THE UNIVERSITY OF TEXAS AT AUSTIN
Department of Statistics and Data Sciences
College of Natural Sciences

SDS 321: Introduction to Probability and Statistics

Lecture 19: Continuous random variables: Independence, covariance

Purnamrita Sarkar
Department of Statistics and Data Science
The University of Texas at Austin

www.cs.cmu.edu/∼psarkar/teaching

Roadmap

- Independence
- Covariance and correlation.
- More than two random variables.
- Function of random variables

Independent random variables-example

- You have two random variables X, Y with joint PDF

$$f_{XY}(x, y) = \begin{cases} cxy & x, y \in [0, 1] \\ 0 & \text{Otherwise} \end{cases}$$

- What is c?
- Are X, Y independent?

Independent random variables-example

- You have two random variables X, Y with joint PDF

$$f_{XY}(x,y) = \begin{cases} cxy & x, y \in [0,1] \\ 0 & \text{Otherwise} \end{cases}$$

- What is c?
- Are X, Y independent?

- $c \int_{x=0}^{1} \int_{y=0}^{1} xydydx = c/4 = 1$. So $c = 4$.

Independent random variables-example

- You have two random variables X, Y with joint PDF

$$f_{XY}(x,y) = \begin{cases} cxy & x,y \in [0,1] \\ 0 & \text{Otherwise} \end{cases}$$

- What is c?
- Are X, Y independent?

- $c\int_{x=0}^{1}\int_{y=0}^{1} xydydx = c/4 = 1$. So $c = 4$.
- $f_X(x) = \begin{cases} \int_0^1 4xydy = 2x & x \in [0,1] \\ 0 & \text{Otherwise} \end{cases}$

Independent random variables-example

- You have two random variables X, Y with joint PDF

$$f_{XY}(x,y) = \begin{cases} cxy & x,y \in [0,1] \\ 0 & \text{Otherwise} \end{cases}$$

- What is c?
- Are X, Y independent?

- $c \int_{x=0}^{1} \int_{y=0}^{1} xydydx = c/4 = 1$. So $c = 4$.
- $f_X(x) = \begin{cases} \int_0^1 4xydy = 2x & x \in [0,1] \\ 0 & \text{Otherwise} \end{cases}$
- $f_Y(y) = \begin{cases} 2y & y \in [0,1] \\ 0 & \text{Otherwise} \end{cases}$
- $f_{X,Y}(x,y) = f_X(x)f_Y(y)$ for all $x, y \in [0,1]$.

Independent random variables-example

- You have two random variables X, Y with joint PDF

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

- What is c?
- Are X, Y independent?

Independent random variables-example

- You have two random variables X, Y with joint PDF

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

- What is c?
- Are X, Y independent?

- $c\int_{x=0}^{1}\int_{y=0}^{x} xydydx = c\int_{x=0}^{1} x^3/2dx = c/8$. So $c = 8$.

Independent random variables-example

- You have two random variables X, Y with joint PDF

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

- What is c?
- Are X, Y independent?

- $c\int_{x=0}^{1}\int_{y=0}^{x} xydydx = c\int_{x=0}^{1} x^3/2dx = c/8$. So $c = 8$.
- $f_X(x) = \begin{cases} \int_0^x 8xydy = 4x^3 & x \in [0,1] \\ 0 & \text{Otherwise} \end{cases}$

Independent random variables-example

- You have two random variables X, Y with joint PDF

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

- What is c?
- Are X, Y independent?

- $c\int_{x=0}^{1}\int_{y=0}^{x} xydydx = c\int_{x=0}^{1} x^3/2dx = c/8$. So $c = 8$.
- $f_X(x) = \begin{cases} \int_0^x 8xydy = 4x^3 & x \in [0,1] \\ 0 & \text{Otherwise} \end{cases}$
- $f_Y(y) = \begin{cases} \int_y^1 8xydx = 4y(1-y^2) & y \in [0,1] \\ 0 & \text{Otherwise} \end{cases}$
- $f_{X,Y}(x,y) \neq f_X(x)f_Y(y)$ for all $x, y \in [0,1]$.

Independent random variables-example

- You have two random variables X, Y with joint PDF

$$f_{XY}(x, y) = \begin{cases} ce^{-2x}e^{-3y} & x, y \geq 0 \\ 0 & \text{Otherwise} \end{cases}$$

- What is c?
- Are X, Y independent?
- Compute $E[XY]$.

Independent random variables-example

- You have two random variables X, Y with joint PDF

$$f_{XY}(x,y) = \begin{cases} ce^{-2x}e^{-3y} & x, y \geq 0 \\ 0 & \text{Otherwise} \end{cases}$$

- What is c?
- Are X, Y independent?
- Compute $E[XY]$.

- First note that the X Y are not constrained by each other.

Independent random variables-example

- You have two random variables X, Y with joint PDF

$$f_{XY}(x,y) = \begin{cases} ce^{-2x}e^{-3y} & x, y \geq 0 \\ 0 & \text{Otherwise} \end{cases}$$

- What is c?
- Are X, Y independent?
- Compute $E[XY]$.

- First note that the X Y are not constrained by each other.
- Next note that $e^{-2x}e^{-3y}$ is basically the product of a function of x and a function of y. If I gave you $f_{X,Y}(x,y) = c(x+y)$, then its not a product of a function of x and a function of y. But in the given problem, X, Y are independent.

