TABLE DES MATIÈRES

NOTATION ET TERMINOLOGIE

$\mathbb{R}^n$	:	L'ensemble des vecteurs avec n composants.
$\mathbb{R}^n_+$	:	L'orthant positif de l'espace $\mathbb{R}^n$.
$\mathbb{R}^{n\times m}$	:	L'espace vectoriel des matrices réelles de taille $(n \times m)$
$\mathbb{R}^n_{++}$	:	l'orthant strictement positif de l'espace $\mathbb{R}^n$.
$s.c$	:	Sous les contraintes.
x^*	:	La solution optimale du problème.
$\emptyset$	:	L'ensemble vide.
x^t	:	Le transposé du vecteur x de $\mathbb{R}^n$.
(PM)	:	Programmation mathématique.
(PL)	:	Programmation linéaire.
(DL)	:	Le problème dual de Programmation linéaire.
K.K.T	:	Karush-Kuhn-Tucker.
$\Delta x, \Delta y, \Delta s$	:	les directions de Newton.
e	:	le vecteur de $\mathbb{R}$, dont toutes les composantes sont égales à 1.
$\psi(t)$	:	Fonction noyau.
$\phi(v)$	=	$\sum_{i=1}^{n} \psi(v_i)$:La fonction barrière logarithmique de type primal-dual.

INTRODUCTION PROBABILITY AND STATISTICS

CHAPTER 1

COMBINATORIAL ANALYSIS

1.1

EXEMPLE

How many ways to order 52 cards?

Answer: 52.51.50.....1 = 52!

n hats, n people, how many ways to assign each person a hat?

Answer: $n!$

n hats, $k < n$ people, how many ways to assign each person a hat?

$n(n-1)(n-2)...(n-k+1) = n!/(n-k)!$

1.1.1 Arrangements

Définition 1.1.1 *[?] Let E be a set with n elements, an arrangement of p of these objects is an ordered sequence of p objects taken from these n objects.*

There are two types of arrangements: with and without repetition.

Arrangement without repetition

We call an arrangement without repetition of p objects chosen from n objects any ordered layout (disposition) of p objects taken from the n objects without repetitions.
The number of arrangements without repetition, noted A_p^n, is as follows:

$$A_n^p = \frac{n!}{(n-p)!} = n \times (n-1) \times (n-2) \cdots \times (n-p+1),$$

where $1 \leq p \leq n$.
In an arrangement without repetition, the p objects in the list are all distinct. This corresponds to a draw without replacement and with order.
Example How many three-letter words containing no more than one letter can be formed using the letters of the alphabet?

$$A_{26}^3 = \frac{26!}{(26-3)!} = 26 \times 25 \times 24 = 15600 mots.$$

Arrangement with repetition

We call an arrangement with repetition of p objects chosen from n objects any ordered layout (disposition) of p objects taken from the n objects with repetitions.
The number of arrangements with repetition, noted n^p, is as follows:

$$n^p = n \times n \times n \cdots \times n,$$

where $1 \leq p \leq n$.
In a non-repetition arrangement, the p objects in the list are not necessarily all distinct. This corresponds to a draw with replacement and with order.
Example How many two-letter words can be made with the letters of the alphabet?

$$26^2 = 26 \times 26.$$

1.1.2 Permutations

Définition 1.1.2 *[?] Let E a set of n objects. We call permutation of n distinct objects any ordered sequence of n objects or any arrangement n to n of these objects.*

Permutation without repetition

This is the special case of the arrangement without repetition of p objects among n objects, when p = n.
The number of permutations of n objects is: $n!$

Example The number of ways to seat eight diners (guests) around a table is: 8! = 40320.

Permutation with repetition

In the case where there are k identical objects among the n objects, then

$$\frac{n!}{k!},$$

Example The number of possible words (with or without meaning) that can be formed by permuting the 8 letters of the word "Quantity" is $\frac{8!}{2!}$ = 20160 words, we have 2 t in "Quantity".
Considering the word "Swimming", the number of possible words is $\frac{8!}{2!2!}$ = 10080 words,

because we have the i 2 times and the m 2 times.

1.1.3 Combinations

Combination without repetitions (without discounts)

Définition 1.1.3 *[?] Given a set E of n objects. We call combinations of p objects any set of p objects taken from the n objects without replacement (without discount).*

The number of combinations of p objects among n and without replacement, is:

$$C_n^p = \frac{n!}{p!(n-p)!}$$

where $1 \leq p \leq n$.
Example 1 The random drawing of 5 cards from a deck of 32 cards (poker hand) is a combination with p = 5 and n = 32. The number of possible drawings is: $C_{32}^5 = \frac{32!}{5!(32-5)!} =$ 409696 possibilities.
Example 2 Forming a delegation of 2 students from a group of 20 is a combination with p = 2 and n = 20. The number of possible delegations is $C_{20}^2 = \frac{20!}{2!(20-2)!} = 190$ possibilities.

Combination with repetitions (with discounts)

The number of combinations of p objects among n and with replacement (with discount),is:

$$C_{n+p-1}^p = \frac{(n+p-1)!}{p!(n-1)!}$$

where $1 \leq p \leq n$.

Example1 Let's make up 3-letter words from a 5-letter alphabet with discount.
The number of words is $C_{5+3-1}^3 = C_7^3 = 35$.
There are 3 possible cases:
-C_5^3 number of words of 3 different letters and without order;

-$2C_5^2$ number of words with 2 different letters and one redundant letter;

-C_5^1 number of words with 3 identical letters;

in total, we have $C_5^3 + 2C_5^2 + C_5^1 = C_7^3 = 35$ words.

