Moulay El Mehdi Falloul

Théorie des probabilités et de la statistique

Edilivre

Introduction

La Probabilité et les statistiques sont deux disciplines des mathématiques associées et indépendants à la fois. L'analyse statistique utilise souvent la théorie des probabilités. En outre, beaucoup de sujets dans les statistiques sont indépendants de la théorie des probabilités. La Probabilité est la mesure de la probabilité qu'un événement se produira. La Probabilité est utilisée pour quantifier une attitude d'esprit envers certaines propositions dont la vérité n'est pas certaine. La certitude que nous adoptons peut être décrite en termes de mesure numérique entre 0 et 1 (où 0 indique l'impossibilité et 1 indique la certitude). Un exemple simple du calcul des probabilités est celui du jet d'une pièce de monnaie. Puisque les 2 résultats sont réputées équiprobables, la probabilité de « face » est égale à la probabilité de « pile » et chaque probabilité est égale à 1/2 ou de façon équivalente elle est égale à 50 % de chance de « pile » ou « face ». La théorie des probabilités est largement utilisée dans beaucoup de domaines d'étude comme les mathématiques, les statistiques, les sciences économiques, la biologie, le jeu du hasard, la physique, l'intelligence artificielle, l'actuariat, l'informatique, l'aide à la décision, la sociologie. La théorie des probabilités est également utilisée pour décrire les régularités des systèmes complexes.

Les Statistiques est l'étude de la collecte, de l'analyse, de l'interprétation, de la présentation et l'organisation des données. Dans l'application des statistiques, par exemple, dans un problème scientifique, industriel ou social, il faut tout d'abord une population ou un processus à étudier. Les populations peuvent être des sujets divers tels que « toutes les personnes vivant dans un pays » ou « chaque atome composant un cristal ». Elles abordent tous les aspects des données y compris la planification de la collecte de données sur le plan de la conception des sondages et des études empiriques.

Cet ouvrage comporte 10 chapitres qui portent sur les principales théories des probabilités et des statistiques ; l'analyse combinatoire, les théorèmes fondamentaux des probabilités, les variables aléatoires et lois de probabilités, les théories de l'échantillonnage et de l'estimation, les tests statistiques.

Chapitre I
Concepts et théorèmes généraux

I. L'analyse combinatoire

L'analyse combinatoire est une branche des mathématiques qui étudie comment compter les objets. Elle fournit des méthodes de dénombrements particulièrement utiles en théorie des probabilités. Les probabilités dites combinatoires utilisent constamment les formules de l'analyse combinatoire développées dans ce chapitre. Un exemple des applications intéressantes de cette dernière est la démonstration du développement du binôme de Newton utilisé dans le calcul des probabilités d'une loi binomiale.

On suppose que E est un ensemble fini non vide de n éléments.

Par exemple, on peut imaginer que E est une urne contenant n boules numérotées de 1 à n.

I.1 Les p-listes d'éléments d'un ensemble de n éléments

(Modèle : p tirages d'une boule parmi n, avec ordre et remise.)
p désigne un naturel supérieur ou égal à 1. Il peut être supérieur à n.
(Les p-listes d'éléments de E sont les éléments de E^p.)

On tire une boule. On note son numéro. On la remet dans l'urne.
On fait de même pour une $2^{\text{ième}}$ boule, puis pour une $3^{\text{ième}}$,…, enfin pour une $p^{\text{ième}}$.
On obtient ainsi une suite ordonnée de p numéros compris entre 1 et n, avec d'éventuelles répétitions.
C'est une **p-liste** d'éléments de {1,2,3,…,n}.

Il y a n choix possibles du premier numéro.
Pour chacun de ces n choix, il y a n choix possibles du second numéro.
Il y a donc n^2 façons de choisir les 2 premiers numéros.
Pour chacun de ces n^2 choix, il y a n choix possibles du $3^{\text{ième}}$ numéro.
Il y a donc n^3 façons de choisir les 3 premiers numéros.
Pour chacun de ces n^3 choix, il y a n choix possibles du $4^{\text{ième}}$ numéro.
Il y a donc n^4 façons de choisir les 4 premiers numéros.
etc.

On constate que :

Les p-listes d'éléments d'un ensemble de n éléments sont au nombre de np.

** Exercice : Le loto sportif*

Dans le jeu du loto sportif, le parieur doit remplir une grille où il indique les résultats qu'il prévoit pour treize matchs de football. Pour chacun des treize matchs, trois réponses sont possibles :

l'équipe 1 est annoncée comme gagnante (réponse [1]),
le résultat prévu est un match nul (réponse [N]),
L'équipe 2 est annoncée comme gagnante (réponse [2]).

