COURS DE STATISTIQUE INFERENTIELLE

I. Objectifs du cours

Initier l'étudiant à :

- Utiliser la logique des propositions et techniques de dénombrements dans le calcul des probabilités
- Appliquer le calcul des probabilités dans l'étude des prévisions statistiques ;
- Etudier la théorie de l'échantillonnage et d'estimation
- Utiliser les charpentes mathématiques pour tester les hypothèses de recherche ;

II. Plan du cours

Chapitre 1 : les techniques de dénombrements

1.1. Rappel sur la logique des propositions et les techniques de dénombrements

1.2 Les techniques de dénombrements

1.3 Applications (Exercices, TD ou TP)

Chapitre 2 : Notions de Probabilités

2.1 Mesures

2.2 calcul des probabilités

2.3 Applications (Exercices, TD, TP ou Interrogation)

Chapitre 3 : Les lois des probabilités et estimations

3.1 Variables aléatoires, Espérance mathématique et écart-type

3.2 Lois des probabilités (uniforme, binomiale, multinomiale, de poisson et normale

3.3 Echantillonnage

3.4 Estimation

3.5 Applications (Exercices, TP ou Interrogation)

Chapitre 4 : Tests d'hypothèses

4.1 Hypothèse nulle et Hypothèse alternative

4.2 Test unilatéral et multilatéral

4.3 Méthodes des tests d'Hypothèse

4.4 Applications (Exercices, TD ou TP)

III. COURS DETAILLE

CHAPITRE 1. LES TECHNIQUES DE DENOMBREMENTS

1.1. Rappel sur la logique des propositions

1.3 Applications (Exercices, TD ou TP)

Une proposition est une affirmation orale ou écrite. Elle peut être simple ou composée. La propriété fondamentale de toute proposition est d'être vraie, soit fausse (et de ne pas pouvoir être à la fois vraie et faux).

Notre problème est donc double :

(1) De combien de façons différentes peut on composer des propositions ?

(2) Com ment détermine-t-on la valeur logique d'une proposition composée à partir des valeurs logiques des ses composantes ?

Dans toute formule mathématique, on trouve trois sortes de symboles : des constantes, des variables et des symboles auxiliaires. Dans ce chapitre nous utiliserons des variables d'un même type que nous désignerons parles lettres p, q, r etc., qui représentent des propositions non spécifiées. Puisque chaque variable remplace une proposition, dans tous les cas une valeur logique (inconnue lui sera attachée.

Comme exemples des propositions simples, nous prenons : « il fait beau » et il « fait très chaud ». Nous désignerons la première par p et la deuxième par q. Les compositions composées issues de ces propositions simples peuvent par exemple être :

- « Il fait beau » et « il fait très chaud ». nous symboliserons cette proposition par **p**$\boldsymbol{\Lambda}$**q**. le symbole $\boldsymbol{\Lambda}$ peut se lire **« et »** ;
- « Il fait beau ou il fait très chaud ». Nous symboliserons **V** pour se lire ***« ou ».***

Nous pouvons penser que l'une des propositions ci-dessus est fausse, par **exemple « il ne fait** pas très chaud ». Le symbole employé sera ˥ p. ce qui peut se lire *« non p »*. A partir de ces trois symbole ou opérateurs logiques Λ, V et ˥ nous pouvons construire des propositions plus complexes.

Un moyen commode de résumer ce lien est l'établissement d'une table de vérité. Par exemple :

- ✓ La conjonction de p et q aura pour table de vérité :

p	p**Λ** q	p**Λ** q
V	V	V
V	F	F
F	V	F
F	F	F

- ✓ La disjonction inclusive de p et q aura pour table de vérité :

p	p**V**q	p**V** q
V	V	V
V	F	V
F	V	V
F	F	F

V signifie l'une ou l'autre ou les deux.

- ✓ La négation d'une proposition p est une proposition par ˥**p** et dont la table de vérité est :

p	˥**p**
V	F
V	V

On peut introduire une disjonction au sens strict dite exclusive et notée $\underline{V}$ ou W. elle a pour table de vérité :

p	q	p$\underline{V}$ q
V	V	F
V	F	V
F	V	V
F	V	F

Elle signifie : (p ou q, mais pas les deux).

Il existe encore plusieurs opérateurs logiques secondaires, par exemple l'implication $\Longrightarrow$, l'équivalence $\Leftrightarrow$, etc. dont cours nous ne développons pas dans le cadre de cours.

