# Joint Probability Distributions & Independence

Section 6.3

STAT/MATH 395 Spring 2020

Vincent Roulet

Lecture 7, April 13th, 2020

Ask questions via chat on Zoom

Answer quiz via PollEverywhere (username: vincentroulet)

# Overview

**Optional exercises in homeworks**

- Additional material for you to master the course
- Adds up to the grade of the homework up to the total score
- Taken into account for any recommendation letter

**Previous lectures**

- Joint distributions, discrete and continuous cases

**This lecture**

- Joint distributions and independence,
- Discrete independent random variables
- Continuous independent random variables
- Functions of independent random variables
- Minimum, maximum of independent random variables

# Answer Previous Quiz

## Exercise

*I am shooting an arrow at a target on a wall* $W=\{(x,y): -1\leq x\leq 1, 0\leq y\leq 1\}$*.* *Wind from the left and gravity affect my shot such that the position of the arrow has a p.d.f. proportional to* $\frac{e^x}{\sqrt{y+1}}$

*What is the probability*
*that I touch the target* $T=\{(x,y): -0.1\leq x\leq 0.1, 0.4\leq y\leq 0.6\}$*?*

[1] When limits of an integral are not specified this means that the integral goes form $-\infty$ to $+\infty$

# Answer Previous Quiz

## Exercise

*I am shooting an arrow at a target on a wall* $W = \{(x,y) : -1 \le x \le 1, 0 \le y \le 1\}$. *Wind from the left and gravity affect my shot such that the position of the arrow has a p.d.f. proportional to* $\frac{e^x}{\sqrt{y+1}}$

*What is the probability*
*that I touch the target* $T = \{(x,y) : -0.1 \le x \le 0.1, 0.4 \le y \le 0.6\}$*?*

**Solution**

1. $f(x,y) = \frac{1}{\lambda} \frac{e^x}{\sqrt{y+1}} \mathbf{1}_W(x,y)$ with $\lambda \ge 0$, we have $\int_{-\infty}^{+\infty} f(x,y)dxdy = 1$ and so[1]

$$\lambda = \int\int \frac{e^x}{\sqrt{y+1}} \mathbf{1}_W(x,y)dxdy = \int_0^1 \left( \int_{-1}^1 \frac{e^x}{\sqrt{y+1}} dx \right) dy = 2(\sqrt{2}-1)(e - e^{-1})$$

2.

$$\mathbb{P}((X,Y) \in T) = \frac{1}{\lambda} \int_{0.4}^{0.6} \int_{-0.1}^{0.1} \frac{e^x}{\sqrt{y+1}} dxdy = \frac{2(\sqrt{1.6}-\sqrt{1.4})(e^{0.1}-e^{-0.1})}{\lambda} \approx 0.017$$

[1] When limits of an integral are not specified this means that the integral goes form $-\infty$ to $+\infty$

# Independent Random Variables

### Definition (Independent random variables)

Random variables $X_1, \dots, X_n$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ are **independent** if for any[2] subsets $B_1, \dots, B_n \subset \mathbb{R}$,

$$\mathbb{P}(X_1 \in B_1, \dots, X_n \in B_n) = \mathbb{P}(X_1 \in B_1) \dots \mathbb{P}(X_n \in B_n)$$

or equivalently if their joint c.d.f. $F$ factorizes into the marginal c.d.f. as

$$F(t_1, \dots, t_n) = F_{X_1}(t_1) \dots F_{X_n}(t_n)$$

[2]Again a formal definition requires these subsets to be Borel subsets of $\mathbb{R}^n$

# Independent Random Variables

## Definition (Independent random variables)

Random variables $X_1, \ldots, X_n$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ are **independent** if for any[2] subsets $B_1, \ldots, B_n \subset \mathbb{R}$,

$$\mathbb{P}(X_1 \in B_1, \ldots, X_n \in B_n) = \mathbb{P}(X_1 \in B_1) \ldots \mathbb{P}(X_n \in B_n)$$

or equivalently if their joint c.d.f. $F$ factorizes into the marginal c.d.f. as

$$F(t_1, \ldots, t_n) = F_{X_1}(t_1) \ldots F_{X_n}(t_n)$$

**Proof** If they are independent then the joint c.d.f. factorizes by definition
If the c.d.f. factorizes into the marginals, the idea is that all the Borel subsets we want to measure can be generated by intervals of the form $(-\infty, t]$ for $t \in \mathbb{R}$ by taking intersections or unions of these intervals.

