrt{r^2-b^2}}^{\sqrt{r^2-b^2}} \frac{1}{\pi r^2}\,dx = \frac{2\sqrt{r^2-b^2}}{\pi r^2}$$

08/03: circle equation typo corrected to r^2

Joint Continuous Probabilities

Just like we've done with PDFs for single random variables…

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



Analogues for continuous

Everything we saw today has a continuous version.

There are "no surprises"– replace pmf with pdf and sums with integrals.

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

A note on independence

The definition of independence says X, Y independent if and only if

$p_{X,Y}(x, y) = p_X(x)p_Y(y)$ or $f_{X,Y}(x, y) = f_X(x)f_Y(y)$ (as appropriate)

There's often a nice shortcut. If X, Y are independent then joint support of X, Y (denoted $\Omega_{X,Y}$) must be $\Omega_X \times \Omega_Y$.

Joint support is $\{(x, y) : p_{X,Y}(x, y) > 0\}$.

Often easier to verify <u>dependence</u> when those are different (especially in the continuous case).

But note this is a single implication not an if-and-only-if.

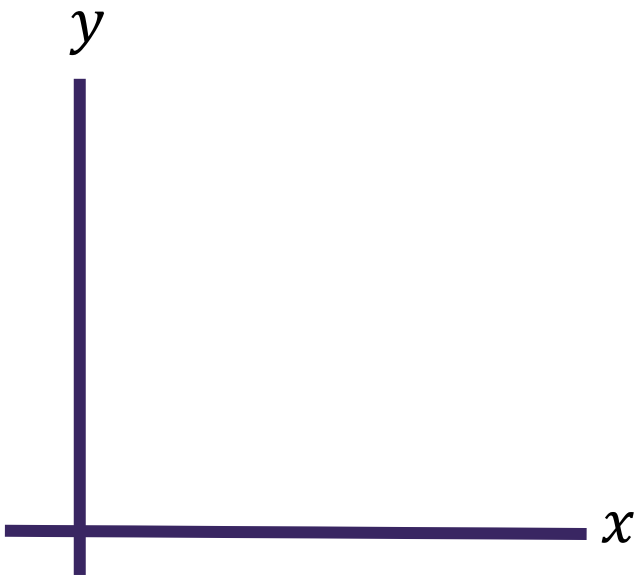
Example: Continuous Servers

The time until server 1 crashes is $X \sim \mathrm{Exp}(u)$, and the time until server 2 crashes is $Y \sim \mathrm{Exp}(v)$. Both servers are independent of each other.

What is the probability server 1 crashes before server 2?

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

pollev.con/jolie312

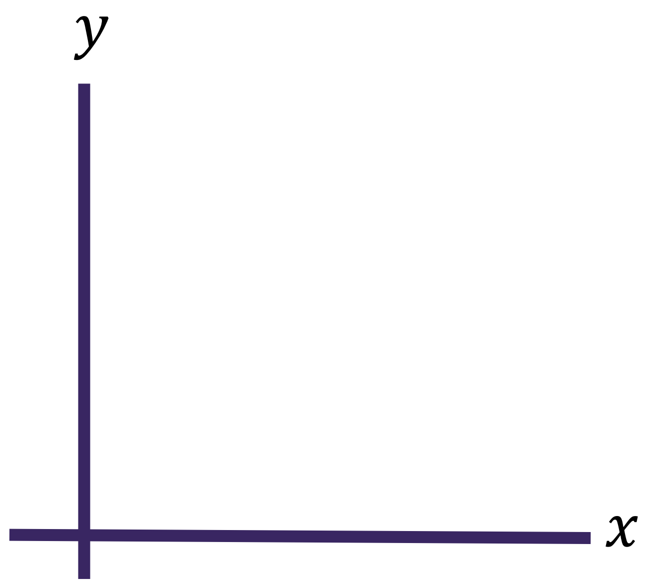


Example: Continuous Servers

The time until server 1 crashes is $X \sim \mathrm{Exp}(u)$, and the time until server 2 crashes is $Y \sim \mathrm{Exp}(v)$. Both servers are independent of each other.

What is the probability server 1 crashes before server 2?

$$\mathbb{P}(X < Y)$$
$$= \int_0^\infty \int_x^\infty f_{X,Y}(x, y)\, dy\, dx$$
$$= \int_0^\infty \int_x^\infty f_X(x) f_Y(y)\, dy\, dx \text{ by independence}$$



Covariance

Covariance

We sometimes want to measure how "intertwined" X and Y are – how much knowing about one of them will affect the other.