1.2 Les techniques de dénombrements

a) Ensembles et sous Ensembles

On appelle ensemble une collection d'objets .cette notion conçue dans sa généralité, a une grande a une grande importance en mathématiques, car elle peut servir de point de départ à l'étude de cette science tout entière.

Les différents meubles d'une pièce forment un ensemble. De même les livres d'une bibliothèque, les nombres entiers compris de 0 et 1000,…

Il y a deux manières essentiellement différentes de déterminer un ensemble. On peut, soit donner une règle permettant de déterminer si oui ou non un objet donné est un membre de l'ensemble.la première manière est appelée « définition » de l'ensemble, la seconde « énumération » des membres ou éléments du même ensemble.

Un ensemble constitué de quelques éléments d'un autre ensemble est appelé ***« sous-ensemble »*** de ce dernier.

Pour l'étude des sous-ensembles, nous utiliserons la terminologie suivante :

- ✓ Ensemble Universel : ensemble d'origine ;
- ✓ Ensembles élémentaires : sous-ensembles constitués d'un seul élément ;
- ✓ Ensemble vide : ensemble ou sous-ensemble qui ne comprend aucun élément.

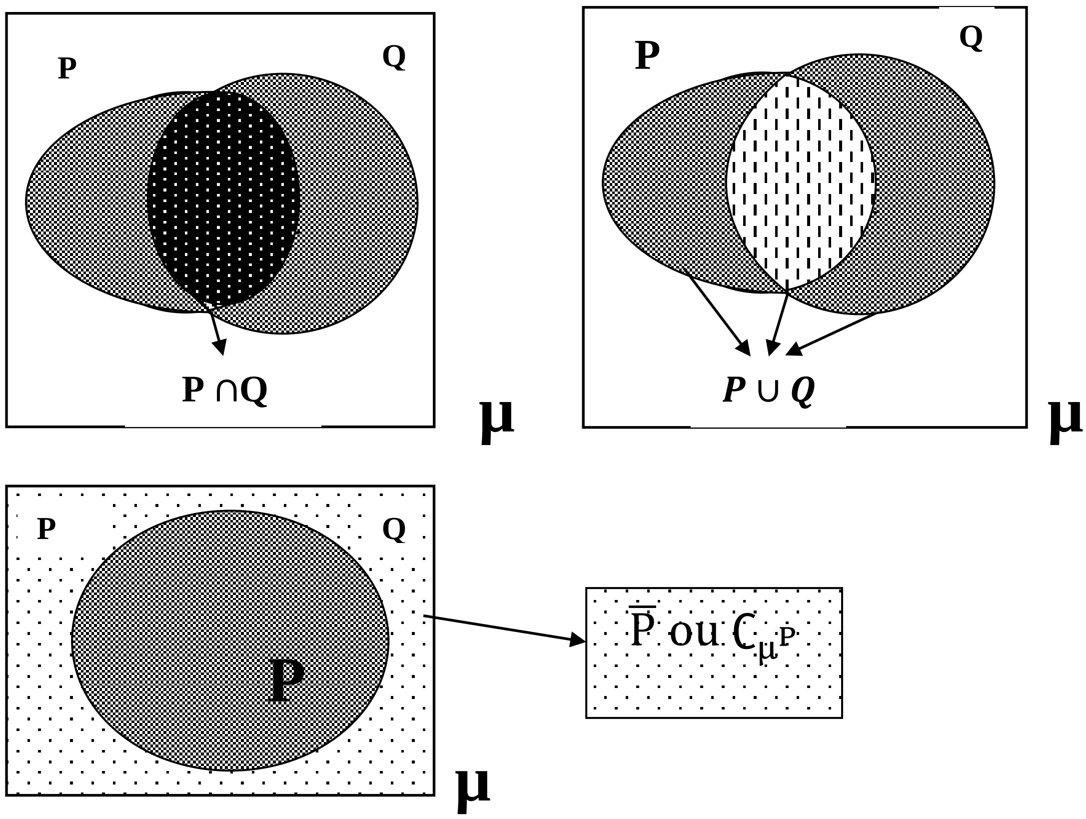
D'une manière générale un ensemble à n éléments possède 2^n sous-ensembles. L'ensemble comprend tous les sous-ensembles d'un μ n éléments et sera symbolisé par P (μ).