Ces trois réponses recouvrent toutes les éventualités et, à l'issue du match, une et une seule se trouvera réalisée.

Voici un extrait de grille :

N°	Equipe 1	Equipe 2	Pronostic
1	Nantes	Marseille	[1] [N] [2]
2	Strasbourg	Auxerre	[1] [N] [2]
……	…………….	……………	…………….
13	Bordeaux	Metz	[1] [N] [2]

La règle du jeu est la suivante : sur chacune des treize lignes, le parieur coche une et une seule des trois cases [1], [N], [2] correspondant au résultat qu'il prévoit. C'est ce qu'on appelle remplir la grille.

1 De combien de façons différentes peut-on remplir la grille ?
2 Dénombrer les grilles pour lesquelles, à l'issue des matchs :
a) toutes les réponses sont exactes ;

b) toutes les réponses sont fausses ;

c) les trois premières réponses sont fausses et les dix autres exactes ;

d) trois réponses et trois seulement sont fausses.

3 Pour gagner au loto sportif, il faut avoir au moins onze réponses exactes. Quel est le nombre de grilles gagnantes ?

Il est possible de calculer le nombre des parties d'un ensemble de n éléments par une méthode analogue :

On imagine que les n éléments sont numérotés de 1 à n.

On se propose de définir une partie de l'ensemble.

Il y a 2 possibilités pour le premier élément : le prendre ou le laisser.

Pour chacune de ces 2 possibilités, il y a 2 possibilités pour le second : le prendre ou pas.

Il y a donc 2^2 possibilités pour les 2 premiers éléments.

Pour chacune de ces 2^2 possibilités, il y a 2 possibilités pour le $3^{\text{ième}}$: le prendre ou pas.

Il y a donc 2^3 possibilités pour les 3 premiers éléments. etc.

Ainsi :

Il y a 2n parties dans un ensemble de n éléments.

I.2 Les p-arrangements d'éléments d'un ensemble de n éléments

(Modèle : p tirages d'une boule parmi n, avec ordre mais sans remise.)

p désigne un naturel compris entre 1 et n.

On tire une boule. On note son numéro. On ne la remet pas dans l'urne.

On fait de même pour une $2^{\text{ième}}$ boule, puis pour une $3^{\text{ième}}$,…, enfin pour une pième.

On obtient ainsi une suite ordonnée de p numéros compris entre 1 et n, deux à deux distincts.

C'est un **p-arrangement** d'éléments de {1,2,3,…,n}.

Il y a n choix possibles du premier numéro.

Pour chacun de ces n choix, il y a (n – 1) choix possibles du second numéro.

Il y a donc n(n – 1) façons de choisir les 2 premiers numéros.

Pour chacun de ces n(n – 1) choix, il y a (n – 2) choix possibles du $3^{\text{ième}}$ numéro.

Il y a donc n(n – 1)(n – 2) façons de choisir les 3 premiers numéros.

Pour chacun de ces n(n – 1)(n – 2) choix, il y a (n – 3) choix possibles du $4^{\text{ième}}$ numéro.

Il y a donc n(n – 1)(n – 2)(n – 3) façons de choisir les 4 premiers numéros.

etc.

On constate que :

Les p-arrangements d'éléments d'un ensemble de n éléments sont au nombre de :

$$\underbrace{n(n-1)(n-2)...(n-p+1)}_{\text{p facteurs}}$$

La différence entre une p-liste et un p-arrangement est que les répétitions sont possibles pour les p-listes, mais impossibles pour les p-arrangements. Par exemple (1, 1, 2) est une 3-liste mais pas un 3-arrangement.

(2, 1, 3) est à la fois une 3-liste et un 3-arrangement.

Tout p-arrangement est une p-liste.

Les p-arrangements sont les p-listes sans répétition.

** Exercice : le tiercé*

20 chevaux sont au départ.

Jouer, c'est prévoir dans l'ordre les numéros des 3 chevaux qui arriveront en tête.

Combien y a-t-il de jeux ? De jeux gagnants dans l'ordre ? De jeux gagnants dans le désordre ?

I.3 Les permutations des éléments d'un ensemble fini. Les factorielles

Le nombre de façons de ranger les n éléments de E est aussi le nombre de n-arrangements d'éléments de E : il s'agit en effet de choisir sans remise un 1^{er} élément puis un $2^{\text{ième}}$ puis un $3^{\text{ième}}$, etc., jusqu'à l'épuisement de l'ensemble.

