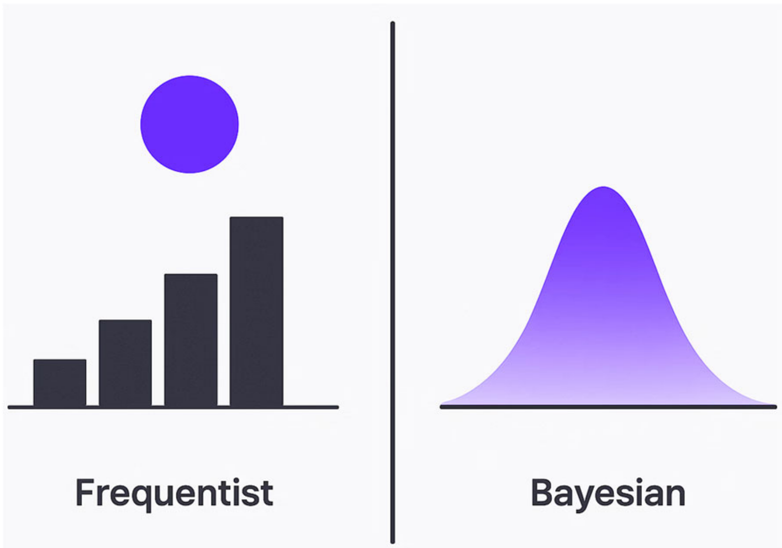
1.2 El rol de la probabilidad y estadística

Probability measure on random variables

REMEMBER!!! – What is probability???
What are the Axioms of Probability??

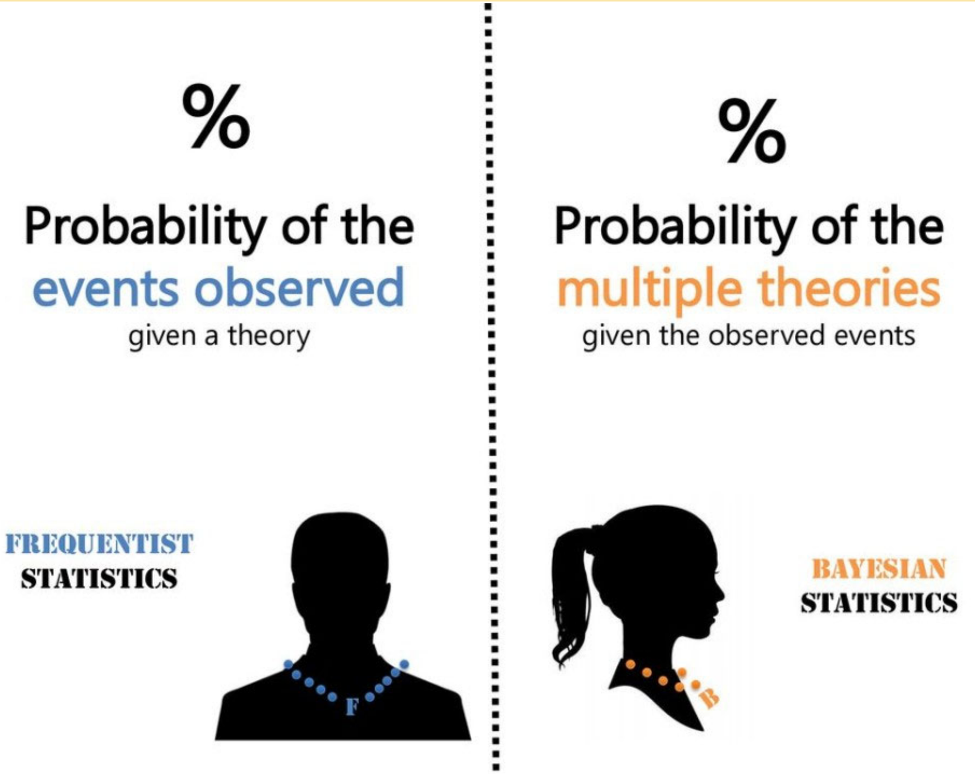
Student´s answers



1.2 El rol de la probabilidad y estadística

Probability measure on random variables

REMEMBER!!! – What is probability???



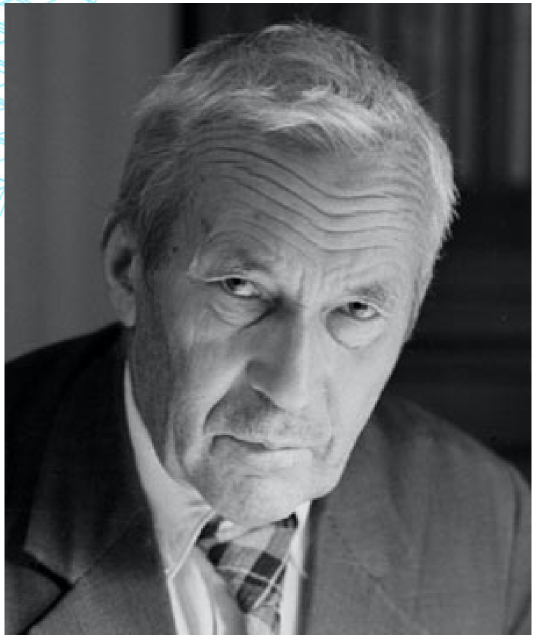
TECNOLÓGICO NACIONAL DE MÉXICO

1.2 El rol de la probabilidad y estadística

EDUCACIÓN
SECRETARÍA DE EDUCACIÓN PÚBLICA

Probability measure on random variables

REMEMBER!!! – What is probability???