On peut aussi former de nouveaux ensembles à l'aide d'ensembles donnés. On utilise comme en logique (1.1.1) les opérateurs P et Q étant deux ensembles, nous appellerons :

- ✓ $P \cap Q$, l'intersection de P et Q c'est-à-dire l'ensemble qui contient uniquement les éléments communs à P et à Q ;
- ✓ $P \cup Q$, la réunion de P et Q c'est-à-dire l'ensemble qui comprend uniquement les éléments appartenant soit à P, soit à Q, soit aux deux.

- ✓ $C_P\mu$ ou $\overline{P}$ complémentaire d'un sous-ensemble P dans un grand ensemble, c'est-à-dire l'ensemble qui comprend uniquement des μ qui ne sont pas dans P.

Nous illustrerons ces opérateurs(ou opérations) par des diagrammes dits « diagrammes de Venn » Par exemple, si nous illustrons l'ensemble universel par un rectangle et ses sous-ensembles par des cercles situés à l'intérieur de ce triangle, on aura :



Deux ensembles P et Q dont l'intersection est vide sont appelés « ensembles disjoints ». Remarquons qu'en comparant la théorie des propositions et celles des ensembles on peut avoir des liens de similitude suivants :

Propositions	Ensembles
p∧q →	P∩Q
p∨q →	P∪Q
$\overline{P}$ →	$\overline{P}$

b) Partition d'un ensemble

Une partition d'un ensemble μ doit appartenir à un et un des sous-ensembles C_i **(i = 1,2,…, r)** de la partition sont appelés des « cases ».Toute partition $[C_1, C_2, \ldots C_r]$ de μ vérifie les deux conditions suivantes :

(1) $C_i \cap C_j = \emptyset$, si $i \neq j$ (les cases sont disjointes)

(2) $U_{i=1}^{r} C_i = \mu$ (les cases sont exhaustives)

Par exemple, si $\mu = \{1, 2,3\}$, les partitions de μ sont **:**

$P_1 \equiv [\{1\}, \{2\}, \{3\}]$, $P_2 \equiv [\{1\}, \{2,3\}]$

$P_3 \equiv [\{2\}, \{1,2\}]$, $P_4 \equiv [\{3\}, \{1,2\}]$

P_1 est une partition de μ en ensembles élémentaires si $[C_1, C_2, \ldots C_r]$ et $[C'_2, C'_3, C'_4]$ sont deux partitions du même ensemble μ , nous obtiendrons une nouvelle partition en considérant la collection de tous les sous-ensembles de μ de la forme $C_i \cap C'_j$ .cette nouvelle partition est appelée « partition produit » de deux partitions originales.

Par exemple à partir de l'ensemble μ de toutes les formes de vies nous pouvons former la partition P_1 où V est l'ensemble de tous les végétaux, et A celui de tous les animaux ainsi que la partition $P_2 = [D, E]$, où D est l'ensemble de formes de vie disparues, et E celui de formes de vies existantes. La partition produit :

$P_1 \cap \mathbf{P}_2 = [V \cap D, V \cap E, A \cap D, A \cap E]$

fournit une classification complète, conforme aux deux classifications séparées.

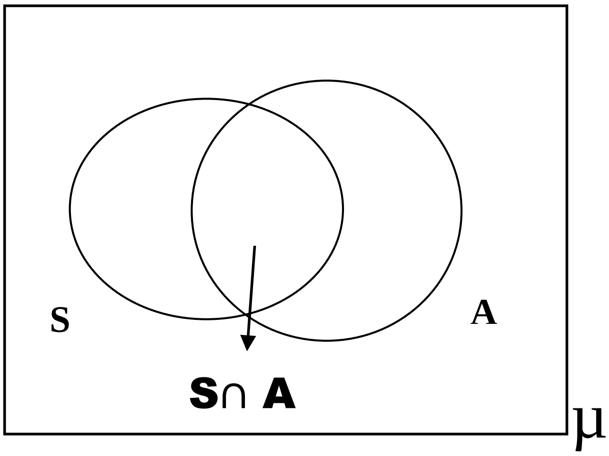
c) Nombres d'éléments d'un Ensemble

Soit χ un ensemble .désignons par $n(\chi)$ les nombres d'éléments de l'ensemble χ.

Comme exemple considérons le problème suivant :

On nous indique 60 étudiants de l'ECOPO/Lubumbashi ont échoué la statistique inductive, et que 80 ont échoué l'Anglais.

Pouvons nous dire combien ont échoué la Statistique inductive où l'Anglais ? La réponse est « **non** », parce qu'il faudrait aussi savoir combien d'étudiants ont échoué les deux matières à la fois.