Let E be a set with n elements. p-lists :We call a p-list of E any ordered sequence (the order is important) of p elements taken from n elements of E. np: number of p-lists of E. Arrangements : On appelle arrangement de p éléments, toute suite ordonnée (l?ordre est important) de p éléments distincts pris parmi n éléments de E. : number of arrangements of p elements among n elements of E. Permutations: If p = n, the arrangements of n elements among n elements will be called permutations of n elements.(A permutation is a map from 1, 2,..., n to 1, 2,..., n. There are n! permutations of n elements.) Pn = Ap n = n!: number of permutations of n elements. Combinations: A combination is a p-element subset of n elements of E. Here, the order is not important and repetition of elements is prohibited. : number of combinations of p elements among n elements of E.

CHAPTER 2

PROBABILITY SPACES

2.1 Sample spaces and events

Définition 2.1.1 *(Random Experiment:) An experiment whose outcome is uncertain before it is performed is called a random experiment.*

Définition 2.1.2 *(Sample Space) The set of all possible outcomes of the given experiment is called the sample space and is denoted by* Ω.

Each member of a sample space (an element of Ω) is called an outcome or an elementary event and is denoted by ω. **Examples**

- Coin toss: $\Omega = \{Heads, Tails\} = \{H, T\}$; $\omega = H$ is an elementary event
- Roll of a die:$\Omega = \{1, 2, 3, 4, 5, 6\}$
- Tossing of two coins: $\Omega = \{(H, H), (T, H), (H, T), (T, T)\}$
- If the experiment consists of rolling a pair of dice, then the sample space Ω consists of

the 36 pairs in the set

$$\Omega = D \times D$$

with

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

- Coin is tossed until heads appear. What is Ω? - Life expectancy of a random person. $\Omega = [1, 120]$ years

Définition 2.1.3 *(Events) An event, A, is a subset of the sample space.*

This means that event A is simply a collection of outcomes. Events are typically denoted by upper case letters, usually from the beginning of the alphabet.
An event is said to have occurred if the outcome of the experiment belongs to it.
Examples
- sure event = sample space Ω (an event for sure to occur)
- impossible event = empty set $\emptyset$ (an event impossible to occur)
- Coin toss: $\Omega = \{Heads, Tails\}$
$E = \{Heads\}$ is the event that a head appears on the flip of a coin.
- Roll of a die: $\Omega = \{1, 2, 3, 4, 5, 6\}$. Some possible events are
$E_1 = \{1\}$ (single outcome), $E_2 = \{\textit{an even number shows up}\} = \{2, 4, 6\}$ (3 outcomes), $E_3 = \{\textit{the outcome is} \geq 3\} = \{3, 4, 5, 6\}$, $E_4 = \{\textit{the outcome is} \leq 0\} = \emptyset$, (no outcome).
- Life expectancy. $\Omega = [1, 120]$.
$E = [50, 120]$ is the event that a random person lives beyond 50 years.

2.1.1 Language of Events

Typical Notation	Language of Sets	Language of Events
Ω	Whole space	Certain event
$\emptyset$	Empty set	Impossible event
A	Subset of Ω	Event that some outcome in A occurs
A^c	Complement of A	Event that no outcome in A occurs
$A \cup B$	Union	Event that an outcome in A or B or both occurs
$A \cap B$	Intersection	Event that an outcome in both A and B occurs
$A \cap B = \emptyset$	Disjoint sets	Mutually exclusive events

Définition 2.1.4 *Events A and B are disjoint if their intersection is empty,*

$$A \cap B = \emptyset.$$

Events $A_1, A_2, ..., A_n$ are mutually exclusive or pairwise disjoint if any two of these events are disjoint, i.e.,

$$A_i \cap A_j = \emptyset \; for \; any \; i \neq j.$$

Définition 2.1.5 *(sigma-algebra (σ-algebra) (σ-field)) A collection $\mathcal{F}$ of subsets of Ω is called a sigma-algebra if it satisfies*
(a) it includes the sample space : $\Omega \in \mathcal{F}$
(b) $\mathcal{F}$ is stable by countable union, i.e. if $A_1, A_2, ... \in \mathcal{F}$, then $\cup_{i=1}^{\infty} A_i \in \mathcal{F}$
(c) $\mathcal{F}$ is stable by complement, if $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$

Examples - $\mathcal{F} = \{\Phi, \Omega\}$ is the smallest sigma-algebra (Degenerate sigma-algebra).
- If $A \subset \Omega, \mathcal{F} = \{\Phi, A, A^c, \Omega\}$ is a σ-field
- $\mathcal{P}(\Omega)$ is the richest sigma-algebra. This sigma-algebra is called a power set (is the set whose elements are all the subsets of Ω).
- the Borel sigma-algebra, if $\Omega = \mathbb{R}$, and is denoted by $\mathcal{B}_{\mathbb{R}}$

Définition 2.1.6 *(Indicator function) Indicator function of an event denoted by $I_A(\omega)$ and defined as*

$$I_A(\omega) \quad = \quad \begin{cases} 1, & if \ \omega \in A \\ 0, & si \ \omega \notin A \end{cases}$$

Définition 2.1.7 *(Partition of set) A partition of a set A is a set $\{A_1, A_2, ..., A_n\}$ with the following properties:*
a. $A_i \subseteq A$, $i = 1, 2, . . . , n$, which means that A is a set of subsets.
b. $A_i \cap A_j = \emptyset$, for every $i \neq j$, $i = 1, 2, . . . , n$; $j = 1, 2, . . . , n$, which means that the subsets are mutually (or pairwise) disjoint; that is, no two subsets have any element in common.
c. $\cup_{i=1}^{n} A_i = A$, which means that the subsets are collectively exhaustive. That is, the subsets together include all possible values of the set A.