Independent random variables-example

- You have two random variables X, Y with joint PDF

$$f_{XY}(x, y) = \begin{cases} ce^{-2x}e^{-3y} & x, y \geq 0 \\ 0 & \text{Otherwise} \end{cases}$$

- What is c?
- Are X, Y independent?
- Compute $E[XY]$.

- First note that the X Y are not constrained by each other.
- Next note that $e^{-2x}e^{-3y}$ is basically the product of a function of x and a function of y. If I gave you $f_{X,Y}(x, y) = c(x + y)$, then its not a product of a function of x and a function of y. But in the given problem, X, Y are independent.
- Next look at the two bits, for x we have e^{-2x}. So this is sort of like an exponential. So $f_X(x) = 2e^{-2x}$ and $f_Y(y) = 3e^{-3x}$.

Independent random variables-example

- You have two random variables X, Y with joint PDF

$$f_{XY}(x,y) = \begin{cases} ce^{-2x}e^{-3y} & x, y \geq 0 \\ 0 & \text{Otherwise} \end{cases}$$

- What is c?
- Are X, Y independent?
- Compute $E[XY]$.

- First note that the X Y are not constrained by each other.
- Next note that $e^{-2x}e^{-3y}$ is basically the product of a function of x and a function of y. If I gave you $f_{X,Y}(x,y) = c(x+y)$, then its not a product of a function of x and a function of y. But in the given problem, X, Y are independent.
- Next look at the two bits, for x we have e^{-2x}. So this is sort of like an exponential. So $f_X(x) = 2e^{-2x}$ and $f_Y(y) = 3e^{-3x}$.
- $c = 6$.

Independent random variables-example

- You have two random variables X, Y with joint PDF

$$f_{XY}(x,y) = \begin{cases} ce^{-2x}e^{-3y} & x, y \geq 0 \\ 0 & \text{Otherwise} \end{cases}$$

- What is c?
- Are X, Y independent?
- Compute $E[XY]$.

- First note that the X Y are not constrained by each other.
- Next note that $e^{-2x}e^{-3y}$ is basically the product of a function of x and a function of y. If I gave you $f_{X,Y}(x,y) = c(x+y)$, then its not a product of a function of x and a function of y. But in the given problem, X, Y are independent.
- Next look at the two bits, for x we have e^{-2x}. So this is sort of like an exponential. So $f_X(x) = 2e^{-2x}$ and $f_Y(y) = 3e^{-3x}$.
- $c = 6$.
- So $E[X] = 1/2$ and $E[Y] = 1/3$ and via independence $E[XY] = 1/6$

Independent random variables-example

- You have two random variables X, Y with joint PDF

$$f_{XY}(x,y) = \begin{cases} ce^{-2x}e^{-y^2/2} & x \geq 0, -\infty \leq y \leq \infty \\ 0 & \text{Otherwise} \end{cases}$$

- What is c?
- Are X, Y independent?
- Compute $E[X], E[Y]$.

Independent random variables-example

- You have two random variables X, Y with joint PDF

$$f_{XY}(x,y) = \begin{cases} ce^{-2x}e^{-y^2/2} & x \geq 0, -\infty \leq y \leq \infty \\ 0 & \text{Otherwise} \end{cases}$$

- What is c?
- Are X, Y independent?
- Compute $E[X], E[Y]$.

- Yes, they are independent.
- $f_X(x) = 2e^{-2x}$, i.e. $X \sim \mathit{Exponential}(2)$

Independent random variables-example

- You have two random variables X, Y with joint PDF

$$f_{XY}(x,y) = \begin{cases} ce^{-2x}e^{-y^2/2} & x \geq 0, -\infty \leq y \leq \infty \\ 0 & \text{Otherwise} \end{cases}$$

- What is c?
- Are X, Y independent?
- Compute $E[X], E[Y]$.

- Yes, they are independent.
- $f_X(x) = 2e^{-2x}$, i.e. $X \sim \mathit{Exponential}(2)$
- $f_Y(y) = 1/\sqrt{2\pi}e^{-y^2/2}$, i.e. $Y \sim N(0,1)$

Independent random variables-example

- You have two random variables X, Y with joint PDF

$$f_{XY}(x,y) = \begin{cases} ce^{-2x}e^{-y^2/2} & x \geq 0, -\infty \leq y \leq \infty \\ 0 & \text{Otherwise} \end{cases}$$

- What is c?
- Are X, Y independent?
- Compute $E[X], E[Y]$.

- Yes, they are independent.
- $f_X(x) = 2e^{-2x}$, i.e. $X \sim \mathit{Exponential}(2)$
- $f_Y(y) = 1/\sqrt{2\pi}e^{-y^2/2}$, i.e. $Y \sim N(0,1)$
- $c = 2/\sqrt{2\pi}$.

Independent random variables-example

- You have two random variables X, Y with joint PDF

$$f_{XY}(x,y) = \begin{cases} ce^{-2x}e^{-y^2/2} & x \geq 0, -\infty \leq y \leq \infty \\ 0 & \text{Otherwise} \end{cases}$$

- What is c?
- Are X, Y independent?
- Compute $E[X], E[Y]$.