Andrey Kolmogorov (1903-1987) was one of the most influential mathematicians of the 20th century, especially renowned for his fundamental contributions to probability theory. He is considered the father of modern probability.

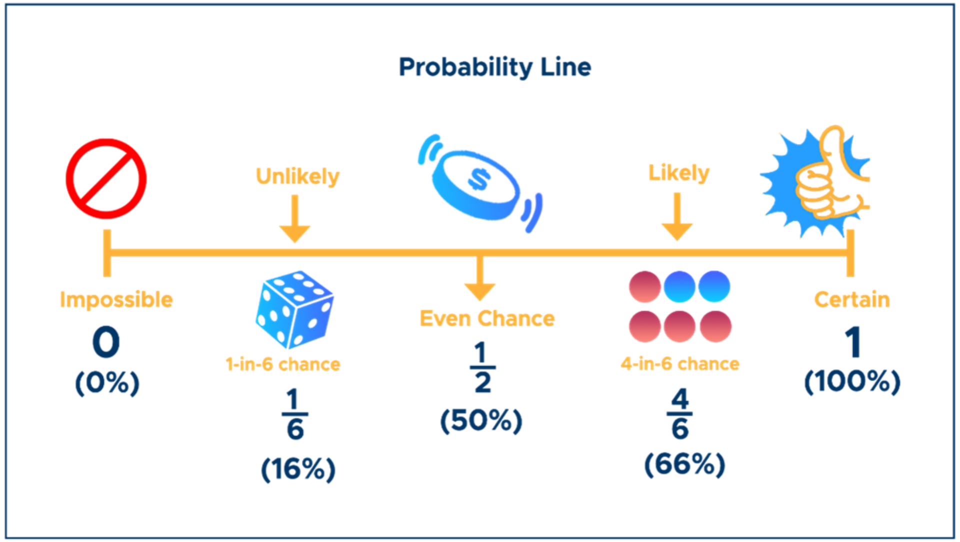
In 1933, he published his work "Foundations of the Theory of Probability," in which he formulated Kolmogorov's axioms, which establish the rigorous mathematical foundations of probability as a branch of set theory.

1.2 El rol de la probabilidad y estadística

Probability measure on random variables

REMEMBER!!! – What is probability???

Probability is a measure of the size of a set.



TECNOLÓGICO NACIONAL DE MÉXICO

1.2 El rol de la probabilidad y estadística

EDUCACIÓN
SECRETARÍA DE EDUCACIÓN PÚBLICA

Probability measure on random variables

REMEMBER!!! – What is probability???

Suppose that you have a fair die. It has 6 faces: {⚀, ⚁, ⚂, ⚃, ⚄, ⚅}

Students

What is the probability that you get a number that is "less than 5" and is "an even number"?

TECNOLÓGICO NACIONAL DE MÉXICO

1.2 El rol de la probabilidad y estadística

EDUCACIÓN
SECRETARÍA DE EDUCACIÓN PÚBLICA

Probability measure on random variables

REMEMBER!!! – What is probability???

Suppose that you have a fair die. It has 6 faces: {⚀, ⚁, ⚂, ⚃, ⚄, ⚅}

Students
What is the probability that you get a number that is "less than 5" and is "an even number"?

ANSWER: 2/6

1.2 El rol de la probabilidad y estadística

TECNOLÓGICO NACIONAL DE MÉXICO

EDUCACIÓN
SECRETARÍA DE EDUCACIÓN PÚBLICA

Probability measure on random variables

REMEMBER!!! – What is probability???

Sample Space

A sample space is the set containing all possible outcomes.

Ω = {⚀, ⚁, ⚂, ⚃, ⚄, ⚅}

Events

E_1 = "less than 5" = {⚀, ⚁, ⚂, ⚃}

E_2 = "an even number" = {⚁, ⚃, ⚅}

Then we take the intersection between these two events to conclude the event

E = {⚁, ⚃}

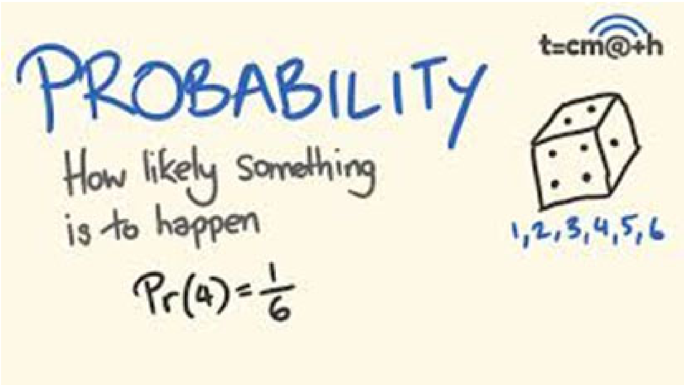
The numerical value "2" is the size of this event E.

1.2 El rol de la probabilidad y estadística

Probability measure on random variables

REMEMBER!!! – What is probability???

So, when we say that "the probability is $\frac{2}{6}$," we are saying that the size of the event E relative to the sample space Ω is the ratio $\frac{2}{6}$. This process involves **measuring** the size of E and Ω. In this particular example, the measure we use is a "counter" that counts the number of elements.