Définition 2.1.8 *(Probability and probability space) Given a σ-algebra $\mathcal{F}$ on a set Ω. A probability $\mathbb{P}$ is a function $\mathbb{P} : \mathcal{F} \longrightarrow [0, 1]$ satisfying (s.t.)*
(a) (Unit measure) The sample space has unit probability, $P(\Omega) = 1$
(b) (Sigma-additivity) if $A_1, A_2, ... \in \mathcal{F}$ is a collection of disjoint members in $\mathcal{F}$, then

$$\mathbb{P}(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$$

($\mathbb{P}$ is said to be countably additive)

The triplet $(\Omega, \mathcal{F}, \mathbb{P})$ is called a probability space.

Remarque For any event A,

$$\mathbb{P}(A) = \frac{The\ number\ of\ outcomes\ in\ A}{The\ number\ of\ outcomes\ in\ \Omega}$$

Example 0

Which of the following are Probability functions?

(i) $\Omega = \{1, 2, 3, \ldots\}, , \mathcal{F}$ is σ - field on Ω. A function $\mathbb{P}$ defined on space $(\Omega, \mathcal{F})$ as

$$\mathbb{P}(i) = \frac{1}{2^i}$$

for $i \in \Omega$

Solution

a)

$$\mathbb{P}(\Omega) = \sum_{i=1}^{\infty} \frac{1}{2^i} = 1$$

b)

$$\mathbb{P}(A) \geq 0 \ for\ all\ A \in \mathcal{F}$$

c)Let us define mutually exclusive events, $A_i = i$ we can verify countable additivity.

$$\mathbb{P}(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$$

By a) ,b)and c) $\mathbb{P}$ is Probability function.

Example 1

- Coin toss:$\Omega = \{H, T\}, \mathcal{F} = \{\Phi, H, T, \Omega\}$

$$\mathbb{P}(\Omega) = 1, \mathbb{P}(\emptyset) = 0, \mathbb{P}(H) = \frac{1}{2}, \mathbb{P}(T) = \frac{1}{2}$$

- Roll of a die:$\Omega = \{1, 2, 3, 4, 5, 6\}$.

The probability of obtaining the 1 is $1/6$. The probability of obtaining the 2 is $1/6$. In fact, the probability of obtaining any particular integer from $1, 2, ...6$, is $1/6$.

I defined the events $A = \{5, 6\}$, and $B = \{1, 3, 5\}$. We can now see that $\mathbb{P}(A) = 2/6$ and

$\mathbb{P}(B) = 3/6$.

Example 2

Two fair dice are tossed. Find the probability of each of the following events:

a. The sum of the outcomes of the two dice is equal to 7.

b. The sum of the outcomes of the two dice is equal to 7 or 11.

c. The outcome of the second die is greater than the outcome of the first die.

d. Both dice come up with even numbers.

Solution

We first define the sample space of the experiment. If we let the pair (x, y) denote the outcome, first die comes up x and second die comes up y, where $x, y \in \{1, 2, 3, 4, 5, 6\}$, then $\Omega = \{(1, 1), (1, 2), (1, 3), \cdots, (6, 6)\}$. The total number of sample points is 36.

(a) Let A_1 denote the event that the sum of the outcomes of the two dice is equal to seven. Then $A_1 = \{(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)\}$. Since the number of sample points in the event is 6, we have that $\mathbb{P}(A_1) = 6/36 = 1/6$.

(b) Let B denote the event that the sum of the outcomes of the two dice is either seven or eleven, and let A_2 denote the event that the sum of the outcomes of the two dice is eleven. Then, $A_2 = \{(5, 6), (6, 5)\}$ with 2 sample points. Thus, $\mathbb{P}(A_2) = 2/36 = 1/18$. Since B is the union of A_1 and A_2, which are mutually exclusive events, we obtain $\mathbb{P}(B) = \mathbb{P}(A1 \cup A2) = \mathbb{P}(A_1) + \mathbb{P}(A_2) = \frac{1}{6} + \frac{1}{18} = \frac{2}{9}$

(c) Let C denote the event that the outcome of the second die is greater than the outcome of the first die. Then $C = \{(1, 2), (1, 3), (1, 4), (1, 5), (1, 6), (2, 3), (2, 4), (2, 5), (2, 6), (3, 4), (3, 5), (3, 6), (4, 5), ($ with 15 sample points. Thus, $\mathbb{P}(C) = 15/36 = 5/12$.

(d) Let D denote the event that both dice come up with even numbers. Then $D = \{(2, 2), (2, 4), (2, 6), (4, 2), (4, 4), (4, 6), (6, 2), (6, 4), (6, 6)\}$ with 9 sample points. Thus, $\mathbb{P}(D) = 9/36 = 1/4$.

Proposition 2.1.1 *1-* $\mathbb{P}(\emptyset) = 0$.

2- $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$.

3- if $A_1 \subset A_2$ *then* $\mathbb{P}(A_1) \leqslant \mathbb{P}(A_2)$.

4- $\mathbb{P}(A_1 \cup A_2) = \mathbb{P}(A_1) + \mathbb{P}(A_2) - \mathbb{P}(A_1 \cap A_2)$.

5- $\mathbb{P}(\cup_{i\geq 1} A_i) \leq \sum_{i\geq 1} \mathbb{P}(A_i)$.

Preuve. 1- By the same definition, $\mathbb{P}(\Omega) = \mathbb{P}(\Omega \cup \emptyset) = \mathbb{P}(\Omega) + \mathbb{P}(\emptyset)$, because Ω and $\emptyset$ are mutually exclusive. Therefore, $\mathbb{P}(\emptyset) = 0$.