- Yes, they are independent.
- $f_X(x) = 2e^{-2x}$, i.e. $X \sim \mathit{Exponential}(2)$
- $f_Y(y) = 1/\sqrt{2\pi}e^{-y^2/2}$, i.e. $Y \sim N(0,1)$
- $c = 2/\sqrt{2\pi}$.
- $E[X] = 1/2$ and $E[Y] = 0$.

Covariance

- The **covariance** of two random variables X and Y is given by

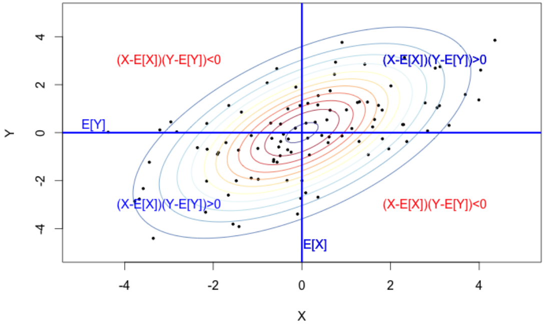
$$\text{cov}(X, Y) = E\left[(X - E[X])(Y - E[Y])\right]$$

- We can simplify this a little

$$\begin{aligned}\text{cov}(X, Y) =& E[(X - E[X])(Y - E[Y])] \\ =& E[XY - XE[Y] - YE[X] + E[X]E[Y]] \\ =& E[XY] - E[X]E[Y] - E[X]E[Y] + E[X]E[Y] \\ =& E[XY] - E[X][E[Y]\end{aligned}$$

- It is a measure of how much X and Y change together.
- A **positive** covariance means that, if $X > E[X]$, we are likely to have $Y > E[Y]$
- A **negative** covariance means that, if $X > E[X]$, we are likely to have $Y < E[Y]$.

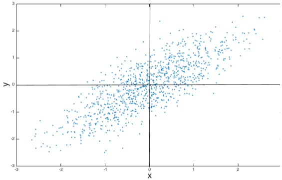
Covariance

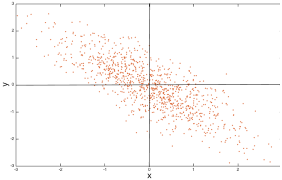


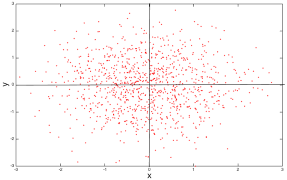
- A **positive** covariance means that we have most mass in the upper right and lower left quadrants.
- A **negative** covariance means that we have most mass in the upper left and lower right quadrants.
- A **zero** covariance means that we have about an equal mass in the upper left and upper right quadrants.

Covariance

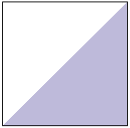
- We are plotting two random variables X and Y below. Which one corresponds to a positive, negative or zero covariance?





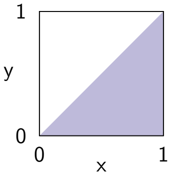


Example: Continuous case



- Let $f_{X,Y}(x,y) = \begin{cases} 2 & 0 \le y \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$
- What is $\text{cov}(X, Y)$?

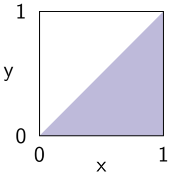
Example: Continuous case



- Let $f_{X,Y}(x,y) = \begin{cases} 2 & 0 \leq y \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$
- What is $\text{cov}(X, Y)$?
  - We first need to calculate the expectation of X and Y.

Example: Continuous case

- Let $f_{X,Y}(x,y) = \begin{cases} 2 & 0 \le y \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$
- What is $\mathrm{cov}(X,Y)$?



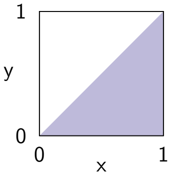
- We first need to calculate the expectation of X and Y.
- The marginal PDF of X and Y are:

$$f_X(x) =$$

$$f_Y(y) =$$

Example: Continuous case

- Let $f_{X,Y}(x,y) = \begin{cases} 2 & 0 \leq y \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$
- What is $\text{cov}(X, Y)$?



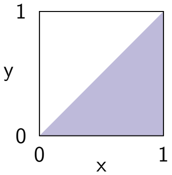
- We first need to calculate the expectation of X and Y.
- The marginal PDF of X and Y are:

$$f_X(x) = \int_0^1 f_{X,Y}(x,y)dy = 2\int_0^x dy = 2x \qquad 0 \leq x \leq 1$$

$$f_Y(y) =$$

Example: Continuous case

- Let $f_{X,Y}(x,y) = \begin{cases} 2 & 0 \le y \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$
- What is cov(X, Y)?



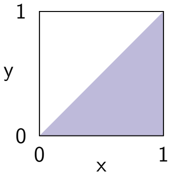
- We first need to calculate the expectation of X and Y.
- The marginal PDF of X and Y are:

$$f_X(x) = \int_0^1 f_{X,Y}(x,y)dy = 2\int_0^x dy = 2x \qquad 0 \le x \le 1$$

$$f_Y(y) = \int_0^1 f_{X,Y}(x,y)dx = 2\int_y^1 dx = 2(1-y) \qquad 0 \le y \le 1$$

Example: Continuous case

- Let $f_{X,Y}(x,y) = \begin{cases} 2 & 0 \le y \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$
- What is $\mathrm{cov}(X,Y)$?