TECNOLÓGICO NACIONAL DE MÉXICO

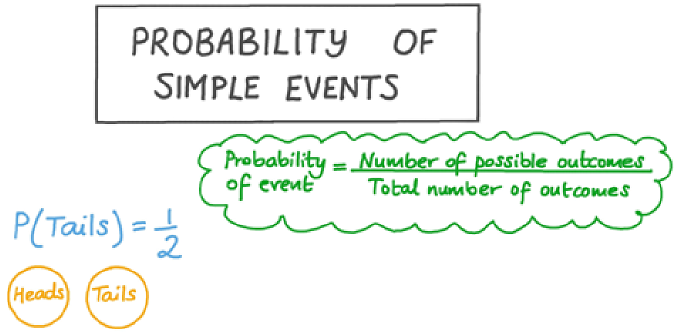
1.2 El rol de la probabilidad y estadística

EDUCACIÓN
SECRETARÍA DE EDUCACIÓN PÚBLICA

Probability measure on random variables

REMEMBER!!! – What is probability???

This example shows us all the necessary components of probability: (i) There is a **sample space**, which is the set that contains all the possible outcomes. (ii) There is an **event**, which is a subset inside the sample space. (iii) Two events E_1 and E_2 can be **combined** to construct another event E that is still a subset inside the sample space. (iv) Probability is a number assigned by certain **rules** such that it describes the **relative size** of the event E compared with the sample space Ω. So, when we say that **probability is a measure of the size of a set**, we create a mapping that takes in a set and outputs the size of that set.

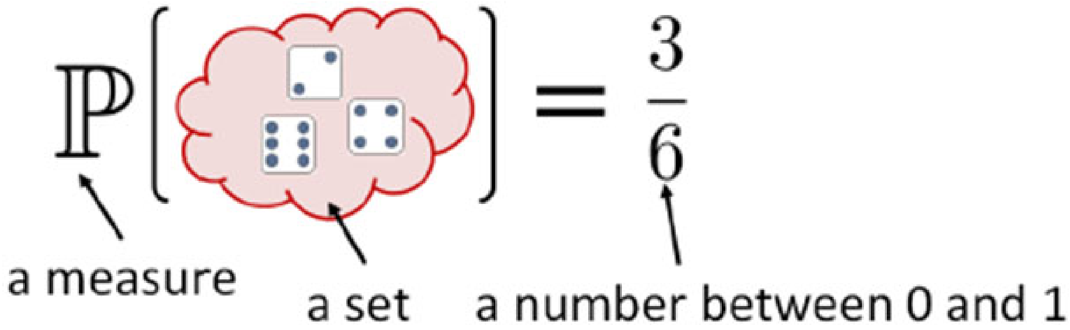


1.2 El rol de la probabilidad y estadística

Probability measure on random variables

REMEMBER!!! – What is probability???

Probability is a measure of the size of a **set**. Whenever we talk about probability, it has to be the probability of a set.



TECNOLÓGICO NACIONAL DE MÉXICO

1.2 El rol de la probabilidad y estadística

EDUCACIÓN
SECRETARÍA DE EDUCACIÓN PÚBLICA

Probability measure on random variables

REMEMBER!!! – What is probability???

Axioms of Probability

Definition 2.21. *A* **probability law** *is a function* $\mathbb{P} : \mathcal{F} \to [0, 1]$ *that maps an event* A *to a real number in* $[0, 1]$. *The function must satisfy the* **axioms of probability**:

I. **Non-negativity**: $\mathbb{P}[A] \geq 0$, *for any* $A \subseteq \Omega$.

II. **Normalization**: $\mathbb{P}[\Omega] = 1$.

III. **Additivity**: *For any disjoint sets* $\{A_1, A_2, \ldots\}$, *it must be true that*

$$\mathbb{P}\left[\bigcup_{i=1}^{\infty} A_i\right] = \sum_{i=1}^{\infty} \mathbb{P}[A_i]. \tag{2.23}$$

1.2 El rol de la probabilidad y estadística

Probability measure on random variables

Homework

Make a presentation (5-10 slides) to explain next concepts:

- Conditional probability.
- Bayes’ Theorem.
- Law of Total Probability.

1.2 El rol de la probabilidad y estadística

Probability measure on random variables

***Optional**

By now, we hope that you understand Key Concept 1: **A random variable is a mapping from a statement to a number**. However, we are now facing another difficulty. We knew how to measure the size of an event using the probability law $\mathbb{P}$ because $\mathbb{P}(\cdot)$ takes an event $E \in \mathcal{F}$ and sends it to a number between $[0, 1]$. After the translation X, we cannot send the output $X(\xi)$ to $\mathbb{P}(\cdot)$ because $\mathbb{P}(\cdot)$ "eats" a set $E \in \mathcal{F}$ and not a number $X(\xi) \in \mathbb{R}$. Therefore, when we write $\mathbb{P}[X = 1]$, how do we measure the size of the event $X = 1$?

Probability measure on random variables

*Optional

This question appears difficult but is actually quite easy to answer. Since the probability law $\mathbb{P}(\cdot)$ is always applied to an **event**, we need to define an event for the random variable X. If we write the sets clearly, we note that "$X = a$" is equivalent to the set

$$E = \left\{ \xi \in \Omega \,\middle|\, X(\xi) = a \right\}.$$

This is the set that contains all possible ξ's such that $X(\xi) = a$. Therefore, when we say "find the probability of $X = a$," we are effectively asking the size of the set $E = \{\xi \in \Omega \,|\, X(\xi) = a\}$.

How then do we measure the size of E? Since E is a subset in the sample space, E is measurable by $\mathbb{P}$. All we need to do is to determine what E is for a given a. This, in turn, requires us to find the **pre-image** $X^{-1}(a)$, which is defined as

$$X^{-1}(a) \stackrel{\text{def}}{=} \left\{ \xi \in \Omega \,\middle|\, X(\xi) = a \right\}.$$

1.2 El rol de la probabilidad y estadística

Probability measure on random variables

*Optional

Wait a minute, is this set just equal to E? Yes, the event E we are seeking is exactly the pre-image $X^{-1}(a)$. As such, the probability measure of E is

$$\mathbb{P}[X = a] = \mathbb{P}[X^{-1}(a)].$$

Figure 3.4 illustrates a situation where two outcomes ξ_1 and ξ_2 are mapped to the same value a on the real line. The corresponding event is the set $X^{-1}(a) = \{\xi_1, \xi_2\}$.

