Overview

Definition of Independence

Two independent RVs

Two random variables X and Y, defined on the same probability space, are called **independent** if for any two sets $A, B \subset \mathbb{R}$

$$P(X \in A, Y \in B) = P(X \in A) \cdot P(Y \in B).$$

The independence of X and Y is denoted by $X \perp\!\!\!\perp Y$.

In words, X and Y are independent if knowing the value of one does not change the distribution of the other. Random variables that are not independent are said to be **dependent**.
Again, as it was for independent events, if independence is clear from the context then use

$$P(X \in A, Y \in B) = P(X \in A) \cdot P(Y \in B),$$

otherwise, check it to confirm/reject independence.

Equivalent Conditions for Independence

Theorem

The following conditions are equivalent:

1. **(In terms of definition)** $X \perp\!\!\!\perp Y$;
2. **(In terms of joint and marginal CDFs)** $F_{(X,Y)}(x,y) = F_X(x)F_Y(y)$, *for any* $(x,y) \in \mathbb{R}^2$;
3. **(In terms of joint and marginal PMFs)** *If X and Y are discrete random variables then* $P(X = x_i, Y = y_j) = P(X = x_i)P(Y = y_j)$, *for any* $(x_i, y_j) \in Range(X) \times Range(Y)$;
4. **(In terms of joint and marginal PDFs)** *If X and Y are continuous random variables then* $f_{(X,Y)}(x,y) = f_X(x)f_Y(y)$, *for any* $(x,y) \in \mathbb{R}^2$.

Examples

Example

Recall the piecewise function

$$f(x,y) = \begin{cases} xy \cdot e^{-\frac{x^2+y^2}{2}}, & \text{if } x \geq 0, y \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

we considered in the previous lecture. It is a joint PDF of some random vector (X, Y). We had established that

$$f_X(x) = \begin{cases} x \cdot e^{-\frac{x^2}{2}}, & \text{if } x \geq 0 \\ 0, & \text{otherwise} \end{cases} \text{ and } f_Y(y) = \begin{cases} y \cdot e^{-\frac{y^2}{2}}, & \text{if } y \geq 0 \\ 0, & \text{otherwise} \end{cases}.$$

Since $f_{(X,Y)}(x,y) = f_X(x)f_Y(y)$, then $X \perp\!\!\!\perp Y$.

Examples

Example (Uniform bivariate distribution over a disk with dependent variables)

Recall another example from the previous lecture: picking a random point from the disk D of radius 10, centered at the origin assumes $(X, Y) \sim U(D)$. In this case $f_{(X,Y)}(x, y) \neq f_X(x)f_Y(y)$ because

$$f_{(X,Y)}(x, y) = \begin{cases} \frac{1}{100\pi}, & \text{if } x^2 + y^2 \leq 100 \\ 0, & \text{otherwise.} \end{cases}$$

$$f_X(x) = \begin{cases} \frac{1}{50\pi}\sqrt{100 - x^2}, & \text{if } |x| \leq 10 \\ 0, & \text{otherwise} \end{cases}$$

$$f_Y(y) = \begin{cases} \frac{1}{50\pi}\sqrt{100 - y^2}, & \text{if } |y| \leq 10 \\ 0, & \text{otherwise} \end{cases}.$$

Examples

Replacing disk by a rectangle, uniform bivariate distribution will have independent coordinates. Unlike the previous example, the marginal distributions here are also uniform.

Uniform bivariate distribution over a rectangle

If $(X, Y) \sim U([a, b] \times [c, d])$ then

$$X \sim U[a, b], Y \sim U[c, d], \text{and } X \perp\!\!\!\perp Y.$$

To prove this we need the joint PDF first:

$$f_{(X,Y)}(x, y) = \begin{cases} \frac{1}{(b-a)(d-c)}, & \text{if } (x, y) \in [a, b] \times [c, d] \\ 0, & \text{otherwise.} \end{cases}$$

Examples. Cont'd

Now compute the marginal PDF of X:

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

$$\begin{cases} \int_c^d \frac{1}{(b-a)(d-c)}dy, & \text{if } x \in [a,b] \\ 0, & \text{otherwise} \end{cases} = \begin{cases} \frac{1}{b-a}, & \text{if } x \in [a,b] \\ 0, & \text{otherwise} \end{cases} \Rightarrow X \sim U[a,b].$$

Absolutely similarly, $Y \sim U[c,d]$, i.e.

$$f_Y(y) = \begin{cases} \frac{1}{d-c}, & \text{if } y \in [c,d] \\ 0, & \text{otherwise} \end{cases}.$$

$$f_{(X,Y)}(x,y) = f_X(x)f_Y(y) \Rightarrow X \perp\!\!\!\perp Y.$$

Examples

Example (Normal bivariate distribution with independent coordinates)