- We first need to calculate the expectation of X and Y.
- The marginal PDF of X and Y are:

$$f_X(x) = \int_0^1 f_{X,Y}(x,y)dy = 2\int_0^x dy = 2x \qquad 0 \le x \le 1$$

$$f_Y(y) = \int_0^1 f_{X,Y}(x,y)dx = 2\int_y^1 dx = 2(1-y) \qquad 0 \le y \le 1$$

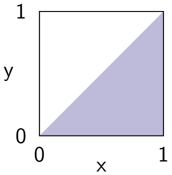
- So the expectations are:

$$E[X] =$$

$$E[Y] =$$

Example: Continuous case

- Let $f_{X,Y}(x,y) = \begin{cases} 2 & 0 \le y \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$
- What is $\text{cov}(X, Y)$?



- We first need to calculate the expectation of X and Y.
- The marginal PDF of X and Y are:

$$f_X(x) = \int_0^1 f_{X,Y}(x,y)dy = 2\int_0^x dy = 2x \qquad 0 \le x \le 1$$

$$f_Y(y) = \int_0^1 f_{X,Y}(x,y)dx = 2\int_y^1 dx = 2(1-y) \qquad 0 \le y \le 1$$

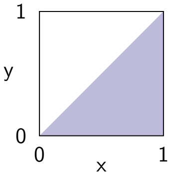
- So the expectations are:

$$E[X] = \int_0^1 2x^2 dx = 2/3$$

$$E[Y] =$$

Example: Continuous case

- Let $f_{X,Y}(x,y) = \begin{cases} 2 & 0 \le y \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$
- What is $\text{cov}(X, Y)$?



- We first need to calculate the expectation of X and Y.
- The marginal PDF of X and Y are:

$$f_X(x) = \int_0^1 f_{X,Y}(x,y)dy = 2\int_0^x dy = 2x \qquad 0 \le x \le 1$$

$$f_Y(y) = \int_0^1 f_{X,Y}(x,y)dx = 2\int_y^1 dx = 2(1-y) \qquad 0 \le y \le 1$$

- So the expectations are:

$$E[X] = \int_0^1 2x^2 dx = 2/3$$

$$E[Y] = \int_0^1 (2y - 2y^2)dy = 1/3$$

Example: Continuous case

- We next need to calculate $E[XY]$.
- This is just the expectation of a function of two random variables

$$E[XY] = \int_0^1 \int_0^1 xy f_{X,Y}(x,y)\,dx\,dy$$

Example: Continuous case

- We next need to calculate $E[XY]$.
- This is just the expectation of a function of two random variables

$$
\begin{aligned}
E[XY] &= \int_0^1 \int_0^1 xyf_{X,Y}(x,y)dxdy \\
&= \int_{x=0}^1 \int_{y=0}^x 2xydx\,dy
\end{aligned}
$$

Example: Continuous case

- We next need to calculate $E[XY]$.
- This is just the expectation of a function of two random variables

$$\begin{aligned} E[XY] &= \int_0^1 \int_0^1 xyf_{X,Y}(x,y)dxdy \\ &= \int_{x=0}^1 \int_{y=0}^x 2xydx\,dy \\ &= \int_{x=0}^1 x(\int_{y=0}^x 2ydy)\,dx \end{aligned}$$

Example: Continuous case

- We next need to calculate $E[XY]$.
- This is just the expectation of a function of two random variables

$$
\begin{aligned}
E[XY] &= \int_0^1 \int_0^1 xyf_{X,Y}(x,y)dxdy \\
&= \int_{x=0}^1 \int_{y=0}^x 2xydx\,dy \\
&= \int_{x=0}^1 x(\int_{y=0}^x 2ydy)\,dx \\
&= \int_{x=0}^1 x\left[y^2\right]_0^x dx = \int_{x=0}^1 x^3 dx = 1/4
\end{aligned}
$$

Example: Continuous case

- We next need to calculate $E[XY]$.
- This is just the expectation of a function of two random variables

$$
\begin{aligned}
E[XY] &= \int_0^1 \int_0^1 xyf_{X,Y}(x,y)dxdy \\
&= \int_{x=0}^1 \int_{y=0}^x 2xydx\,dy \\
&= \int_{x=0}^1 x(\int_{y=0}^x 2ydy)\,dx \\
&= \int_{x=0}^1 x\left[y^2\right]_0^x dx = \int_{x=0}^1 x^3 dx = 1/4
\end{aligned}
$$

- So, $\text{cov}(X,Y) = E[XY] - E[X]E[Y] = \frac{1}{4} - \frac{1}{3}\cdot\frac{2}{3} = \frac{1}{36}$

Covariance and Independence

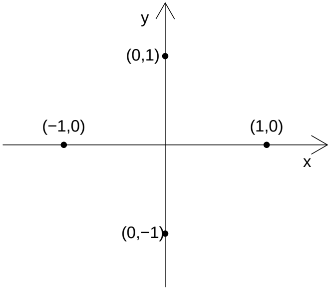
- If two random variables are **independent**, knowing one tells us nothing about the other!
- In this case, $E[XY] = E[X]E[Y]$
- We know that $\text{cov}(X, Y) = E[XY] - E[X]E[Y]$... so if two random variables are independent, their covariance is zero.
- This shouldn't be surprising... we know X can't tell us anything about Y.
- What about the converse? If $\text{cov}(X, Y) = 0$, does that mean that X and Y are independent?