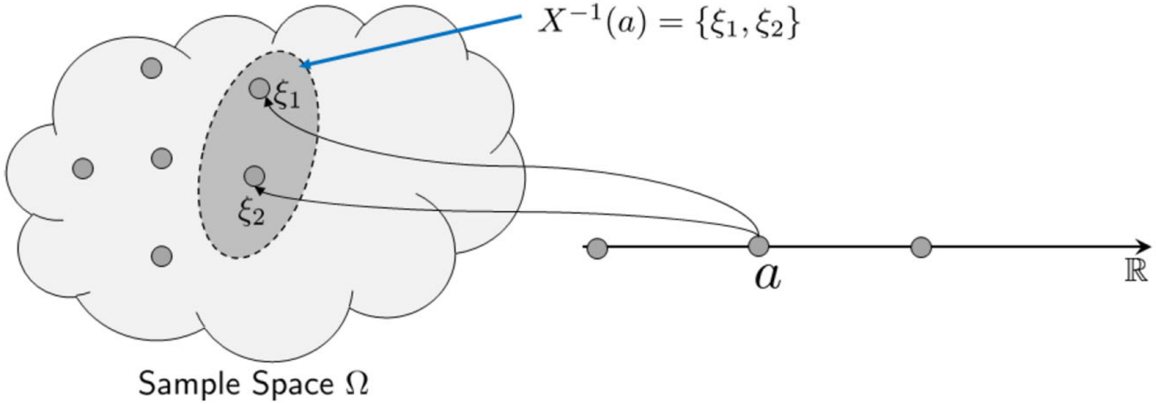


Figure 3.4: When computing the probability of $\mathbb{P}[\{\xi \in \Omega \mid X(\xi) = a\}]$, we effectively take the inverse mapping $X^{-1}(a)$ and compute the probability of the event $\mathbb{P}[\{\xi \in X^{-1}(a)\}] = \mathbb{P}[\{\xi_1, \xi_2\}]$.

TECNOLÓGICO NACIONAL DE MÉXICO

1.2 El rol de la probabilidad y estadística

EDUCACIÓN
SECRETARÍA DE EDUCACIÓN PÚBLICA

Probability measure on random variables

Examples

Example 3.3. Suppose we throw a die. The sample space is

$$\Omega = \{\epsdice{1}, \epsdice{2}, \epsdice{3}, \epsdice{4}, \epsdice{5}, \epsdice{6}\}.$$

There is a natural mapping X that maps $X(\epsdice{1}) = 1$, $X(\epsdice{2}) = 2$ and so on. Thus,

$$\begin{aligned} \mathbb{P}[X \leq 3] &\overset{(a)}{=} \mathbb{P}[X = 1] + \mathbb{P}[X = 2] + \mathbb{P}[X = 3] \\ &\overset{(b)}{=} \mathbb{P}[X^{-1}(1)] + \mathbb{P}[X^{-1}(2)] + \mathbb{P}[X^{-1}(3)] \\ &\overset{(c)}{=} \mathbb{P}[\{\epsdice{1}\}] + \mathbb{P}[\{\epsdice{2}\}] + \mathbb{P}[\{\epsdice{3}\}] = \frac{3}{6}. \end{aligned}$$

In this derivation, step (a) is based on Axiom III, where the three events are disjoint.
Step (b) is the pre-image due to the random variable X.
Step (c) is the list of actual events in the event space.
Every step is justified by the concepts and theorems we have learned so far.

Probability measure on random variables

Exercise

Consider the experiment of throwing a fair die. Let X be the r.v. which assigns 1 if the number that appears is even and 0 if the number that appears is odd.

(a) What is the range of X?

(b) Find $P(X = 1)$ and $P(X = 0)$.

The sample space S on which X is defined consists of 6 points which are equally likely:

$$S = \{1, 2, 3, 4, 5, 6\}$$

1.2 El rol de la probabilidad y estadística

Probability measure on random variables

SOLUTION

(a) The range of X is $R_X = \{0, 1\}$.

(b) $(X = 1) = \{2, 4, 6\}$. Thus, $P(X = 1) = \frac{3}{6} = \frac{1}{2}$. Similarly, $(X = 0) = \{1, 3, 5\}$, and $P(X = 0) = \frac{1}{2}$.

TECNOLÓGICO NACIONAL DE MÉXICO

1.2 El rol de la probabilidad y estadística

EDUCACIÓN

SECRETARÍA DE EDUCACIÓN PÚBLICA

Probability measure on random variables

Exercise

Suppose throw a die twice. The sample space is then

$$\Omega = \{(⚀, ⚀), (⚀, ⚁), \ldots, (⚅, ⚅)\}.$$

Find

- The elements $\xi_i \in \Omega$.
- Assign the random variable $X = sum\ of\ two\ numbers$, and find the set $X(\xi) = 7$.
- The $P[X = 7]$

1.2 El rol de la probabilidad y estadística

TECNOLÓGICO NACIONAL DE MÉXICO

EDUCACIÓN
SECRETARÍA DE EDUCACIÓN PÚBLICA

Probability measure on random variables

SOLUTION

These elements can be translated to 36 outcomes:

$$\xi_1 = (\text{⚀}, \text{⚀}), \xi_2 = (\text{⚀}, \text{⚁}), \ldots, \xi_{36} = (\text{⚅}, \text{⚅}).$$