If X turns out "big" how likely is it that Y will be "big",
how much do they "vary together"

Covariance

$$\mathbf{Cov}(\mathbf{X}, \mathbf{Y}) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$$

Covariance

$$\mathbf{Cov}(\mathbf{X}, \mathbf{Y}) = \mathbb{E}[(\boldsymbol{X} - \mathbb{E}[\boldsymbol{X}])(\boldsymbol{Y} - \mathbb{E}[\boldsymbol{Y}])] = \mathbb{E}[\boldsymbol{XY}] - \mathbb{E}[\boldsymbol{X}]\mathbb{E}[\boldsymbol{Y}]$$

If X, Y go in the same direction

If X, Y go in the opposite directions

Covariance

$$\mathrm{Var}(X+Y)=\mathrm{Var}(X)+\mathrm{Var}(Y)+2\mathrm{Cov}(X,Y)$$

That's consistent with our previous knowledge for independent variables. (for X,Y independent, $\mathbb{E}[XY]=\mathbb{E}[X]\mathbb{E}[Y]$).

Covariance

$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\mathrm{Cov}(X,Y)$$

You and your friend are playing a game, you flip a coin: if heads you pay your friend a dollar, if tails they pay you a dollar. Let X be your profit and Y be your friend's profit.

What is $\mathrm{Var}(X+Y)$?

Before you calculate, make a prediction. What should it be?

Covariance

You and your friend are playing a game, you flip a coin: if heads you pay your friend a dollar, if tails they pay you a dollar. Let X be your profit and Y be your friend's profit.

What is $\mathrm{Var}(X+Y)$?

$\mathrm{Var}(X) = \mathrm{Var}(Y) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = 1 - 0^2 = 1$

$\mathrm{Cov}(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$

$\mathbb{E}[XY] = \frac{1}{2} \cdot (-1 \cdot 1) + \frac{1}{2}(1 \cdot -1) = -1$

$\mathrm{Cov}(X,Y) = -1 - 0 \cdot 0 = -1.$

$\mathrm{Var}(X+Y) = 1 + 1 + 2 \cdot -1 = 0$

Covariance, Another example

Let X be a Bernoulli RV with probability p of success.

Let $Y = X$ (Y **is** X, not an iid copy, literally the same experiment)

Let $Z = -X$

Let W be an independent Bernoulli, indentically distributed to X

Find

$\text{Cov}(X, Y)$, $\text{Cov}(X, Z)$, $\text{Cov}(X, W)$

Covariance, Another example

Let X be a Bernoulli RV with probability p of success.

Let $Y = X$ (Y **is** X, not an iid copy, literally the same experiment)

Let $Z = -X$

Let W be an independent Bernoulli, indentically distributed to X

$\mathrm{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$

$= (1 \cdot 1 \cdot p + 0 \cdot 0 \cdot [1 - p]) - p \cdot p$

$= p - p^2 = p(1 - p)$

Hey, that's the variance of X. This is a pattern: $\mathrm{Cov}(X, X) = \mathrm{Var}(X)$

Covariance, Another example

Let X be a Bernoulli RV with probability p of success.

Let $Y = X$ (Y **is** X, not an iid copy, literally the same experiment)

Let $Z = -X$

Let W be an independent Bernoulli, indentically distributed to X

$\text{Cov}(X, Z) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$

$= (1 \cdot -1 \cdot p + 0 \cdot -0 \cdot [1 - p]) - (p \cdot [-p])$

$= -p - [-p^2] = -p(1 - p)$

General pattern: $\text{Cov}(X, -Y) = -\text{Cov}(X, Y)$

Covariance, Another example

Let X be a Bernoulli RV with probability p of success.

Let $Y = X$ (Y **is** X, not an iid copy, literally the same experiment)

Let $Z = -X$

Let W be an independent Bernoulli, indentically distributed to X

$\text{Cov}(X, W) = \mathbb{E}[XW] - \mathbb{E}[X]\mathbb{E}[W]$

$= (1 \cdot 1 \cdot p^2 + 1 \cdot 0 \cdot p[1-p] + 0 \cdot 1 \cdot [1-p]p + 0 \cdot 0 \cdot [1-p]^2) - (p \cdot [p])$

$= (p^2) - p^2 = 0$

General pattern: if X, Y independent $\text{Cov}(X, Y) = 0$

A Few Notes

Covariance is an un-normalized number.

It measures both how intertwined X, Y are and in some sense how much X, Y vary in the first place (if you multiply both X, Y by 2, the strength of the relationship intuitively is the same, but covariance increases).

If you want just the strength of the relationship, you probably want the "correlation coefficient": $\frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}}$ always between -1 and 1.

Covariance directly measures only "linear" relationships; if Y depends on X^2, the covariance might not be as high as you expect.

If dealing with real data, look at a plot to see if you should be looking for a linear relationship in the first place.