2- Recall that events A and A^c are exhaustive, hence $A \cup A^c = \Omega$. Also, they are disjoint, hence

$$\mathbb{P}(A) + \mathbb{P}(A^c) = \mathbb{P}(A \cup A^c) = \mathbb{P}(\Omega) = 1.$$

Solving this for $\mathbb{P}(A^c)$, we obtain a rule that perfectly agrees with the common sense,

$$\mathbb{P}(A^c) = 1 - \mathbb{P}(A).$$

3- If $A_1 \subset A_2$, then A_2 can be written as the union of the disjoint subsets A_1 and $A_2 - A_1$. Therefore,

$$\mathbb{P}(A_2) = \mathbb{P}(A_1) + \mathbb{P}(A_2 - A_1) \geq \mathbb{P}(A_1).$$

4- $A_1 \cup A_2$ can be written as the disjoint union of $A_1 - A_2$, $A_2 - A_1$ and $A_1 \cap A_2$. Therefore,

$$\mathbb{P}(A_1 \cup A_2) = \mathbb{P}(A_1 - A_2) + \mathbb{P}(A_2 - A_1) + \mathbb{P}(A_1 \cap A_2)$$

$$= [\mathbb{P}(A_1) - \mathbb{P}(A_1 \cap A_2)] + [\mathbb{P}(A_2) - \mathbb{P}(A_1 \cap A_2)] + \mathbb{P}(A_1 \cap A_2)$$

$$= \mathbb{P}(A_1) + \mathbb{P}(A_2) - \mathbb{P}(A_1 \cap A_2).$$

■ **Example**

Three tulip bulbs are planted in a window box. Find the probability that at least one will flower if the probability that all will fail to flower is $\frac{1}{8}$.

Sometimes calculations are made easier by using complementary events.

Solution

Définition 2.1.9 *$f(.)$ is a function defined on $\mathbb{R}$ is a called as Borel function if inverse image is a Borel set.*

2.2 Conditional Probability and independence

2.2.1 Conditional probability

Définition 2.2.1 *If E and F are two events associated with the same sample space of a random experiment, then the conditional probability of the event E under the condition that the event F has occurred (the conditional probability of E given F), written as* $\mathbb{P}(E|F)$*, is given by*

$$\mathbb{P}(E|F) = \frac{\mathbb{P}(E \cap F)}{\mathbb{P}(F)}$$

Example

A bag contains eight red balls, four green balls, and eight yellow balls. A ball is drawn at random from the bag, and it is not a red ball. What is the probability that it is a green ball.

Solution

Let G denote the event that the selected ball is a green ball, and let $\bar{R}$ denote the event that it is not a red ball. Then, $P(G) = 4/20 = 1/5$, since there are 4 green balls out of a total of 20 balls, and $P(\bar{R}) = 12/20 = 3/5$, since there are 12 balls out of 20 that are not red. Now,

$$P(G|\bar{R}) = \frac{P(G \cap \bar{R})}{P(\bar{R})}$$

But if the ball is green and not red, it must be green. Thus, we obtain that $G \cap \bar{R} = G$ and

$$P(G|\bar{R}) = \frac{P(G \cap \bar{R})}{P(\bar{R})} = \frac{P(G)}{P(\bar{R})} = \frac{1/5}{3/5} = \frac{1}{3}$$

Properties

Let E and F be events associated with the sample space Ω of an experiment. Then:

(1) $\mathbb{P}(\Omega|F) = \mathbb{P}(F|F) = 1$

(2) $\mathbb{P}(A \cup B|F) = \mathbb{P}(A|F) + \mathbb{P}(B|F) - \mathbb{P}(A \cap B|F)$.

(3) $\mathbb{P}(E^C|F) = 1 - \mathbb{P}(E|F)$

(4) $\mathbb{P}(E \cap F) = \mathbb{P}(E|F)\mathbb{P}(F) = \mathbb{P}(F|E)\mathbb{P}(E)$.

(5) $\mathbb{P}(E|F) \neq \mathbb{P}(F|E)$

Examples

Show that

$$\mathbb{P}(A \cup B|F) = \mathbb{P}(A|F) + \mathbb{P}(B|F) - \mathbb{P}(A \cap B|F)$$

Solution

$\mathbb{P}(A \cup B|F) = \dfrac{\mathbb{P}((A \cup B) \cap F)}{\mathbb{P}(F)}$

By definition of conditional prob.

$= \dfrac{\mathbb{P}((A \cap F) \cup (B \cap F))}{\mathbb{P}(F)}$

By Distributive law

$= \dfrac{\mathbb{P}(A \cap F) + \mathbb{P}(B \cap F) - \mathbb{P}(A \cap B \cap F)}{\mathbb{P}(F)}$

By Addition theorem on probability.

$= \mathbb{P}(A|F) + \mathbb{P}(B|F) - \mathbb{P}(A \cap B|F)$

By definition of conditional prob.

Examples

Probability that it rains today is 0.4; probability that it will rain tomorrow is 0.5, probability that it will rain tomorrow and rains today is 0.3.

Given that it has rained today, what is the probability that it will rain tomorrow?

Solution

Denote the events, say

A :" it rains today", $\mathbb{P}(A) = 0.4$

B : "it will rain tomorrow", $\mathbb{P}(B) = 0.5$, $\mathbb{P}(A \cap B) = 0.3$.