Let

$$X = \text{sum of two numbers.}$$

Then, if we want to find the probability of getting $X = 7$, we can trace back and ask: Among the 36 outcomes, which of those ξ_i's will give us $X(\xi) = 7$? Or, what is the set $X^{-1}(7)$? To this end, we can write

$$\begin{aligned}
\mathbb{P}[X = 7] &= \mathbb{P}[\{(\text{⚀}, \text{⚅}), (\text{⚁}, \text{⚄}), (\text{⚂}, \text{⚃}), (\text{⚃}, \text{⚂}), (\text{⚄}, \text{⚁}), (\text{⚅}, \text{⚀})\}] \\
&= \mathbb{P}[(\text{⚀}, \text{⚅})] + \mathbb{P}[(\text{⚁}, \text{⚄})] + \mathbb{P}[(\text{⚂}, \text{⚃})] \\
&\quad + \mathbb{P}[(\text{⚃}, \text{⚂})] + \mathbb{P}[(\text{⚄}, \text{⚁})] + \mathbb{P}[(\text{⚅}, \text{⚀})] \\
&= \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36} = \frac{1}{6}.
\end{aligned}$$

Again, in this example, you can see that all the steps are fully justified by the concepts we have learned so far.

TECNOLÓGICO NACIONAL DE MÉXICO

1.2 El rol de la probabilidad y estadística

EDUCACIÓN
SECRETARÍA DE EDUCACIÓN PÚBLICA

Probability measure on random variables

HW Exercise

Consider the experiment of tossing a coin three times. Let *X* be the r.v. giving the number of heads obtained. We assume that the tosses are independent and the probability of a head is *p*.

(a) What is the range of *X*?

(b) Find the probabilities P(*X* = 0), P(*X* = 1), P(*X* = 2), and P(*X* = 3).

(c) Make a Python solution proposal to compute probabilities for an unfair coin.

Key concept

Definition 2.23. *Two events A and B are statistically* **independent** *if*

$$\mathbb{P}[A \cap B] = \mathbb{P}[A]\mathbb{P}[B].$$

Probability measure on random variables

SOLUTION

The sample space S on which X is defined consists of eight sample points
S = {HHH, HHT, …, TTT}

(*a*) The range of X is $R_X = \{0, 1, 2, 3\}$.

(*b*) If $P(H) = p$, then $P(T) = 1 - p$. Since the tosses are independent, we have

$$P(X = 0) = P[\{TTT\}] = (1 - p)^3$$

$$P(X = 1) = P[\{HTT\}] + P[\{THT\}] + P[\{TTH\}] = 3(1 - p)^2 p$$

$$P(X = 2) = P[\{HHT\}] + P[\{HTH\}] + P[\{THH\}] = 3(1 - p)p^2$$

$$P(X = 3) = P[\{HHH\}] = p^3$$

1.2 El rol de la probabilidad y estadística

Probability Mass Function (PMF)

Random variables are mappings that translate events to numbers. After the translation, we have a set of numbers denoting the **states** of the random variables. Each state has a different probability of occurring. The probabilities are summarized by a function known as the probability mass function (PMF).

Definition 3.2. *The* **probability mass function (PMF)** *of a random variable X is a function which specifies the probability of obtaining a number $X(\xi) = x$. We denote a PMF as*

$$p_X(x) = \mathbb{P}[X = x]. \tag{3.1}$$

The set of all possible states of X is denoted as $X(\Omega)$.

The probability mass function is a histogram summarizing the probability of each of the states X takes. Since it is a histogram, a PMF can be easily drawn as a bar chart.

TECNOLÓGICO NACIONAL DE MÉXICO

1.2 El rol de la probabilidad y estadística

EDUCACIÓN

SECRETARÍA DE EDUCACIÓN PÚBLICA

Probability Mass Function (PMF)

Example

Example 3.5. Flip a coin twice. The sample space is $\Omega = \{\text{HH}, \text{HT}, \text{TH}, \text{TT}\}$. We can assign a random variable X = number of heads. Therefore,

$$X(\text{“HH”}) = 2, X(\text{“TH”}) = 1, X(\text{“HT”}) = 1, X(\text{“TT”}) = 0.$$

So the random variable X takes three states: 0, 1, 2. The PMF is therefore

$$p_X(0) = \mathbb{P}[X = 0] = \mathbb{P}[\{\text{“TT”}\}] = \frac{1}{4},$$
$$p_X(1) = \mathbb{P}[X = 1] = \mathbb{P}[\{\text{“TH”}, \text{“HT”}\}] = \frac{1}{2},$$
$$p_X(2) = \mathbb{P}[X = 2] = \mathbb{P}[\{\text{“HH”}\}] = \frac{1}{4}.$$

1.2 El rol de la probabilidad y estadística

TECNOLÓGICO NACIONAL DE MÉXICO

EDUCACIÓN
SECRETARÍA DE EDUCACIÓN PÚBLICA

Probability Mass Function (PMF)

Example

Figure 3.5 shows another example of two different measures $\mathbb{F}$ and $\mathbb{G}$ on the same sample space $\Omega = \{\clubsuit, \diamondsuit, \heartsuit, \spadesuit\}$. Since the PMFs of the two measures are different, even when given the same event E, the resulting probabilities will be different.