Consider one more example from the past lecture. Let $(X, Y) \sim N(\boldsymbol{\mu}, \Sigma)$ where $\boldsymbol{\mu} = \begin{bmatrix} -3 \\ 1 \end{bmatrix}$ and $\Sigma = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$. Lets make sure that $X \perp\!\!\!\perp Y$ and compute $P(X > -4, 0 \leq Y \leq 1)$ as $P(X > -4) \cdot P(0 \leq Y \leq 1)$.
We had computed the joint PDF

$$f_{(X,Y)}(x, y) = \frac{1}{2\sqrt{2}\pi} \cdot e^{-\frac{(x+3)^2}{4} - \frac{(y-1)^2}{2}}.$$

Let's calculate the marginal density functions:

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

Examples: Cont'd

Example

$$= \frac{1}{2\sqrt{\pi}} e^{-\frac{(x+3)^2}{4}} \underbrace{\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{(y-1)^2}{2}} dy}_{=1} = \frac{1}{2\sqrt{\pi}} e^{-\frac{(x+3)^2}{4}}.$$

This means that $X \sim N(-3, 2)$. Similarly, for the other marginal density function we will have $Y \sim N(1, 1)$ because of $f_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(y-1)^2}{2}}$. Indeed, $f_{(X,Y)}(x, y) = f_X(x) f_Y(y) \Rightarrow X \perp\!\!\!\perp Y$. Therefore

$$P(X > -4, 0 \leq Y \leq 1) = P(X > -4) \cdot P(0 \leq Y \leq 1) = ... =$$

$$0.761 \cdot 0.341 = 0.259.$$

Examples: Romeo and Juliet - 2

Example (Romeo and Juliet - 2)

Romeo and Juliet agree to meet at a certain location about 12:30 pm. If Romeo arrives at a time uniformly distributed between 12:15 and 12:45, and if Juliet **independently** arrives at a time uniformly distributed between 12:00 and 1pm, find the probability that the first to arrive waits no longer than 5 minutes. What is the probability that Romeo arrives first?

Solution: Let X and Y be arrival times in minutes after 12:00 of Romeo and Juliet correspondingly. Then

$$f_X(x) = \begin{cases} \frac{1}{30}, & \text{if } x \in [15, 45] \\ 0, & \text{otherwise} \end{cases} \text{ and } f_Y(y) = \begin{cases} \frac{1}{60}, & \text{if } y \in [0, 60] \\ 0, & \text{otherwise} \end{cases}.$$

Solution ...

Since $X \perp\!\!\!\perp Y$,

$$f_{(X,Y)}(x,y) = \begin{cases} \frac{1}{1800}, & \text{if } (x,y) \in [15,45] \times [0,60] \\ 0, & \text{otherwise} \end{cases}.$$

$$P(|Y-X|<5) = \iint_{|Y-X|<5} f_{(X,Y)}(x,y)dxdy =$$

$$\int_{15}^{45} dx \int_{x-5}^{x+5} 1/1800 dy = 1/6.$$

$$P(X<Y) = \iint_{Y>X} f_{(X,Y)}(x,y)dxdy =$$

$$\int_{15}^{45} dx \int_{x}^{60} 1/1800 dy = 1/2.$$

Example: Buffon's Needle

Example (Buffon's needle problem)

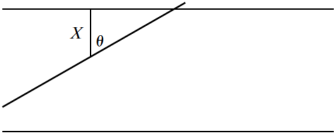
A table is ruled with equidistant parallel lines a distance D apart. A needle of length L, where $L \leq D$, is randomly thrown on the table. What is the probability that the needle will intersect one of the lines (the other possibility being that the needle will be completely contained in the strip between two lines)?

To solve this problem let's consider two random variables:

1. The distance X from the middle point of the needle to the nearest parallel line. It is natural to assume that $X \sim U[0, \frac{D}{2}]$.

2. The angle θ between the needle and the vertical direction (see the figure below). Here we will assume $\theta \sim U[0, \pi/2]$.

Solution ...



The needle will intersect a horizontal line if the hypotenuse of the right triangle above is less than $L/2$, i.e. $X < \frac{L}{2} cos\theta$.

$$P(X < \frac{L}{2} cos\theta) = \iint_{x < \frac{L}{2} cos y} f_{(X,\theta)}(x, y) dxdy =$$

$$\iint_{x < \frac{L}{2} cos y} f_X(x) f_\theta(y) dxdy = \frac{2}{D}\frac{2}{\pi}\int_0^{\pi/2} dy \int_0^{\frac{L}{2} cos y} dx =$$

$$\frac{4}{\pi D}\int_0^{\pi/2} \frac{L}{2} cos y dy = \frac{2L}{\pi D}.$$