[2] Again a formal definition requires these subsets to be Borel subsets of $\mathbb{R}^n$

# Independent Random Variables

## Definition (Independent random variables)

Random variables $X_1, \ldots, X_n$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ are **independent** if for any[2] subsets $B_1, \ldots, B_n \subset \mathbb{R}$,

$$\mathbb{P}(X_1 \in B_1, \ldots, X_n \in B_n) = \mathbb{P}(X_1 \in B_1) \ldots \mathbb{P}(X_n \in B_n)$$

or equivalently if their joint c.d.f. $F$ factorizes into the marginal c.d.f. as

$$F(t_1, \ldots, t_n) = F_{X_1}(t_1) \ldots F_{X_n}(t_n)$$

How can we understand independence
by simply looking at the joint distribution?

What are the consequences in terms of p.m.f., p.d.f., c.d.f.?

[2] Again a formal definition requires these subsets to be Borel subsets of $\mathbb{R}^n$

# Independent Discrete Random Variables

**Lemma**

*Let $X_1, \dots X_n$ be n discrete random variables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Then $X_1, \dots X_n$ are independent if and only if their joint p.m.f. p factorizes into the marginals $p_{X_i}$,*

$$p(k_1, \dots, k_n) = p_{X_1}(k_1) \dots p_{X_n}(k_n)$$

## Independent Discrete Random Variables

### Lemma

*Let $X_1, \dots X_n$ be $n$ discrete random variables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Then $X_1, \dots X_n$ are independent if and only if their joint p.m.f. $p$ factorizes into the marginals $p_{X_i}$,*

$$p(k_1, \dots, k_n) = p_{X_1}(k_1) \dots p_{X_n}(k_n)$$

**Proof** If $X_1, \dots, X_n$ are independent the result comes from the definition.
If the joint p.m.f. factorizes into the marginal distributions, then

$$\begin{aligned}
\mathbb{P}(X_1 \in B_1, \dots, X_n \in B_n) &= \sum_{k_1 \in B_1, \dots, k_n \in B_n} p(k_1, \dots, k_n) \\
&= \sum_{k_1 \in B_1, \dots, k_n \in B_n} p_{X_1}(k_1) \dots p_{X_n}(k_n) \\
&= \left( \sum_{k_1 \in B_1} p_{X_1}(k_1) \right) \dots \left( \sum_{k_n \in B_n} p_{X_n}(k_n) \right) = \prod_{i=1}^{n} \mathbb{P}(X_i \in B_i)
\end{aligned}$$

# Independent Discrete Random Variables

## Example

1. Roll two dices with **4 faces**, denote
   (i) $S$ the sum of the two dices
   (ii) $Y$ the indicator variable that you get a pair
2. Are $S, Y$ independent?

| | | Y | |
|---|---|---|---|
| | | 0 | 1 |
| | 2 | 0 | 1/16 |
| | 3 | 1/8 | 0 |
| | 4 | 1/8 | 1/16 |
| S | 5 | 1/4 | 0 |
| | 6 | 1/8 | 1/16 |
| | 7 | 1/8 | 0 |
| | 8 | 0 | 1/16 |

# Independent Discrete Random Variables

## Example

1. Roll two dices with **4 faces**, denote
   (i) $S$ the sum of the two dices
   (ii) $Y$ the indicator variable that you get a pair
2. Are $S, Y$ independent?