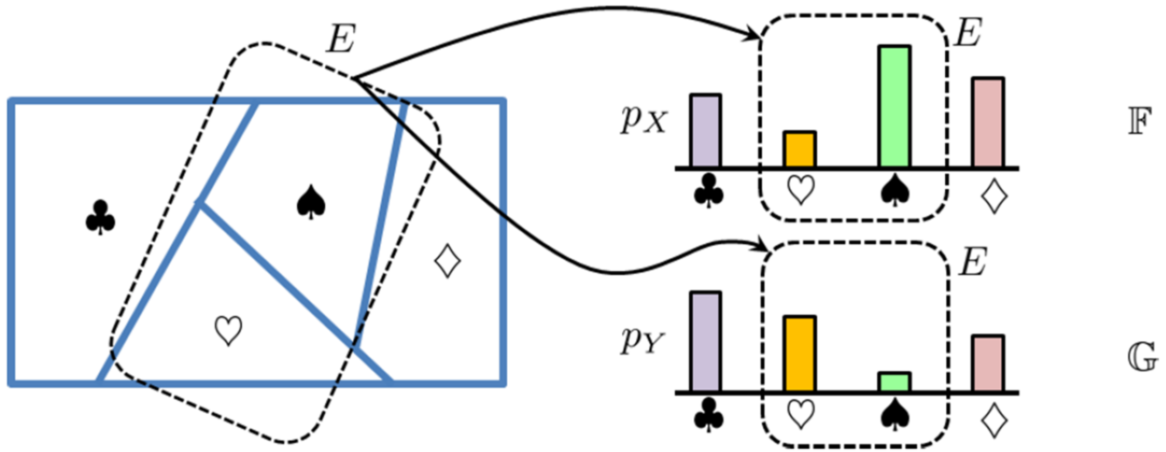


Figure 3.5: If we want to measure the size of a set E, using two different PMFs is equivalent to using two different measures. Therefore, the probabilities will be different.

1.2 El rol de la probabilidad y estadística

Probability Mass Function (PMF)

Example

Theorem 3.1. *A PMF should satisfy the condition that*

$$\sum_{x \in X(\Omega)} p_X(x) = 1.$$

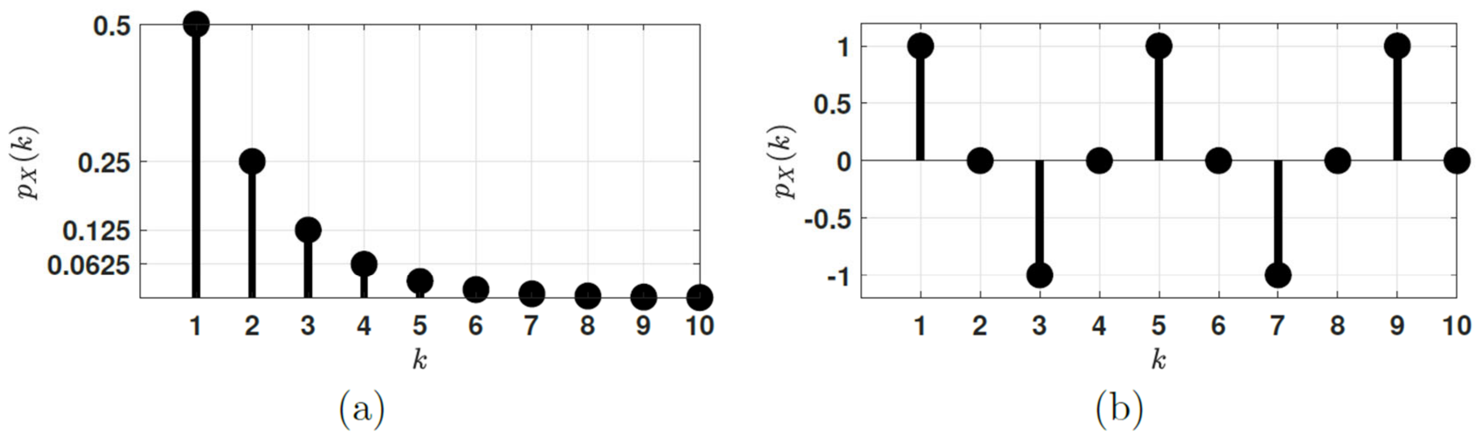


Figure 3.6: (a) The PMF of $p_X(k) = c\left(\frac{1}{2}\right)^k$, for $k = 1, 2, \ldots$. (b) The PMF of $p_X(k) = \sin\left(\frac{\pi}{2}k\right)$, where $k = 1, 2, \ldots$. Note that this is not a valid PMF because probability cannot have negative values.

1.2 El rol de la probabilidad y estadística

Expectation

Def.

Definition 3.4. *The* **expectation** *of a random variable X is*

$$\mathbb{E}[X] = \sum_{x \in X(\Omega)} x\, p_X(x).$$

Expectation is the mean of the random variable X. Intuitively, we can think of $p_X(x)$ as the percentage of times that the random variable X attains the value x. When this percentage is multiplied by x, we obtain the contribution of each x. Summing over all possible values of x then yields the mean. To see this more clearly, we can write the definition as

$$\mathbb{E}[X] = \underbrace{\sum_{x \in X(\Omega)}}_{\text{sum over all states}} \underbrace{x}_{\text{a state } X \text{ takes}} \underbrace{p_X(x)}_{\text{the percentage}}.$$