Covariance and Independence

- If two random variables are **independent**, knowing one tells us nothing about the other!
- In this case, $E[XY] = E[X]E[Y]$
- We know that $\text{cov}(X, Y) = E[XY] - E[X]E[Y]$... so if two random variables are independent, their covariance is zero.
- This shouldn't be surprising... we know X can't tell us anything about Y.
- What about the converse? If $\text{cov}(X, Y) = 0$, does that mean that X and Y are independent?
- Another way of asking this is, does $E[XY] = E[X]E[Y]$ imply X and Y are independent?

Covariance and Independence

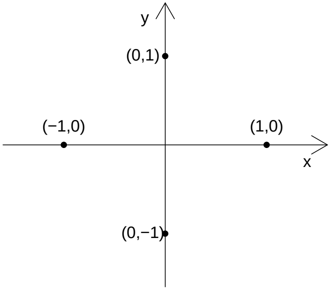
- I start at co-ordinates (0,0). I pick a compass direction (N,S,E,W) uniformly at random, and walk 1 unit in that direction.
- Let (X, Y) be my new coordinates. My sample space is $\{(0,1),(1,0),(0,-1),(-1,0)\}$.



- What are $E[X]$ and $E[Y]$?

Covariance and Independence

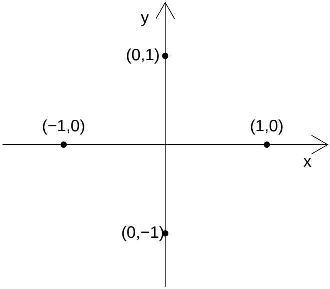
- I start at co-ordinates (0,0). I pick a compass direction (N,S,E,W) uniformly at random, and walk 1 unit in that direction.
- Let (X, Y) be my new coordinates. My sample space is $\{(0,1),(1,0),(0,-1),(-1,0)\}$.



- What are $E[X]$ and $E[Y]$?

Covariance and Independence

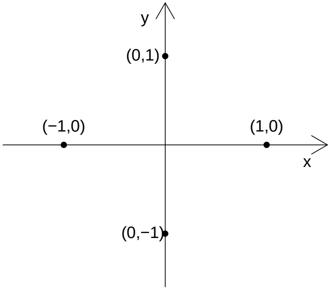
- I start at co-ordinates (0,0). I pick a compass direction (N,S,E,W) uniformly at random, and walk 1 unit in that direction.
- Let (X, Y) be my new coordinates. My sample space is $\{(0,1),(1,0),(0,-1),(-1,0)\}$.



- What are $E[X]$ and $E[Y]$? 0.

Covariance and Independence

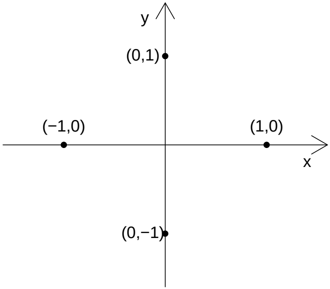
- I start at co-ordinates (0,0). I pick a compass direction (N,S,E,W) uniformly at random, and walk 1 unit in that direction.
- Let (X, Y) be my new coordinates. My sample space is $\{(0,1),(1,0),(0,-1),(-1,0)\}$.



- What are $E[X]$ and $E[Y]$? 0.
- $XY =$

Covariance and Independence

- I start at co-ordinates (0,0). I pick a compass direction (N,S,E,W) uniformly at random, and walk 1 unit in that direction.
- Let (X, Y) be my new coordinates. My sample space is $\{(0,1),(1,0),(0,-1),(-1,0)\}$.



- What are $E[X]$ and $E[Y]$? 0.
- $XY = 0$.
- So, $\text{cov}(X, Y) = 0$.
- But, if I know $X = 1$, then I *must* have $Y = 0$. So, they are not independent!

Independence implies zero correlation... but zero correlation does not imply independence!

Properties of covariance

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

Properties of covariance

- What is the covariance of X and X?

Properties of covariance

- What is the covariance of X and X?
- $\text{cov}(X, X) = \text{var}(X)$

Properties of covariance

- What is the covariance of X and a for some constant a?

Properties of covariance

- What is the covariance of X and a for some constant a?
- $$\begin{aligned} \text{cov}(X, a) &= E[Xa] - E[X]E[a] \\ &= aE[X] - aE[X] = 0 \end{aligned}$$

Properties of covariance

- What is the covariance of X and aY?

Properties of covariance

- What is the covariance of X and aY?
- $\text{cov}(X, aY) = E[aXY] - E[X]E[aY] = a\text{cov}(X, Y)$

Properties of covariance

- What is the covariance of X and $Y + b$?

Properties of covariance

- What is the covariance of X and $Y + b$?
$$\text{cov}(X, Y + b) = E[X(Y + b)] - E[X]E[Y + b]$$
- $$= E[XY] + bE[X] - (E[X]E[Y] + bE[X])$$
$$= \text{cov}(X, Y)$$

Properties of covariance

- What is the covariance of $X_1 + X_2$ and $Y_1 + Y_2$?