Required probability is

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)} = 0.75$$

Proposition 2.2.1 *If $\{A_1, A_2, ..., A_n\}$ form a partition of Ω. Let A be any event. Then*

$$\mathbb{P}(A) = \sum_{i=1}^{n} \mathbb{P}(A \cap A_i) = \sum_{i=1}^{n} \mathbb{P}(A_i) P(A|A_i)$$

Preuve. See chapter 1 (Basic Probability Concepts) ■

As a special case, B and $\bar{B}$ is a partition of Ω, so:

$$\mathbb{P}(A) = \mathbb{P}(A \cap B) + \mathbb{P}(A \cap \bar{B})$$

$$= \mathbb{P}(B)P(A|B) + \mathbb{P}(\bar{B})\mathbb{P}(A|\bar{B}) \; for \; any \; A, B.$$

2.2.2 Independent Events

Now we can give an intuitively very clear definition of independence.

Définition 2.2.2 *Let E and F be two events associated with a sample space Ω. If the probability of occurrence of one of them is not affected by the occurrence of the other, then we say that the two events are independent. Thus, two events E and F will be independent, if*
(a) $\mathbb{P}(F|E) = \mathbb{P}(F)$.
(b) $\mathbb{P}(E|F) = \mathbb{P}(E)$.

Using the multiplication theorem on probability, we have

$$\mathbb{P}(E \cap F) = \mathbb{P}(E)\mathbb{P}(F).$$

Example
There is a $0,01$ probability for a hard drive to crash. Therefore, it has two backups, each having a $0,02$ probability to crash, and all three components are independent of each other. The stored information is lost only in an unfortunate situation when all three devices crash. What is the probability that the information is saved?
Solution
Denote the events, say,
H :" hard drive crashes", B_1 = "first backup crashes", B_2 = "second backup crashes".
It is given that H, B_1, and B_2 are independent,
$\mathbb{P}(H) = 0.01$, and $\mathbb{P}(B_1) = \mathbb{P}(B_2) = 0.02$.

Applying rules for the complement and for the intersection of independent events,
$\mathbb{P}(saved) = 1 - \mathbb{P}(lost) = 1 - \mathbb{P}(H \cap B_1 \cap B_2)$
$= 1 - \mathbb{P}(H)\mathbb{P}(B_1)\mathbb{P}(B_2)$
$= 1 - (0.01)(0.02)(0.02) = 0.999996.$

Théorème 2.2.1 *If events A and B are independent so are (i) A and $\bar{B}$ (ii) B and $\bar{A}$ (iii) $\bar{A}$ and $\bar{B}$.*

Preuve. (i) Consider

$$\mathbb{P}(A \cap \bar{B}) = \mathbb{P}(A) - \mathbb{P}(A \cap B)$$

Since A and B are independent

$$\mathbb{P}(A \cap \bar{B}) = \mathbb{P}(A) - \mathbb{P}(A)\mathbb{P}(B) = \mathbb{P}(A)\mathbb{P}(\bar{B})$$

So, A and $\bar{B}$ are independent. Similarly we can prove (ii)
(iii) Consider

$$\mathbb{P}(\bar{A} \cap \bar{B}) = \mathbb{P}(A \bar{\cup} B) = 1 - \mathbb{P}(A \cup B)$$

$= 1 - [\mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)]$
$= [1 - \mathbb{P}(A)][1 - \mathbb{P}(B)]$
$= \mathbb{P}(\bar{A})\mathbb{P}(\bar{B})$
So, $\bar{A}$ and $\bar{B}$ are independent ■

Remarque 2.2.2 *If A,B,C are three events*
-They are pairwise independent if
$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$
$\mathbb{P}(A \cap C) = \mathbb{P}(A)\mathbb{P}(C)$
$\mathbb{P}(B \cap C) = \mathbb{P}(B)\mathbb{P}(C)$
-They are completely independent (are said to be mutually independent) if
$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$
$\mathbb{P}(A \cap C) = \mathbb{P}(A)\mathbb{P}(C)$

$\mathbb{P}(B \cap C) = \mathbb{P}(B)\mathbb{P}(C)$

and $\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C)$

Théorème 2.2.3 (Theorem of Total Probability) *Let $\{A_1, A_2, ..., A_n\}$ be a partition of the sample space Ω. Let A be any event associated with Ω, then*

$$\mathbb{P}(A) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)\mathbb{P}(A|A_i)$$

Théorème 2.2.4 (Bayes's Theorem) *If $\{A_1, A_2, ..., A_n\}$ are mutually exclusive and exhaustive events associated with a sample space, and A is any event of non zero probability, then*

$$\mathbb{P}(A_i|A) = \frac{\mathbb{P}(A_i)\mathbb{P}(A|A_i)}{\sum_{i=1}^{\infty} \mathbb{P}(A_i)\mathbb{P}(A|A_i)}$$

Example

Three people X ,Y ,Z have been nominated for the Manager's post. The chances for getting elected for them are 0.4, 0.35 and 0.25 respectively. If X will be selected the probability that he will introduce Bonus scheme is 0.6 the respective chances in respect of Y and Z are 0.3 and 0.4 respectively. If it w known that Bonus scheme has been introduced, what is the probability that X is selected as a Manager?

Solution

see probability-theory.pdf+++++ P47

CHAPTER 3

RANDOM VARIABLES

Définition 3.0.1 *A random variable X (r.v.) is a measurable function from a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *to the reals, i.e., it is a function*

$$X : \Omega \longrightarrow \mathbb{R}$$

such that $\forall B \in \mathcal{B}_{\mathbb{R}}$

$$X^{-1}(B) \in \mathcal{F}.$$

A random variable (r.v.) is defined as a function that associates a number to each element of the outcome space. Hence, any X,

$$X : \Omega \longrightarrow \mathbb{R}$$

is a random variable.

Random variables are usually denoted by $X, Y, Z, \cdots$.