Ainsi on a :

$$\boxed{n(S\cap A) = n(S) + n(A) - n(S\cap A)}$$

En généralisant cette formule aux 3,4,… sous ensembles, on aura :

(*) $n(E_1 \cup E_2 \cup E_3) = n(E_1) + n(E_2) + n(E_3) - n(E_1 \cap E_2) - n(E_1 \cap E_3) - n(E_2 \cap E_3) + n(E_1 \cap E_2 \cap E_3)$

(**) $n(E_1 \cup E_2 \cup E_3 \cup E_4) = n(E_1) + n(E_2) + n(E_3) + n(E_4) - n(E_1 \cap E_2) - n(E_1 \cap E_3) - n(E_1 \cap E_4) - n(E_2 \cap E_3) - n(E_2 \cap E_4) - n(E_3 \cap E_4) + n(E_1 \cap E_2 \cap E_3) + n(E_1 \cap E_2 \cap E_4) + n(E_1 \cap E_3 \cap E_4) + n(E_2 \cap E_3 \cap E_4) - n(E_1 \cap E_2 \cap E_3 \cap E_4)$.

(***) $n(E_1 \cup E_2 \cup E_3 \cup E_4 \cup E_5) = ?$

(*…*) $n(E_1 \cup E_2 \ldots \cup E_r) - n(E_1) + n(E_2) + \ldots n(E_r) - n(E_1 \cap E_r) - n(E_1 \cap E_3) - n(E_{r-1} \cap E_r) + n(E_1 \cap E_2 \cap E_3) + n(E_1 \cap E_2 \cap E_4) + \ldots n(E_{r-2} \cap E_{r-1} \cap E_r) - \ldots \mp n(E_1 \cap E_2 \cap \ldots E_r)$.

On appelle cette égalité la formule ***« d'inclusion-exclusion »*** ou le nombre d'éléments de la réunion d'ensembles.

Par exemple dans une école on constate que les langues vivantes choisies se repartissent comme suit :

Anglais (150 étudiants), Français (75 étudiants), Portugais (25 étudiants) et Espagnol (75 étudiants). On note aussi que les étudiants choisissent plusieurs langues : Anglais et Français (70), Anglais et Portugais(30), Anglais et Espagnol(40), Français et Portugais(5), Anglais, Français et Portugais(2). Si chaque étudiant prend au moins une langue, combien d'étudiants y-a-t-il ?

En appliquant la formule ci-dessus et en ignorant l'ensemble vide on a :

$n(A,F,P,E) = n(A) + n(F) + n(P) + n(E) - n(A \cap F) - n(A \cap F) - n(A \cap P) - n(A \cap E) - n(F \cap P) - n(F \cap P) + n(A \cap F \cap P) =$ 150+75+35+50-70-30-40-5+2=167

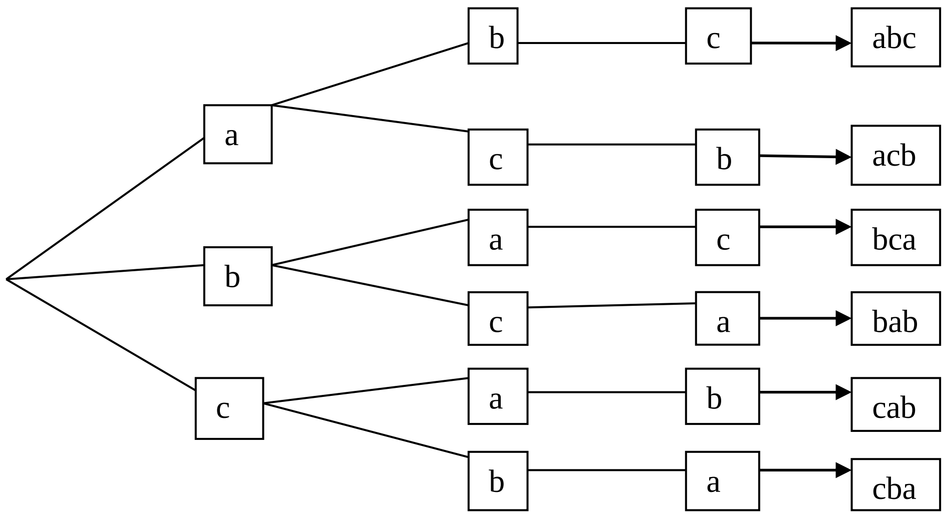
Comme chaque étudiant prend au moins une langue, le nombre d'étudiant est de 167.

d) Nombres de permutation d'éléments d'un ensemble

On appelle ***« permutation »*** de n objets différents d'un ensemble le nombre de manières dont on peut placer ces n objets dans un ordre donné. Par exemple si on veut placer trois objets a,b et c, nous pouvons énumérer ces permutations comme suit :

a b c a c b	b a c b c c	c a b c b a

Au total nous aurons 6 permutations possibles. Nous pouvons aussi obtenir les permutations en construisant un arbre semblable pour ces trois objets :



En généralisant cet arbre à n objets, nous trouverons le nombre des chemins en multipliant entre eux les nombres n, n-1, n-2, jusqu'au nombre 1.