Properties of covariance

- What is the covariance of $X_1 + X_2$ and $Y_1 + Y_2$?
- $$\begin{aligned}\text{cov}(X_1 + X_2, Y_1 + Y_2) &= E[(X_1 + X_2)(Y_1 + Y_2)] - E[X_1 + X_2]E[Y_1 + Y_2] \\ &= E[X_1 Y_1] + E[X_1 Y_2] + E[X_2 Y_1] + E[X_2 Y_2] \\ &\quad - (E[X_1]E[Y_1] + E[X_1]E[Y_2] + E[X_2]E[Y_1] + E[X_2]E[Y_2]) \\ &= (E[X_1 Y_1] - E[X_1]E[Y_1]) + (E[X_1 Y_2] - E[X_1]E[Y_2]) \\ &\quad + (E[X_2 Y_1] - E[X_2]E[Y_1]) + (E[X_2 Y_2] - E[X_2]E[Y_2]) \\ &= \text{cov}(X_1, Y_1) + \text{cov}(X_1, Y_2) + \text{cov}(X_2, Y_1) + \text{cov}(X_2, Y_2)\end{aligned}$$

Properties of covariance

- What is the covariance of $\sum_{i=1}^{n} a_i X_i$ and $\sum_{j=1}^{n} b_j Y_j$?

Properties of covariance

- What is the covariance of $\sum_{i=1}^{n} a_i X_i$ and $\sum_{j=1}^{n} b_j Y_j$?
- $\mathrm{cov}(\sum_{i=1}^{n} a_i X_i, \sum_{j=1}^{n} b_j Y_j) = \sum_{i,j} a_i b_j \mathrm{cov}(X_i, Y_j)$

Correlation

- We know that the sign of a covariance indicates whether $X - E[X]$ and $Y - E[Y]$ tend to have the same sign.
- The magnitude gives us some indication of the extent to which this is true... but it is hard to interpret.
  - The magnitude depends not just how much X and Y co-vary, but also on how much X and Y deviate from their expected values.
- The **correlation coefficient** $\rho_{X,Y}$(sometimes referred to as the Pearson's correlation coefficient) is a standardized version of the covariance.

$$\rho_{X,Y} = \frac{\text{cov}(X,Y)}{\sqrt{\text{var}(X)\text{var}(Y)}}$$

- We always have $-1 \leq \rho_{X,Y} \leq 1$
  - $\rho = 0$ implies zero covariance.
  - $|\rho| = 1$ iff there is a linear relationship between X and Y.

Correlation: Example of $|\rho| = 1$

- We throw a biased coin, with probability of heads p, n times. Let X be the number of heads, and let Y be the number of tails.
- $Y = n - X$
- $E[X] =$

Correlation: Example of $|\rho| = 1$

- We throw a biased coin, with probability of heads p, n times. Let X be the number of heads, and let Y be the number of tails.
- $Y = n - X$
- $E[X] = np$, and $E[Y] = n(1 - p) = n - E[X]$.
- $\text{var}(X) = np(1 - p) = \text{var}(Y)$.

Correlation: Example of $|\rho| = 1$

- We throw a biased coin, with probability of heads p, n times. Let X be the number of heads, and let Y be the number of tails.
- $Y = n - X$
- $E[X] = np$, and $E[Y] = n(1 - p) = n - E[X]$.
- $\text{var}(X) = np(1 - p) = \text{var}(Y)$.
- $\text{cov}(X, Y) = \text{cov}(X, n - X) = \text{cov}(X, -X) = -\text{cov}(X, X) = -\text{var}(X)$

Correlation: Example of $|\rho| = 1$

- We throw a biased coin, with probability of heads p, n times. Let X be the number of heads, and let Y be the number of tails.
- $Y = n - X$
- $E[X] = np$, and $E[Y] = n(1-p) = n - E[X]$.
- $\text{var}(X) = np(1-p) = \text{var}(Y)$.
- $\text{cov}(X, Y) = \text{cov}(X, n - X) = \text{cov}(X, -X) = -\text{cov}(X, X) = -\text{var}(X)$
- The correlation coefficient is therefore

$$\rho_{X,Y} = \frac{\text{cov}(X, Y)}{\sqrt{\text{var}(X)\text{var}(Y)}} = \frac{-\text{var}(X)}{\sqrt{\text{var}(X)\text{var}(X)}} = -1$$

Correlation: Example of $|\rho| = 1$

- We throw a biased coin, with probability of heads p, n times. Let X be the number of heads, and let Y be the number of tails.
- $Y = n - X$
- $E[X] = np$, and $E[Y] = n(1-p) = n - E[X]$.
- $\text{var}(X) = np(1-p) = \text{var}(Y)$.
- $\text{cov}(X, Y) = \text{cov}(X, n - X) = \text{cov}(X, -X) = -\text{cov}(X, X) = -\text{var}(X)$
- The correlation coefficient is therefore

$$\rho_{X,Y} = \frac{\text{cov}(X, Y)}{\sqrt{\text{var}(X)\text{var}(Y)}} = \frac{-\text{var}(X)}{\sqrt{\text{var}(X)\text{var}(X)}} = -1$$