3.1 Discrete random variables

As before, suppose Ω is a sample space.

Définition 3.1.1 *A Random Variable X is said to be discrete if it takes only the values of a finite or countably infinite set* $\{0, 1, 2, \cdots\}$

Exemples 3.1.1 *1. Tossing 2 coins simultaneously*

$$\Omega = \{HH, HT, TH, TT\}$$

Let the random variable be getting number of heads then
$X(HH) = 2; X(HT) = 1; X(TH) = 1; X(TT) = 0$
So

$$X(\Omega) = \{0, 1, 2\}$$

2. Flip a coin three times, let X be the number of heads in three tosses.
$\Omega = \{HHH, HHT, HTH, THH, TTH, THT, HTT, TTT\}$
$X(HHH) = 3;, , , , , , X(TTT) = 0$
So

$$X(\Omega) = \{0, 1, 2, 3\}$$

3. Sum of the two numbers on throwing 2 dice

$$X(\Omega) = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$$

4. Number of children in a family

3.1.1 Discrete Probability Distributions

Définition 3.1.2 *a probability distribution consists of value that a random variable can assume and the corresponding probabilities of the values.*

Value of X	x_1	x_2	x_3	$\cdots$	x_k
Probability	p_1	p_2	p_3	$\cdots$	p_K

where

$$0 \leq pi \leq 1 \; and \; p_1 + p_2 + \cdots + p_k = 1$$

We have

$$P_X(x) = \mathbb{P}(X = x) = \mathbb{P}(X^{-1}(x)) = \mathbb{P}(\{\omega \in \Omega / X(\omega) = x\})$$

Note that the pre-image $X^{-1}(x)$ is the event $\{\omega \in \Omega : X(\omega) = x\}$.

Example: Consider the experiment of tossing a coin three times. Let X is the number of heads. Construct the probability distribution of X.

Solution: - First identify the possible value that X can assume. -Calculate the probability of each possible distinct value of X and express X in the form of frequency distribution.

Value of X (X=x)	0	1	2	3
Probability (P(X=x))	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

3.1.2 Cumulative Distribution Function (C. D. F.) of a disc. r.v.

Définition 3.1.3 *The cumulative distribution function F of a discrete random variable X is*

$$F(x_i) = \mathbb{P}(\{\omega \in \Omega / X(\omega) \leqslant x_i\}) = \sum_{x \leqslant x_i} \mathbb{P}(\{\omega \in \Omega / X(\omega) = x\}) = \mathbb{P}(X \leq x) = \sum_{x \leqslant x_i} \mathbb{P}(X = x)$$

Example Suppose that X has the following probability distribution

$p_X(0) = \mathbb{P}(X = 0) = \frac{1}{8}$

$p_X(1) = \mathbb{P}(X = 1) = \frac{3}{8}$

$p_X(2) = \mathbb{P}(X = 2) = \frac{3}{8}$

$p_X(3) = \mathbb{P}(X = 3) = \frac{1}{8}$.

Find the CDF for X and plot the graph of the CDF.

Solution: summing up the probabilities up to the value of x we get the following

$$F_X(x) = \begin{cases} 0, & for\ x < 0 \\ \frac{1}{8}, & for\ 0 \leqslant x < 1 \\ \frac{4}{8}, & for\ 1 \leqslant x < 2 \\ \frac{7}{8}, & for\ 2 \leqslant x < 3 \\ 1, & si\ \ x \geqslant 3. \end{cases}$$

3.2 Continuous Random Variables

Définition 3.2.1 *Its set of possible values is the set of real numbers* $\mathbb{R}$*, one interval, or a disjoint union of intervals on the real line (e.g.,* $[0, 10] \cup [20, 30]$*).*

Examples:

- Height of students at certain college.

- Life time of light bulbs.

Définition 3.2.2 *Let X be a continuous r.v. Then a probability distribution or probability density function (pdf) of X is a function* f *such that for any two numbers a and b with* $a \leqslant b$*, we have*

$$\mathbb{P}(a \leqslant X \leqslant b) = \int_a^b f(x)dx$$

The probability that X is in the interval $[a, b]$ can be calculated by integrating the pdf of the r.v. X.

The probability that X takes on a value in the interval [a, b] is the area above this interval and under the graph of the density function ($\mathbb{P}(a \leqslant X \leqslant b)$ = the area under the density curve between a and b)

For $f(x)$ to be a legitimate pdf, it must satisfy the following two conditions:

- $f(x) \geq 0$, for all x,

- $\int_{-\infty}^{+\infty} f(x)dx = 1$.

Example1 Suppose that the error in the reaction temperature, for a controlled laboratory experiment is a continuous random variable X having the probability density function

$$f(x) \quad = \quad \begin{cases} \frac{x^2}{3}, & for \ \ -1x < 2 \\ 0, & otherwise. \end{cases}$$

(i) Verify that $f(x)$ is a density function.

(ii) Find $\mathbb{P}(0 \leqslant X \leqslant 1)$.

(iii) Find $\mathbb{P}(0 < X < 1)$.

Solution(a) $f(x) > 0$ because $f(x)$ is quadratic function.

(b)$\int_{-\infty}^{+\infty} f(x)dx = \int_{-1}^{2} \frac{x^2}{3}dx = !!!!! = 1$.

(ii)$\mathbb{P}(0 \leqslant X \leqslant 1) = \int_0^1 f(x)dx = \int_0^1 \frac{x^2}{3}dx = !!!!! = \frac{1}{9}$.

(iii) By the same way, we have $\mathbb{P}(0 < X < 1) = \frac{1}{9}$.