Ce nombre obtenu est représenté par un symbole **n!** et se lit ***« n factorielle »***

Ainsi par exemple :

3 ! = 3.2.1 = 6

4 ! = 4.3.2.1 = 24, etc.

Par convention, **1! =0!=1**

Nous pouvons donc dire qu'il y a **n!** Permutations de n objets distincts, avec r types différents est :

$$\boxed{\frac{n!}{n_1\,!\,n_2!\ldots n_r!}}$$

Par exemple, combien peut faire des permutations différentes avec les lettres B, B, U, T, T, T, A ?

On aura :

$$\frac{7!}{2!\,1!\,3!\,1!} = \frac{7.6.5.4.3!}{2!\,1!\,3!\,1!} = \frac{7.6.5.4}{2.1.3.1} = 420$$

Nous représentons ce nombre des placements par le symbole :

$$C_n^{n_1, n_2, \dots, n_r} \quad \text{ou} \quad \binom{n}{n_1, n_2, \dots, n_r}$$

Par exemple une équipe a joué six matches dans sa saison.

(1) De combien de manières la saison peut-elle se terminer en deux victoires, trois défaites et un match nul ?

La réponse est $= \binom{6}{2,3,1} = \frac{6!}{2!+3!1!} = 60$

(2) De combien de façons la saison peut-elle se terminer deux victoires et 4 matches nuls ?

La réponse est $C_6^{2,3,1} = C_6^{2,4} = \frac{6!}{2!\,4!} = \frac{6.5.4!}{2!4!} = 15$ façons

e) Combinaisons de n objets pris r à r

Remarquons que le symbole ci-dessus donné du point C n'est défini que si $\mathbf{n_1+n_2+\dots + n_r = n}$

Ce symbole fournit le nombre des partitions en r cases ordonnées $[C_1,C_2,\dots C_r]$, d'un ensemble de n éléments , avec n1 dans la première case, n2 dans la seconde case, etc.

Considérons le cas particulier de deux cases .ici le problème est équivalent à celui du dénombrement des sous ensembles à r éléments que l'on peut choisir dans un ensemble à n éléments. Ceci résulte de ce que n'importe quel choix définit une partition $[C,\bar{C}]$, où **C** est l'ensemble des éléments choisis et $\bar{C}$ l'ensemble de ceux qui restent.

Les partitions de ce genre sont au nombre de :

$$\frac{n!}{r!(n-r)!} \; ;$$

Qui est aussi le nombre des sous ensembles à r éléments .dans le cas de la notation :

$\binom{n}{r, n-r}$ se réduit à $\binom{n}{r}$. Donc $\binom{n}{r} = \binom{n}{n-r} = \frac{\text{n!}}{\text{r!(n−r)!}}$ et on parle du nombre de combinaison de n éléments µ pris r à r .le symbole $\binom{n}{r}$ peut aussi s'écrire :

$$\boxed{C_n^r = \frac{\text{n!}}{\text{r!(n−r)!}}}$$

Donnez une interprétation de C_6^0 et C_6^6 et justifier la notation 0 ! =1 !

f) Généralisation de la formule de binôme

On sait que, d'après les produits remarquables étudiés au secondaire, l'on a :

- $(x+y)^2 = x^2 + 2xy + y^2$,
- $(x+y)^3 = x^3 + 3x^2y + 3xy^2 + y^3$

  On peut aussi écrire les formules comme suit :

- ✓ $(x+y)^2 = C_2^2 x^2 + C_2^1 xy + C_2^0 y^2$,
- ✓ $(x+y)^3 = C_3^3 x^3 + C_3^2 x^2 y + C_2^1 x y^2 + C_3^0 y^3$

  On peut écrire :

- ✓ $(x+y)^n = c_n^n{}_{x^n} + c_n^{n-1} x^{n-1} y + \ldots\ldots..+ c_n^1 x y^{n-1} + c_n^0 y^n \sum_{i=0}^{n} c_n^i x^{n-i} y^i$ (binôme de Newton)