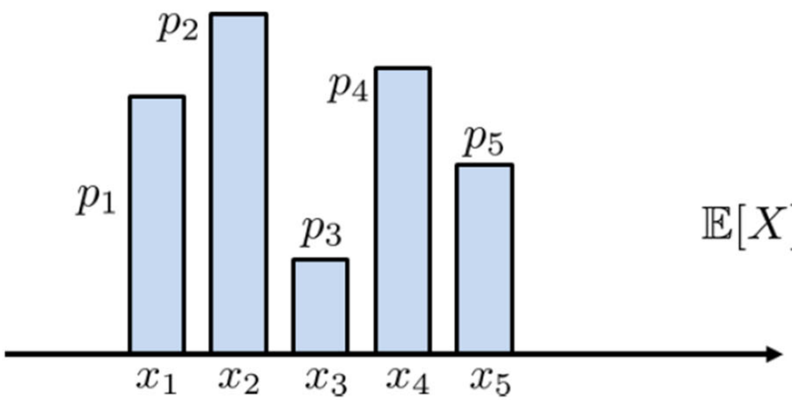
TECNOLÓGICO NACIONAL DE MÉXICO

1.2 El rol de la probabilidad y estadística

EDUCACIÓN
SECRETARÍA DE EDUCACIÓN PÚBLICA

Expectation

Figure 3.15 illustrates a PMF that contains five states $x_1, \ldots, x_5$. Corresponding to each state are $p_X(x_1), \ldots, p_X(x_5)$. For this PMF to make sense, we must assume that $p_X(x_1) + \cdots + p_X(x_5) = 1$. To simplify notation, let us define $p_i \stackrel{\text{def}}{=} p_X(x_i)$. Then the expectation of X is just the sum of the products: value (x_i) times height (p_i). This gives $\mathbb{E}[X] = \sum_{i=1}^{5} x_i p_X(x_i)$.



$$\mathbb{E}[X] = p_1 x_1 + \ldots + p_5 x_5$$

Figure 3.15: The expectation of a random variable is the sum of $x_i p_i$.

1.2 El rol de la probabilidad y estadística

Expectation

The expected value is essentially a **weighted average**.

The expected value of a random variable is also interpreted as the **long-run value** of the random variable. In other words, if we repeat the underlying random experiment several times and take the average of the values of the random variable corresponding to the outcomes, we would get the **expected value**, approximately.

Finally, the expected value of a random variable has a graphical interpretation. The expected value gives the **center of mass** of the probability mass function.

TECNOLÓGICO NACIONAL DE MÉXICO

1.2 El rol de la probabilidad y estadística

EDUCACIÓN
SECRETARÍA DE EDUCACIÓN PÚBLICA

Expectation

Example

Example 3.9. Flip an unfair coin, where the probability of getting a head is $\frac{3}{4}$. Let X be a random variable such that $X = 1$ means getting a head. Then we can show that $p_X(1) = \frac{3}{4}$ and $p_X(0) = \frac{1}{4}$. The expectation of X is therefore

$$\mathbb{E}[X] = (1)p_X(1) + (0)p_X(0) = (1)\left(\frac{3}{4}\right) + (0)\left(\frac{1}{4}\right) = \frac{3}{4}.$$

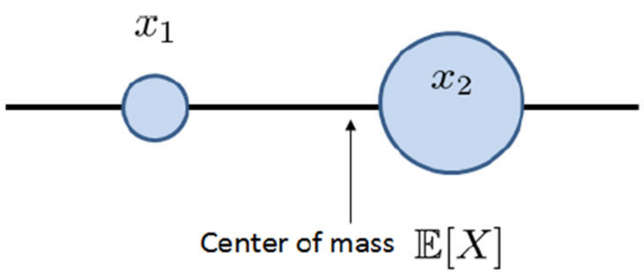


Figure 3.17: Center of mass. If a state x_2 is more influential than another state x_1, the center of mass $\mathbb{E}[X]$ will lean towards x_2.

1.2 El rol de la probabilidad y estadística

Expectation

Example

Example 3.11. Roll a die twice. Let X be the first roll and Y be the second roll. Let $Z = \max(X, Y)$. To compute the expectation $\mathbb{E}[Z]$, we first construct the sample space. Since there are two rolls, we can construct a table listing all possible pairs of outcomes. This will give us $\{(1,1),(1,2),\ldots,(6,6)\}$. Now, we calculate Z, which is the max of the two rolls. So if we have $(1,3)$, then the max will be 3, whereas if we have $(5,2)$, then the max will be 5.

Propose the table and try to find the expectation manually

TECNOLÓGICO NACIONAL DE MÉXICO

1.2 El rol de la probabilidad y estadística

EDUCACIÓN
SECRETARÍA DE EDUCACIÓN PÚBLICA

Expectation

SOLUTION

	1	2	3	4	5	6
1	1	2	3	4	5	6
2	2	2	3	4	5	6
3	3	3	3	4	5	6
4	4	4	4	4	5	6
5	5	5	5	5	5	6
6	6	6	6	6	6	6

This table tell us that Z has 6 states. The PMF of Z can be determined by counting the number of times a state shows up in the table. Thus, we can show that

$$p_Z(1) = \frac{1}{36},\ p_Z(2) = \frac{3}{36},\ p_Z(3) = \frac{5}{36},$$
$$p_Z(4) = \frac{7}{36},\ p_Z(5) = \frac{9}{36},\ p_Z(6) = \frac{11}{36}.$$

The expectation of Z is therefore

$$\begin{aligned}\mathbb{E}[Z] &= (1)\left(\frac{1}{36}\right) + (2)\left(\frac{3}{36}\right) + (3)\left(\frac{5}{36}\right) \\ &\quad + (4)\left(\frac{7}{36}\right) + (5)\left(\frac{9}{36}\right) + (6)\left(\frac{11}{36}\right) \\ &= \frac{161}{36}.\end{aligned}$$

1.2 El rol de la probabilidad y estadística

Expectation

Example

Example 3.12. Consider a game in which we flip a coin 3 times. The reward of the game is

- \$1 if there are 2 heads
- \$8 if there are 3 heads
- \$0 if there are 0 or 1 head

There is a cost associated with the game. To enter the game, the player has to pay \$1.50. We want to compute the net gain, on average.