The (cumulative) distribution function (cdf) for random variable X is

$$F(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^{x} f(t)dt,$$

and has properties

-$\lim_{x \longrightarrow -\infty} F(x) = 0$,

-$\lim_{x \longrightarrow +\infty} F(x) = 1$,

-if $x_1 < x_2$, then $F(x_1) < F(x_2)$; that is, F is nondecreasing,

-$\mathbb{P}(a < X < b) = \mathbb{P}(X < b) - \mathbb{P}(X < a) = F(b) - F(a) = \int_a^b f(x)dx$,

-$F^{'}(x) = f(x)$

- However, for a continuous random variable, $\mathbb{P}(X = a) = 0$

Example 2 For the density function of Example 1, find F(x)

Solution By definition, we have

$F(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^{x} f(t)dt = \int_{-1}^{x} \frac{t^2}{3}dt = *****= \frac{x^3}{9} + \frac{1}{9}$

Remeque

Endpoints are not important for continuous r.v.s.

Endpoints are very important for discrete r.v.s.

Example of calculations with the p.d.f.

Let

$$f(x) = \begin{cases} ke^{-2x}, & for\ 0 < x < \infty \\ 0, & otherwise. \end{cases}$$

(i) Find the constant k.

(ii) Find $\mathbb{P}(1 < X \leq 3)$.

(iii) Find the cumulative distribution function, $F_X(x)$, for all x.

3.3 Expectation

The expectation of a r.v. X is a real number computed by

$$E[X] = \sum_{k \in X(\Omega)} k\mathbb{P}(X = k).$$

if X is discrete and

$$E[X] = \int_{\mathbb{R}} x f(x) dx.$$

if X is continuous. Intuitively, the expectation is the average value of the r.v., more precisely it is a weighted average of the values of the r.v., where the weights are the probabilities of the outcomes. The expectation is also called mean, expected value and sometimes average.

Example If we toss a coin and X is 1 if we have heads and 0 if we have tails, what is the expectation of X?

Solution:

$$p_X(x) = \begin{cases} \frac{1}{2}, & for\ x = 1 \\ \frac{1}{2}, & for\ x = 0 \\ 0, & si\ \ otherwise. \end{cases}$$

Hence $E(X) = (1)(\frac{1}{2}) + (0)(\frac{1}{2}) = \frac{1}{2}$.

Expected Value Rule for Functions of Random Variables

Let X be a random variable with PMF p_X,and let $g(X)$ be a function of X. Then,the expected value of the random variable $g(X)$ is given by

$$E[g(X)] = \sum g(x)\mathbb{P}_X(x).$$

3.4 Variance

The variance $var(X)$ of a random variable X is defined by

$$var(X) = E[X^2] - (E[X])^2.$$

The square root of$Var(X)$ is called the standard deviation (SD) of X

$$\sigma_X = \sqrt{var(X)}.$$

Example We toss a fair coin and let $X = 1$ if we get heads, $X = -1$ if we get tails. Then E(X) = 0, so $VarX = E(X)^2 = (1)^2\frac{1}{2} + (-1)^2\frac{1}{2} = 1$.

3.4.1 Properties of expectation and variance

All properties of expectation and variance are exactly the same for continuous and discrete random variables.
For any random variables, X, Y , and $X_1, ..., X_n$, continuous or discrete, and for constants a and b:

- E(aX + b) =aE(X) + b.
- E(ag(X) + b) =aE(g(X)) + b.
- E(X + Y) =E(X) + E(Y).
- $E(X_1 + ... + X_n) = E(X_1) + ... + E(X_n)$.
- $Var(aX + b) = a^2Var(X)$.
- $Var(ag(X) + b) = a^2Var(g(X))$.
-$\sigma_{aX+b} =| a | \sigma_X$

The following statements are generally true only when X and Y are INDEPENDENT:
- E(XY) =E(X)E(Y) when X, Y independent.
- Var(X + Y) =Var(X) + Var(Y) when X, Y independent.

Question 1
The probability distribution of a discrete random variable X is given by

x	0	1	2	3	4
P(X=x)	$\frac{3}{8}$	$\frac{1}{3}$	$\frac{1}{4}$	a	$\frac{1}{24}$

where a is a positive constant.
a) Explain why $a = 0$.
b) Find the value of $E(X)$.
c) Calculate $Var(X)$.
Solution $\sum P(X = x) = 1$, $E(X) = 1$, $Var(X) + 1$
Question 2

The probability distribution of a discrete random variable X is given by

x	0	1	2	3
P(X=x)	$\frac{1}{12}$	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{3}$

Find, showing full workings where appropriate, the value of

a) $P(1 < X \leqslant 3)$.

b) $F(1.8)$.

c) $E(X)$.

d) $Var(X)$.

e) $E(2X - 3)$.

f) $Var(2X - 3)$.

Solution

$P(1 < X \leqslant 3) = \frac{2}{3}$, $F(1.8) = \frac{1}{8}$, $E(X)= \frac{23}{12}$, $Var(X)$.

$E(2X?3)$. e) $E(2X - 3)$.

f) $Var(2X - 3)$.

See Discrete- rando dans le birou

CHAPTER 4

IMPORTANT PROBABILITY DISTRIBUTIONS

4.1 Discrete probability distributions

4.1.1 The uniform distribution

If all the outcomes have equal probability, i.e $\mathbb{P}_1 = \mathbb{P}_2 = \cdots = \mathbb{P}_n = \frac{1}{n}$, then the distribution is called a **uniform distribution**.