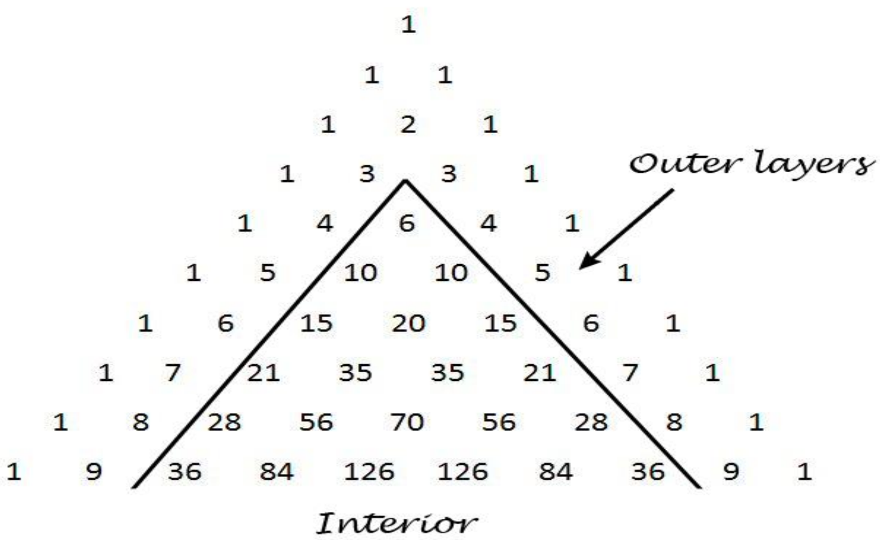
Nous obtenons ainsi la formule généralisée de $(x_{1+} x_2 + \cdots + x_r)^n$ en additionnant tous les termes de la formule

$$c_n^{n_1, n_2, \ldots\ldots\ldots n_r} {x_1}^{n_1} {x_2}^{n_2} \ldots\ldots\ldots\ldots {x_r}^{n_r} \text{, où } n_{1+} n_{2+\cdots\ldots\ldots+} n_r = n$$

Par exemple, on a :

$$(x+y+z)^2$$
$$= c_2^{2,0,0}x^2y^0z^0 + c_2^{1,1,0}x^1y^1z^0 + c_2^{1,0,1}x^1y^0z^1$$
$$+ c_2^{0,2,0}x^0y^2z^0 + c_2^{0,1,1}x^0y^1z^1 + c_2^{0,0,2}x^0y^0z^2$$
$$= x^2 + 2xy + 2xz + y^2 + 2yz + z^2$$

Notons aussi qu'une méthode commode pour obtenir les coefficients de $(x+y)^n$ est fourni par le fameux triangle dit ***«Triangle de Pascal »*** suivant :



La propriété des nombres des nombres C_n^r sur laquelle est fondée ce triangle est :

$$C_{n+1}^r = C_n^{r-1} + C_n^r$$

g) Arrangements de n objets pris r à r

Arrangement les objets différents a, b, c, d pris 2 à 2 et puis 3 à 3. On a successivement

(1)

a	db	ad	a	
a	cb	cd	b	ou $A_4^2 \ \frac{4!}{(4-2)!} = 12$
a	db	dd	c	

Donc il y a 12 façons d'arranger 4 objets 2 à 2

(2)

a	b	c	b	b	c	c	a	b	d	a	b	
a	b	d	b	b	d	c	a	d	d	a	c	
a	c	b	b	c	a	c	b	a	d	b	a	ou $A_4^3 \ \frac{4!}{(4-3)!} = 24$
a	c	d	b	c	d	c	b	d	d	b	c	
a	d	b	b	d	a	c	c	a	d	c	a	
a	d	d	b	d	c	c	c	b	d	c	b	

Donc, il y a 24 façons d'arranger 4 objets 3 à 3 .Nous pouvons ainsi arranger n objets d'un ensemble donné **μ** pris r à r en utilisant la formule :

$$A_n^r \ \frac{n!}{(n-r)!} = 24$$

1.3 Applications

Cfr. Exercices : Serie I

CHAPITRE 2 : NOTIONS DE PROBABILITES

2.1. MESURES

a) Tribu ou signe sigma-algèbre

Soit E un ensemble et $\mathcal{P}(E)$ son e ρson ensemble des parties. Un ensemble T$\subset$ $\mathcal{P}(E)$ est appelé tribu sur E si et seulement si les conditions suivantes sont vérifiées :

Par exemple, les intervalles ouverts de $\mathbb{R}$ $]-\infty$, a [,] a, b [et] b, $+\infty$ [, avec a $<$ b et a, b réels engendrent des tribus **B**($\mathbb{R}$)appelés tribus de Borel sur $\mathbb{R}$.