Ce nombre est : n(n – 1)(n – 2)(n – 3)…1. Effectuons le produit de la droite vers la gauche :

nous reconnaissons le produit des entiers depuis 1 jusqu'à n compris.

Par définition, ce nombre est **la factorielle de n.**

Il y a n! façons de ranger n éléments. n! = 1 2 3 … (n-1) n
La **suite** des factorielles peut être définie de la façon suivante :
1 ! = 1 et, pour tout entier strictement positif n : (n + 1)! = n! × (n + 1).

En effet, le produit des entiers de 1 à n + 1 est
Le produit par (n + 1) du produit des entiers de 1 à n.
La seule façon de rendre cette égalité vraie aussi pour n = 0, c'est de poser :

$$0! \times 1 = 1!$$

c'est à dire :

$$0! = 1.$$

La **définition par récurrence** de la suite des factorielles est donc celle-ci :

0! = 1 et (n*N) ((n + 1)! = n! (n + 1))

Il est temps de définir le nombre de p-arrangements d'éléments de E à l'aide des factorielles.

Ce nombre est désigné par A_n^p.

$$A_n^p = n!/(n-p)! \qquad (0 \le p \le n)$$

Ceci est vrai même si p ou n ou les deux sont nuls ; en particulier :

$$A_n^0 = A_0^0 = 1;\ A_n^n = n!$$

I.4 Les p-combinaisons d'éléments d'un ensemble de n éléments

Dans ce paragraphe, l'entier p est inférieur ou égal à l'entier n.
Les p-combinaisons d'éléments d'un ensemble E de n éléments sont les parties de E à p éléments.

On ne tient pas compte de l'ordre.

La différence entre une p-combinaison et un p-arrangement est que dans un p-arrangement on tient compte de l'ordre, alors que dans une p-combinaison, on n'en tient pas compte.
Par exemple, les 3-arrangements (1,2,3) et (2,1,3) sont différents,
alors que les 3-combinaisons {1,2,3} et {2,1,3} sont les mêmes.

Dans les p-combinaisons, il n'y a ni ordre ni répétition.

A la 3-combinaison {1,2,3} on peut associer les 6 3-arrangements (1,2,3), (1,3,2), (2,1,3), (2,3,1), (3,1,2), (3,2,1).

Le nombre 6 est le nombre d'ordres possibles pour les 3 éléments 1,2,3, c'est à dire 3 !

Plus généralement, à toute p-combinaison correspondent autant de p-arrangements que d'ordres possibles pour p éléments, c'est à dire p !.

Il y a donc p! fois plus de p-arrangements que de p-combinaisons.

Il y a p! fois moins de p-combinaisons que de p-arrangements.

Le nombre de p-combinaisons d'un ensemble de p éléments est noté C_n^p ou encore : $\begin{pmatrix} \mathrm{n} \\ \mathrm{p} \end{pmatrix}$

On retiendra les égalités suivantes :

Si ($0 \leq p \leq n$) alors :

$$C_n^p = \begin{pmatrix} \mathrm{n} \\ \mathrm{p} \end{pmatrix} = \frac{A_n^p}{p!} = \frac{n!}{p!(n-p)!} = \frac{\overbrace{n(n-1)(n-2)\ldots(n-p+1)}^{\text{p facteurs}}}{p!}$$

(La dernière écriture est la plus convenable dans les calculs numériques.)

En particulier : $C_0^0 = C_n^0 = C_n^n = 1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \mathrm{n} \\ 0 \end{pmatrix} = \begin{pmatrix} \mathrm{n} \\ \mathrm{n} \end{pmatrix}$

Dans tous les autres cas : $C_n^p = \begin{pmatrix} \mathrm{n} \\ \mathrm{p} \end{pmatrix} = 0$

I.5 Les combinaisons avec répétitions

Considérons la situation suivante : 4 clients viennent se désaltérer dans un débit de boissons. Ils s'assoient à une même table. Chacun souhaite commander une unique boisson.

Il y a 10 types différents de consommations.

Quel est le nombre de plateaux différents que peut composer le barman pour satisfaire les clients ?

Le serveur choisit 4 boissons parmi les 10 types. Il ne tient pas compte de l'ordre.

Mais il accepte les répétitions, puisque plusieurs clients peuvent désirer la même type de boisson. Son choix est ***une 4-combinaison avec répétitions*** d'éléments pris dans un ensemble de 10 éléments.

On peut illustrer les possibilités par des lignes ordonnées de 4 ronds (les clients) et 9 barres (séparant les 10 types) : compléter le tableau ci-dessous.