4.1.2 The Bernoulli distribution

If an experiment has two possible outcomes [success and failure] with their probabilities p and $q = 1 - p$ respectively , then the number of success [x=1 or x=0] has a Bernoulli distribution , and denoted by :

$x \sim B(p)$, where p is the parameter of this distribution.
The probability mass function is given by:

$$f(x;p) = p^x.(1-p)^{1-x}; x = 0;1$$

E(X) = p , Var(X) = pq

4.1.3 The Binomial distribution (or Repeated Bernoulli Trials)

Then the probability of getting k successes in n trials becomes:

$$P(X = k) = C_n^k p^k q^{n-k}$$

We write : $X \sim B(n, p)$
Remark: If X is a binomial random variable with parameters n and p then
E(X) = np , Var(X) = npq

4.1.4 The geometric distribution

The geometric distribution tells us the probability that the first occurrence of success requires k number of independent trials, each with success probability p.
This discrete probability distribution is represented

$$P(X = k) = pq^{k-1}$$

We write : $X \sim G(p)$
For example,
1. Toss a coin repeatedly. Let X = number of tosses to first head.
Remark: If X is a binomial random variable with parameters n and p then
$E(X) = \frac{1}{p}$, $Var(X) = \frac{q}{p^2}$

4.1.5 The Poisson distribution

A random variable X is said to have a Poisson distribution if its probability distribution is given by:

$$P(X = k) = \frac{\lambda^k e^{-1}}{k*}$$

Where $k \in \mathbb{N}$ and λ = the average number.
We write $X \sim \mathcal{P}(\lambda)$ The Poisson distribution depends only on the average number of occurrences per unit time of space.
If X is a Poisson random variable with parameter λ then $E(X) = \lambda$, $Var(X) = \lambda$

4.2 Continuous probability distributions

4.2.1 Continuous Uniform distribution

X has a Uniform distribution on the interval $[a, b]$ if X is equally likely to fall anywhere in the interval $[a, b]$.
We write $X \sim$ Uniform[a, b], or $X \sim U_{[a,b]}$.

Probability density function, $f_X(x)$
The PDF of a uniform distribution is given by:

$$f(x) = \begin{cases} \frac{x-a}{b-a}, & if\ a \leqslant x \leqslant b \\ 0 & otherwise. \end{cases}$$

If $X \sim U_{[a,b]}$, then

Distribution function, $F_X(x)$ The entire cdf is

$$F_X(x) = \begin{cases} 0, & for\ x < a \\ \frac{x-a}{b-a}, & for\ a \leq x < b \\ 1 & si\ \ x \geqslant b. \end{cases}$$

Mean and variance:
If $X \sim U_{[a,b]}$, $E(X) = \frac{a+b}{2}; var = \frac{(b-a)^2}{12}$

4.2.2 The Normal distribution (Gaussian)

The normal distribution is probably the most important distribution in all of probability A continuous r.v. X is said to have a normal distribution with parameters μ and $\sigma > 0$ (or μ and σ^2), if the pdf of X is

$$f(x) = \frac{1}{\sigma\sqrt{2\Pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}$$

We write $X \sim \mathcal{N}(\mu; \sigma^2)$.
The normal distribution with parameter values $\mu = 0$ and $\sigma = 1$ is called the standard normal distribution.

4.2.3 The Exponential distribution

Probability density function:

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & for\ x \geqslant 0 \\ 0 & si\ \ x < 0. \end{cases}$$

Distribution function:

$$F_X(x) = \begin{cases} 1 - e^{-\lambda x}, & for\ x \geqslant 0 \\ 0 & si\ \ x < 0. \end{cases}$$

We write $X \sim e(\lambda)$.

When X $X \sim e(\lambda)$, then:$E(X) = \frac{1}{\lambda}$; $Var(X) = \frac{1}{\lambda^2}$.

Proof : $E[X] = \int_{\mathbb{R}} xf(x)dx = \int_0^\infty x\lambda e^{-\lambda x}dx.$

Integration by parts: recall that $\int uv'dx = uv - \int vu'ddx$.

Let $u = x$, so $u' = 1$, and let $v' = \lambda e^{-\lambda x}$, so $v = e^{-\lambda x}$

Then $E(X) = \int_0^\infty x\lambda e^{-\lambda x}dx =$

CHAPTER 5

CHARACTERISTIC FUNCTIONS

Définition 5.0.1 *The characteristic function of a a random variable X is the function* φ_x : $\mathbb{R} \longrightarrow \mathbb{C}$ *defined by*

$$\varphi_x(t) = \mathbb{E}(e^{itX})$$

CHAPTER 6

LIMIT THEOREMS

6.1 Markov and Chebyshev Inequalities

Théorème 6.1.1 (Markov Inequality) *If arandom variable X can only take nonnegative values,then*

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}(X)}{a}$$

for all $a > 0$

Théorème 6.1.2 (Chebyshev Inequality) *If X is arandom variable with mean* μ *and variance* σ^2*,then*

$$\mathbb{P}(|X - \mu| \geq c) \leq \frac{\sigma^2}{C^2}$$

for all $c > 0$

6.2 The Weak Law of Large Numbers

Théorème 6.2.1 (The Weak Law of Large Numbers) *Let X_1,X_2,...be independent identically distributed random variables with mean μ. For every $\epsilon > 0$, we have*

$$\mathbb{P}(|M_n - \mu| \geq \epsilon) = \mathbb{P}(|\frac{X_1 + X_2 + \cdots + X_n}{n} - \mu| \geq \epsilon) \longrightarrow 0$$

as $n \longrightarrow \infty$

6.3 Convergence in Probability

see : d973b10c2587781f86ca4f2aff49098f P 40

6.4 The Central Limit Theorem

see : d973b10c2587781f86ca4f2aff49098f

CONCLUSION

BIBLIOGRAPHY