De la même manière, on peut définir les tribus de Borel sur $\mathbb{R}_+$,$\overline{\mathbb{R}}$= $\mathbb{R} \cup \{-\infty, \infty +\}$ et $\overline{\mathbb{R}}_+$= $\mathbb{R}_+ \cup \{-\infty, +\infty\}$ notées respectivement B($\mathbb{R}_+$), B($\overline{\mathbb{R}}$) et B($\overline{\mathbb{R}}_+$).

On peut, par exemple, montrer que B ($\overline{\mathbb{R}}_+$) est la tribu de Borel engendrée par les intervalles : I = {] a, $\infty +$ [; a $\in \mathbb{R}_+$}

b) Espace mesurable

Soit E un ensemble et T une tribu définie sur E. Le couple (E, T) est appelé ***« espace mesurable ».***

Soient (E, T) un espace mesurable et A $\in$ T .On appelle « ***fonction caractéristique mesurable*** » de l'ensemble A, notée1_A (ouχ_A) la fonction définie par :

$$1_A(\chi) = \begin{cases} 1, si\ x \in A \\ 0, si\ x \epsilon\ \bar{A} = A^C \end{cases}$$

Soit (E, T) espace mesurable et f : E$\longrightarrow \mathbb{R}$. On dit que f est « étagée » ou « T- étagée » si f est une combinaison linéaire (finie) des fonctions caractéristiques mesurables, c'est-à-dire s'il existe une famille $(A_i)_{i=1,\dots,n}$ éléments de T et n réels, … tels que $f = \sum_{i=1}^{n} a_i\, 1A_i$.

– On dit que f « étagée positive » si f est « étagée » et prend ses valeurs dans

On note l'ensemble des fonctions étagées et celui des fonctions étagées positives.

Notons que la fonction étagée positive permet de définir l'intégrale à partir de la notion de mesure.

c) Mesure

Soit (E, T) un espace mesurable. On appelle « mesure » définie sur (E, T) une application

m : $T \to \overline{\mathbb{R}}_+$, $A \to m(A)$ satisfaisant aux conditions ci-après :

(m1) m(A) = 0, si A= $\emptyset$

(m2) m (A) ≥ 0, $\forall$ A$\in$ T

(m3) m (A U B) = m (A) + m (B), si A $\cap$ B = $\emptyset$ ou

m (A U B) = m (A) + m (B) – m (A $\cap$B), sinon.

D'une manière générale,

(m3) s'écrit : m ($\cup_{i=1}^{n} A_i$) = $\sum_{i=1}^{n} m(A_i)$, si $A_i \cap A_j = \emptyset$, $i \neq j$ ou

m ($\cup_{i=1}^{n} A_i$) = $\sum_{i=1}^{n} m(A_i)$ -$\sum_{i=1}^{n} m(A_i \cap A_j)$ + $\sum_{i,j,k=1}^{n} m(A_i \cap A_j \cap A_k) - \cdots \mp m(A_1 \cap \ldots \cap A_n)$, sinon,

avec $i \neq j$, $i \neq j \neq k$, …

Le triplet (E, T, m) est appelé « espace mesuré ».

Remarquons que si f : E $\to R$ est une fonction étagée positive et m : T$\to R+$ une mesure définie sur(E, T), alors l'intégrale f s'écrit :

$$\boxed{\int f dm = \sum_{i=1}^{n} a_i \; n_i(A_i)}$$

Avec 0< $a_1 < a_2 < a_3 < \cdots < a_n$,$A_i \neq \emptyset$, pour tout i, $A_i \cap A_j = \emptyset, si\ i \neq j\ et\ f = \sum_{i=7}^{n} a_{iA_i}$.

Aussi, afin d'étendre le concept intégrale à une classe de fonctions plus générales que celle des fonctions étagées (positives), on introduit les fonctions mesurables (positives).

Soient (E, T) un espace mesurable et F un ensemble (par exemple F= $\mathbb{R}$ ou $\mathbb{R}_{+}^{-}$). Une fonction f : E$\to$F est une fonction T-mesurable si f^{-1}(A)$\in$ T ,$\forall A \in B(F)$.