Correlation: Example of $|\rho| = 1$

- We throw a biased coin, with probability of heads p, n times. Let X be the number of heads, and let Y be the number of tails.
- $Y = n - X$
- $E[X] = np$, and $E[Y] = n(1 - p) = n - E[X]$.
- $\text{var}(X) = np(1 - p) = \text{var}(Y)$.
- $\text{cov}(X, Y) = \text{cov}(X, n - X) = \text{cov}(X, -X) = -\text{cov}(X, X) = -\text{var}(X)$
- The correlation coefficient is therefore

$$\rho_{X,Y} = \frac{\text{cov}(X, Y)}{\sqrt{\text{var}(X)\text{var}(Y)}} = \frac{-\text{var}(X)}{\sqrt{\text{var}(X)\text{var}(X)}} = -1$$

Correlation: Example of $|\rho| = 1$

- We throw a biased coin, with probability of heads p, n times. Let X be the number of heads, and let Y be the number of tails.
- $Y = n - X$
- $E[X] = np$, and $E[Y] = n(1 - p) = n - E[X]$.
- $\text{var}(X) = np(1 - p) = \text{var}(Y)$.
- $\text{cov}(X, Y) = \text{cov}(X, n - X) = \text{cov}(X, -X) = -\text{cov}(X, X) = -\text{var}(X)$
- The correlation coefficient is therefore

$$\rho_{X,Y} = \frac{\text{cov}(X, Y)}{\sqrt{\text{var}(X)\text{var}(Y)}} = \frac{-\text{var}(X)}{\sqrt{\text{var}(X)\text{var}(X)}} = -1$$

- Remember $X = n - Y$, so they have a linear relationship.

Correlation: Example of $|\rho| = 1$

- Let $Y = aX + b$.

Correlation: Example of $|\rho| = 1$

- Let $Y = aX + b$.
- $$\text{cov}(X, aX + b) = \text{cov}(X, aX) = a\text{var}(X)$$
$$\rho_{X,aX+b} = \frac{\text{cov}(X, aX)}{\sqrt{\text{var}(X)a^2\text{var}(X)}} = \frac{a\text{var}(X)}{|a|\text{var}(X)} = \frac{a}{|a|}$$
$$= \begin{cases} 1 & a > 0 \\ -1 & a < 0 \end{cases}$$

Variance of a sum of random variables

- Earlier in the course, we looked at the variance of the sum of discrete random variables. Same rule holds for continuous random variables.

$$\text{var}(X + Y) = \text{var}(X) + \text{var}(Y) + 2\text{cov}(X, Y)$$

- So for independent X, Y, $\text{var}(X + Y) = \text{var}(X) + \text{var}(Y)$

More than two random variables

- For multiple random variables we have:
  $P((X,Y,Z) \in B) = \int_{(x,y,z)\in B)} f_{X,Y,Z}(x,y,z)dxdydz$
- Marginalization: $f_{X,Y}(x,y) =$

More than two random variables

- For multiple random variables we have:
  $P((X, Y, Z) \in B) = \int_{(x,y,z) \in B} f_{X,Y,Z}(x, y, z)dxdydz$
- Marginalization: $f_{X,Y}(x, y) = \int f_{X,Y,Z}(x, y, z)dz$

More than two random variables

- For multiple random variables we have: $P((X,Y,Z) \in B) = \int_{(x,y,z)\in B)} f_{X,Y,Z}(x,y,z) dxdydz$
- Marginalization: $f_{X,Y}(x,y) = \int f_{X,Y,Z}(x,y,z)dz$
- Marginalization: $f_X(x) = \int\int f_{X,Y,Z}(x,y,z)dy\ dz$

More than two random variables

- For multiple random variables we have:
  $P((X, Y, Z) \in B) = \int_{(x,y,z) \in B)} f_{X,Y,Z}(x, y, z) dxdydz$
- Marginalization: $f_{X,Y}(x, y) = \int f_{X,Y,Z}(x, y, z)dz$
- Marginalization: $f_X(x) = \int \int f_{X,Y,Z}(x, y, z)dy \; dz$
- Conditional PDF: $f_{X,Y|Z}(x, y|z) = \frac{f_{X,Y,Z}(x, y, z)}{f_Z(z)}$, For $f_Z(z) > 0$

More than two random variables

- For multiple random variables we have:
  $P((X,Y,Z) \in B) = \int_{(x,y,z)\in B)} f_{X,Y,Z}(x,y,z)dxdydz$
- Marginalization: $f_{X,Y}(x,y) = \int f_{X,Y,Z}(x,y,z)dz$
- Marginalization: $f_X(x) = \int \int f_{X,Y,Z}(x,y,z)dy\ dz$
- Conditional PDF: $f_{X,Y|Z}(x,y|z) = \frac{f_{X,Y,Z}(x,y,z)}{f_Z(z)}$, For $f_Z(z) > 0$
- Conditional PDF: $f_{X|Y,Z}(x|y,z) = \frac{f_{X,Y,Z}(x,y,z)}{f_{Y,Z}(y,z)}$, For $f_{Y,Z}(y,z) > 0$