| | | Y | |
|---|---|---|---|
| | | 0 | 1 |
| | 2 | 0 | 1/16 |
| | 3 | 1/8 | 0 |
| | 4 | 1/8 | 1/16 |
| S | 5 | 1/4 | 0 |
| | 6 | 1/8 | 1/16 |
| | 7 | 1/8 | 0 |
| | 8 | 0 | 1/16 |

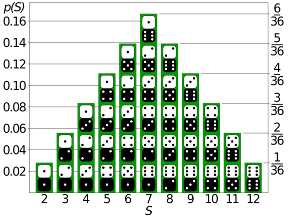
**Solution** Check for example $\mathbb{P}(S = 2, Y = 0) = 0 \neq \mathbb{P}(S = 2)\,\mathbb{P}(Y = 0) > 0$
Note: one counterexample suffices to show that $S, Y$ are dependent,
but to prove independence one would need to show the equality for all values of $S, Y$

# Independent Discrete Random Variables

## Example

Roll repeatedly a pair of dice.
Denote $N$ the number of rolls until the sum of the dice is 2 or a 6

1. What is the distribution of $N$?
2. Denote $X$ the sum you finally get (2 or 6), are $X$ and $N$ independent?

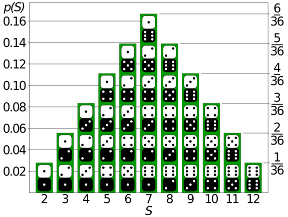


# Independent Discrete Random Variables

## Example

Roll repeatedly a pair of dice.
Denote $N$ the number of rolls until the sum of the dice is 2 or a 6

1. What is the distribution of $N$?
2. Denote $X$ the sum you finally get (2 or 6), are $X$ and $N$ independent?




**Solution**

1. Let $Y_i$ be the sum of the two dice at the i$^{\text{th}}$ roll.
   We have $\mathbb{P}(Y_i \in \{2,6\}) = 1/36 + 5/36 = 1/6$ and so $N \sim \text{Geom}(1/6)$
2. $\mathbb{P}(N = n, X = 6) = \mathbb{P}(Y_1 \notin \{2,6\}, \ldots, Y_{n-1} \notin \{2,6\}, Y_n = 6) = \left(\frac{5}{6}\right)^{n-1} \frac{5}{36}$
   Therefore $\mathbb{P}(X = 6) = \sum_{n=1}^{+\infty} \left(\frac{5}{6}\right)^{n-1} \frac{5}{36} = \frac{5/36}{1-5/6} = 5/6$
   So $\mathbb{P}(N = n, X = 6) = \left(\frac{5}{6}\right)^{n-1} \frac{1}{6}\frac{5}{6} = \mathbb{P}(N = n)\,\mathbb{P}(X = 6)$
   Same argument shows $\mathbb{P}(N = n, X = 2) = \mathbb{P}(N = n)\,\mathbb{P}(X = 2)$
   $\rightarrow$ $N$ and $X$ are independent

# Independent Continuous Random Variables

## Lemma

*Let $X_1, \dots, X_n$ be $n$ r.v. on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Assume that for $j \in \{1, \dots, n\}$, the rv. $X_j$ has p.d.f. $f_{X_j}$.*

1. *If $X_1, \dots, X_n$ have a joint p.d.f. that factorizes in the marginal p.d.f. as*

   $$f(x_1, \dots, x_n) = f_{X_1}(x_1) \dots f_{X_n}(x_n)$$

   *then $X_1, \dots, X_n$ are independent.*
2. *Conversely if $X_1, \dots, X_n$ are independent then they are jointly continuous with joint p.d.f.*

   $$f(x_1, \dots, x_n) = f_{X_1}(x_1) \dots f_{X_n}(x_n)$$

**Note:**

1. Checking if $X_1, \dots X_n$ are independent can be done by looking at the joint p.d.f.
2. Conversely if they are independent, we know that they have a joint p.d.f. (remember that it was not always the case a priori)