Very important topic to introduce the bandits problem

TECNOLÓGICO NACIONAL DE MÉXICO

1.2 El rol de la probabilidad y estadística

EDUCACIÓN
SECRETARÍA DE EDUCACIÓN PÚBLICA

Expectation

SOLUTION

To answer this question, we first note that the sample space contains 8 elements: HHH, HHT, HTH, THH, THT, TTH, HTT, TTT. Let X be the number of heads. Then the PMF of X is

$$p_X(0) = \frac{1}{8}, \quad p_X(1) = \frac{3}{8}, \quad p_X(2) = \frac{3}{8}, \quad p_X(3) = \frac{1}{8}.$$

We then let Y be the reward. The PMF of Y can be found by "adding" the probabilities of X. This yields

$$p_Y(0) = p_X(0) + p_X(1) = \frac{4}{8}, \quad p_Y(1) = p_X(2) = \frac{3}{8}, \quad p_Y(8) = p_X(3) = \frac{1}{8}.$$

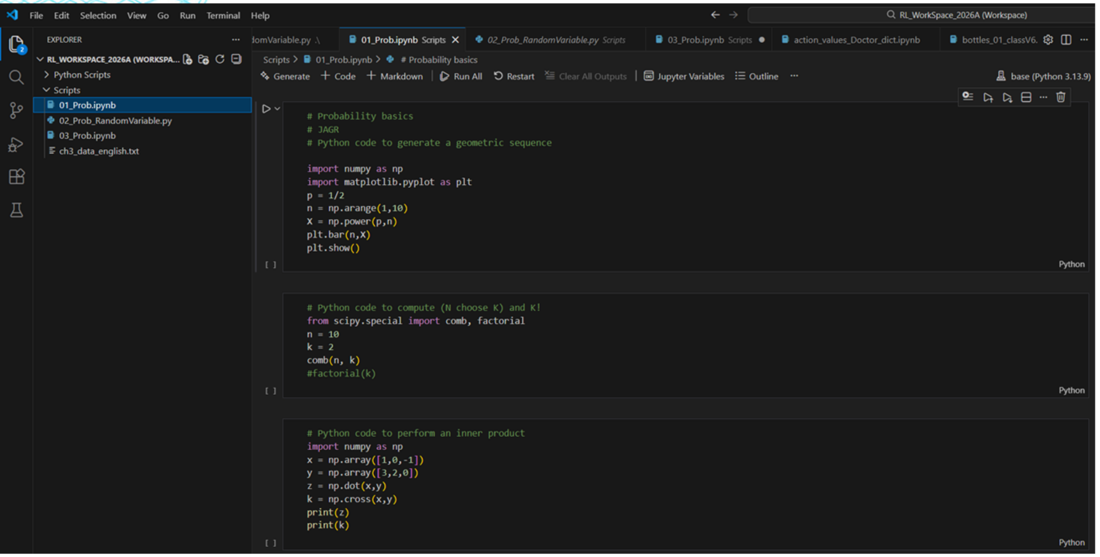
The expectation of Y is

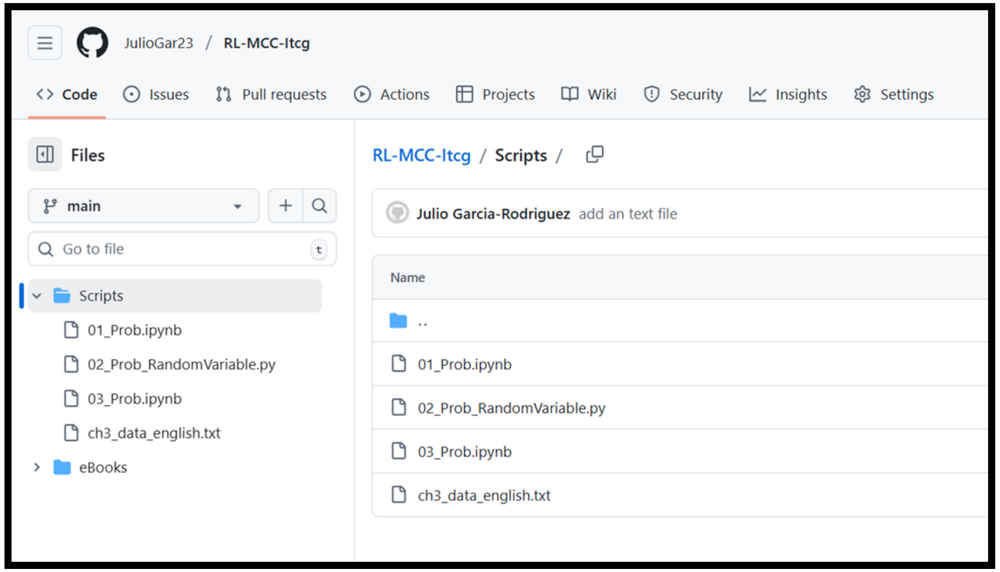
$$\mathbb{E}[X] = (0)\left(\frac{4}{8}\right) + (1)\left(\frac{3}{8}\right) + (8)\left(\frac{1}{8}\right) = \frac{11}{8}.$$

Since the cost of the game is $\frac{12}{8}$, the net gain (on average) is $-\frac{1}{8}$.

1.2 El rol de la probabilidad y estadística

Review and study Python exercises





https://github.com/JulioGar23/RL-MCC-Itcg/tree/main/Scripts