More than two random variables

- For multiple random variables we have:
  $P((X, Y, Z) \in B) = \int_{(x,y,z)\in B)} f_{X,Y,Z}(x, y, z) dxdydz$
- Marginalization: $f_{X,Y}(x, y) = \int f_{X,Y,Z}(x, y, z) dz$
- Marginalization: $f_X(x) = \int \int f_{X,Y,Z}(x, y, z) dy \ dz$
- Conditional PDF: $f_{X,Y|Z}(x, y|z) = \frac{f_{X,Y,Z}(x, y, z)}{f_Z(z)}$, For $f_Z(z) > 0$
- Conditional PDF: $f_{X|Y,Z}(x|y, z) = \frac{f_{X,Y,Z}(x, y, z)}{f_{Y,Z}(y, z)}$, For $f_{Y,Z}(y, z) > 0$
- Multiplication rule:
  $f_{X,Y,Z}(x, y, z) = f_{X|Y,Z}(x|y, z) f_{Y|Z}(y|z) f_Z(z)$, For $f_{Y,Z}(y, z) > 0$

More than two random variables

- For multiple random variables we have:
  $P((X, Y, Z) \in B) = \int_{(x,y,z)\in B)} f_{X,Y,Z}(x, y, z) dxdydz$
- Marginalization: $f_{X,Y}(x, y) = \int f_{X,Y,Z}(x, y, z) dz$
- Marginalization: $f_X(x) = \int \int f_{X,Y,Z}(x, y, z) dy \; dz$
- Conditional PDF: $f_{X,Y|Z}(x, y|z) = \frac{f_{X,Y,Z}(x, y, z)}{f_Z(z)}$, For $f_Z(z) > 0$
- Conditional PDF: $f_{X|Y,Z}(x|y, z) = \frac{f_{X,Y,Z}(x, y, z)}{f_{Y,Z}(y, z)}$, For $f_{Y,Z}(y, z) > 0$
- Multiplication rule:
  $f_{X,Y,Z}(x, y, z) = f_{X|Y,Z}(x|y, z) f_{Y|Z}(y|z) f_Z(z)$, For $f_{Y,Z}(y, z) > 0$
- Independence: $f_{X,Y,Z}(x, y, z) = f_X(x) f_Y(y) f_Z(z)$ For all x, y, z

More than two random variables

- For two random variables X, Y arising out of the same experiment, we define their CDF as:

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

- How do I get $f_{X,Y}(x, y)$ back? $f_{X,Y}(x, y) = \dfrac{d^2 F_{X,Y}(x, y)}{dx\ dy}$
- Let X and Y be jointly uniform on the unit square. $F_{X,Y}(x, y) = xy$ for $0 \leq x, y \leq 1$
- What is $f_{X,Y}(x, y)$?. Differentiate! $\dfrac{d}{dx}\left(\dfrac{d}{dy}(xy)\right)$
- This equals 1 for all $0 \leq x, y \leq 1$!

More than two random variables

- For two random variables X, Y arising out of the same experiment, we define their CDF as:

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

- How do I get $f_{X,Y}(x,y)$ back? $f_{X,Y}(x,y) = \dfrac{d^2 F_{X,Y}(x,y)}{dx\ dy}$
- Let X and Y be jointly uniform on the unit square. $F_{X,Y}(x,y) = xy$ for $0 \leq x, y \leq 1$
- What is $f_{X,Y}(x,y)$?. Differentiate! $\dfrac{d}{dx}\left(\dfrac{d}{dy}(xy)\right)$
- This equals 1 for all $0 \leq x, y \leq 1$!

Function of random variables-I

- $X \sim Exp(1)$. Find the PDF of $Y = X^2$.
- How do we do this?
- Start with the CDF. Find $F_Y(y) = P(X^2 \leq y)$.

Function of random variables-I

- $X \sim Exp(1)$. Find the PDF of $Y = X^2$.
- How do we do this?
- Start with the CDF. Find $F_Y(y) = P(X^2 \leq y)$.
- $F_Y(y) = P(X^2 \leq y) = P(X \leq \sqrt{y}) = 1 - \exp(-\sqrt{y})$.

Function of random variables-I

- $X \sim Exp(1)$. Find the PDF of $Y = X^2$.
- How do we do this?
- Start with the CDF. Find $F_Y(y) = P(X^2 \leq y)$.
- $F_Y(y) = P(X^2 \leq y) = P(X \leq \sqrt{y}) = 1 - \exp(-\sqrt{y})$.
- So $f_Y(y) = \begin{cases} \frac{dF_Y(y)}{dy}x = \frac{\exp(-\sqrt{y})}{2\sqrt{y}} & y \geq 0 \\ 0 & \text{otherwise} \end{cases}$

Function of random variables-II

- $X \sim Uniform([0,1])$. Find the PDF of $Y = X^2$.
- How do we do this?
- Start with the CDF. Find $F_Y(y) = P(X^2 \leq y)$.

Function of random variables-II

- $X \sim Uniform([0,1])$. Find the PDF of $Y = X^2$.
- How do we do this?
- Start with the CDF. Find $F_Y(y) = P(X^2 \leq y)$.
- $F_Y(y) = P(X^2 \leq y) = P(X \leq \sqrt{y}) = \sqrt{y}$.

Function of random variables-II

- $X \sim \textit{Uniform}([0,1])$. Find the PDF of $Y = X^2$.
- How do we do this?
- Start with the CDF. Find $F_Y(y) = P(X^2 \leq y)$.
- $F_Y(y) = P(X^2 \leq y) = P(X \leq \sqrt{y}) = \sqrt{y}$.
- So $f_Y(y) = \begin{cases} \dfrac{dF_Y(y)}{dy} = \dfrac{1}{2\sqrt{y}} & y \in [0,1] \\ 0 & \text{otherwise} \end{cases}$