# Independent Continuous Random Variables

## Lemma

*Let $X_1, \dots, X_n$ be $n$ r.v. on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Assume that for $j \in \{1, \dots, n\}$, the rv. $X_j$ has p.d.f. $f_{X_j}$.*

1. *If $X_1, \dots, X_n$ have a joint p.d.f. that factorizes in the marginal p.d.f. as*

$$f(x_1, \dots, x_n) = f_{X_1}(x_1) \dots f_{X_n}(x_n)$$

   *then $X_1, \dots, X_n$ are independent.*

2. *Conversely if $X_1, \dots, X_n$ are independent then they are jointly continuous with joint p.d.f.*

$$f(x_1, \dots, x_n) = f_{X_1}(x_1) \dots f_{X_n}(x_n)$$

**Proof** For $n = 2$ with two r.v. $(X, Y)$, denote $A, B \subset \mathbb{R}$,

$$\mathbb{P}(X \in A, Y \in B) = \int_A \int_B f_{X,Y}(x,y)dxdy = \int_A \int_B f_X(x) f_Y(y) dxdy$$

$$= \int_A f_X(x)dx \int_B f_Y(y)dy = \mathbb{P}(X \in A)\,\mathbb{P}(X \in B)$$

Conversely, if $X, Y$ are independent

$$\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A)\,\mathbb{P}(X \in B) = \int_A \int_B f_X(x) f_Y(y) dxdy$$

Example (Shooting an arrow)

Consider $X$, $Y$ with p.d.f. $f(x,y) = \frac{1}{\lambda}\frac{e^x}{\sqrt{y+1}}\mathbf{1}_W(x,y)$ for $\lambda{=}2(\sqrt{2}-1)(e-e^{-1})$ where $W{=}\{(x,y) : -1{\leq}x{\leq}1, 0{\leq}y{\leq}1\}$.

1. Are $X$, $Y$ independent?
2. What consequences it had when computing the probability to get the target $T{=}\{(x,y){:}{-0.1}{\leq}x{\leq}0.1, 0.4{\leq}y{\leq}0.6\}$?

## Example (Shooting an arrow)

Consider $X, Y$ with p.d.f. $f(x,y) = \frac{1}{\lambda}\frac{e^x}{\sqrt{y+1}}\mathbf{1}_W(x,y)$ for $\lambda = 2(\sqrt{2}-1)(e-e^{-1})$ where $W = \{(x,y) : -1 \le x \le 1, 0 \le y \le 1\}$.

1. Are $X, Y$ independent?
2. What consequences it had when computing the probability to get the target $T = \{(x,y): -0.1 \le x \le 0.1, 0.4 \le y \le 0.6\}$?

**Solution**

1. Note that $\mathbf{1}_W(x,y) = \mathbf{1}_{[-1,1]}(x)\,\mathbf{1}_{[0,1]}(y)$,
   then one has $f_X(x) = \frac{1}{e-e^{-1}}e^x\,\mathbf{1}_{[-1,1]}(x)$, $f_Y(y) = \frac{1}{2(\sqrt{2}-1)\sqrt{y+1}}\,\mathbf{1}_{[0,1]}(y)$
   So $X, Y$ are independent
2. $\mathbb{P}((X,Y) \in T) = \mathbb{P}(X \in [-0.1, 0.1])\,\mathbb{P}(Y \in [0.4, 0.6])$ where $\mathbb{P}(X \in [-0.1, 0.1])$ , $\mathbb{P}(Y \in [0.4, 0.6])$ can be computed from $f_X$, $f_Y$ respectively.

# Quiz for next lecture

Example

Let $X, Y$ be two jointly continuous r.v., are $X, Y$ independent if

1. their joint p.d.f. is $f_{X,Y}(x,y) = \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}}$?
2. their joint p.d.f. is $f_{X,Y}(x,y) = \frac{1}{2\pi\sqrt{1-\rho^2}} e^{-\frac{x^2+y^2-2\rho xy}{2(1-\rho^2)}}$ for $-1 < \rho < 1$?