SCHEMA	SIGNIFICATION
oo││o│││o││││	2 boissons du type 1, 1 du type 3, 1 du type 6
o│││o││o││││o	1 boisson de chacun des types 1, 4, 6, 10
││││o│││││ooo	1 boisson du type 5, 3 boissons du type 10
	1 boisson de chacun des types 2, 3, 6, 9
│o│││o│││oo││	
	2 boissons de chacun des types 3 et 8
││││oooo│││││	
	3 boissons du type 7, 1 boisson du type 9
o│││o││││││oo	
	1 boisson de chacun des types 4, 5, 6, 10

Dans chaque schéma, il y a 13 positions numérotées. 4 sont occupées par des ronds et 9 par des barres.

Le nombre des schémas est le nombre de façons de choisir 4 positions parmi 13 (ou 9 positions parmi 13, c'est la même chose) sans tenir compte de l'ordre. Ce nombre est :

$$C_{13}^{4} = C_{13}^{9} = \binom{13}{4} = \binom{13}{9}$$

Pour passer au cas général, on remplace 10 par un entier naturel quelconque n et 4 par un entier naturel quelconque p, pas nécessairement inférieur à n.

Le nombre des p-combinaisons avec répétitions d'éléments d'un ensemble E de cardinal n est noté Γ_n^p. Il vérifie les égalités suivantes :

$$\Gamma_n^p = C_{n+p-1}^{p} = C_{n+p-1}^{n-1} = \binom{n+p-1}{p} = \binom{n+p-1}{n-1}$$

1 – Résumé

Il y a n! façons de ranger n éléments. $n! = 1\times 2\times 3\times \ldots\times (n-1)\times n$

$$A_n^p = n(n-1)(n-2)\ldots(n-p+1)$$

Il y a A_n^p façons de choisir p éléments parmi n en tenant compte de l'ordre.

$$C_n^p = A_n^p / p! = n!/ p!(n-p)! = \frac{n(n-1)(n-2)\ldots(n-p+1)}{p!}$$

Choix de p éléments discernables parmi n	avec ordre	sans ordre
avec d'éventuelles répétitions	p-listes n^p	p-combinaisons avec répétitions $\Gamma_n^p = C_{n+p-1}^{p} = \binom{n+p-1}{p}$
sans répétition	p-arrangements A_n^p	p-combinaisons $C_n^p = \binom{n}{p}$

2 – Les partitions

Une partition de E est un ensemble de parties de E 2 à 2 disjointes dont l'union est E.

On souhaite répartir les n éléments de E en p parties ($1 \le p \le n$) numérotées, 2 à 2 disjointes,

de cardinaux respectifs $n_1, n_2, \ldots, n_p$ ($n_1 + n_2 + \ldots + n_p = n$).

L'ensemble de ces répartitions a pour cardinal :

$$\frac{n!}{n_1!n_2!\ldots n_p!}$$

Pour p=2, ce cardinal n'est autre que $C_n^{n_1}$, égal à $C_n^{n_2}$.

Si ($0 \le p \le n$) alors :

$$C_n^p = \binom{n}{p} = \frac{A_n^p}{p!} = \frac{n!}{p!(n-p)!} = \frac{\overbrace{n(n-1)(n-2)\ldots(n-p+1)}^{\text{p facteurs}}}{p!}$$

En particulier : $C_0^0 = C_n^0 = C_n^n = 1 = \binom{0}{0} = \binom{n}{0} = \binom{n}{n}$

Dans tous les autres cas : $C_n^p = \binom{n}{p} = 0$

3 – La symétrie

Supposons que p et q soient 2 entiers naturels de somme n.

Dans un ensemble de n éléments, il y a autant de parties de p éléments que de parties de q éléments, puisque prendre p éléments, c'est laisser les q autres.

D'où

$$C_n^p = C_n^q \qquad (p + q = n)$$

4 – Le triangle de Pascal

Dans un ensemble E de n+1 éléments, on isole un élément a.

Le nombre de parties de E possédant p+1 éléments ($0 \leq p \leq n$) (c'est-à-dire C_{n+1}^{p+1}) est égal au nombre de parties de E de p+1 éléments ne contenant pas a (dans E-{a}, il y en a C_n^{p+1}), augmenté du nombre de parties de E formées de a et de p autres éléments (il y en a C_n^p).

D'où l'égalité

$$C_n^p + C_n^{p+1} = C_{n+1}^{p+1}$$

5 – Le binôme de Newton

On sait que : $(a + b)^2 = a^2 + 2ab + b^2$

$(a + b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$

$(a + b)^4 = a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4$

On reconnaît les coefficients du triangle de Pascal.