Soit (E, T) un espace mesurable et f une fonction mesurable de E dans $\mathbb{R}$.Alors l'ensemble $\{f^{-1}\ (A), A \in B(\mathbb{R})\}$ est une tribu sur E appelée « engendrée par la fonction mesurable T ».

On note par :

- $\mathcal{M}(\mathrm{E}, \mathrm{T}) \equiv \mathcal{M} = \{\mathrm{f} : \mathrm{E} \to \mathbb{R}, mesurable\}$
- $\mathcal{M}_{+}(\mathrm{E}, \mathrm{T}) \equiv \mathcal{M}_{+} = \{ f : E \to \mathbb{R}_{+}^{-}, mesurable \}$

Notons que si (E, T) est un espace mesurable et E→ ℝ une fonction. Alors f est mesurable s'il existe une suite (f_n)n ∈ ℕ ⊂ $\mathcal{E}$ telle que pour tout x ∈ E ,$f_n\ (x) \to f(x)$, quand n → +∞ .

2.3 Probabilités

Nous entendons souvent des phrases connues comme celle-ci :

- Il va probablement pleuvoir aujourd'hui ;
- J'ai une grande chance de réussir cet examen ;
- Il y a une chance sur deux pour qu'une pièce de monnaie tombe côté pile, etc.

Nous constaterons que les propositions se rapprochent chaque fois à une situation dont l'issue n'est certaine, mais nous exprimerons cependant une certaine confiance quant à la justesse de notre prévision. La théorie des probabilités fournit une charpente mathématique à des telles affirmations. Dans cette partie du cours nous cherchons une méthode permettant d'attribuer des probabilités à toutes les issues imaginable pour l'expérience donnée .Nous serons aidés par un principe de base.

Principe fondamental : deux propositions équivalentes ont la mémé probabilité.

a) Définitions et propriétés

Nous procéderons en trois étapes :

(1) Nous déterminerons l'ensemble des possibilités μ,que nous appelons « univers »c'est-à-dire l'ensemble des cas possibles ;
(2) A chaque sous ensemble x de μ nous associerons un nombre appelé la mesure de x, noté m(x), mais ne dépassant pas n ;
(3) A chaque proposition p nous assignerons la mesure m(p) de son ensemble de vérité .Ainsi la possibilité d'une proposition P s'écrira Pr(p).

Par exemple on jette un dé à jouer ordinaire. Quelle est la possibilité pour qu'apparaisse un nombre inférieur à 4 ?

(1) L'ensemble des possibilités est μ= {1,2,3,4,5,6} ;

(2) Nous assignons un nombre positif 1/6 dont le poids ou la mesure à aucune possibilité ;

(3) L'ensemble des vérités de la proposition « le nombre qui sort du jet est inférieur à 4 est {1,2,3}. La probabilité Pr de cette proposition est donc $\frac{3}{6} = \frac{1}{2}$, somme des poids des éléments de l'ensemble de vérité.

La probabilité étant une mesure de valeur maximale 1, elle vérifie les propriétés suivantes :

(P_1) Pr(x) = 0 si et seulement si x=∅ ;

(P_2) 0≤ Pr(x) ≤ 1, pour tout ensemble x≤ μ ;

(P_3)Pr (x∪y) $= \begin{cases} \mathrm{Pr}(x)+\mathrm{Pr}(y), si\ x\cap y=\emptyset \\ \mathrm{Pr}(x)+\mathrm{Pr}(y)-\mathrm{Pr}(x\cap y),\ \text{sinon} \end{cases}$

ou encore dans le cas d'une proposition logique on a :

(P_1) Pr(p)=0, si et seulement si p est une erreur logique ;

(P_2) 0 ≤ Pr (p) ≤ 1, pour toute proposition p ;

(P_3) Pr (p∨q) $= \begin{cases} \mathrm{Pr}(p)+\mathrm{Pr}(q), si\ p\ et q\ sont\ compatibles \\ \mathrm{Pr}(p)+\mathrm{Pr}(q)-\mathrm{Pr}(p\wedge q),\ \text{sinon} \end{cases}$

Une autre propriété des probabilités, souvent utilisée dans l'établissement des prévisions est

Pr(μ) $= 1$ et Pr($\overline{\mathrm{X}}$) =1 - Pr(X), ou en terme de propositions,

P (μ)Pr($\overline{\mathrm{X}}$) = 1-, P(p).

Nous pouvons donc distinguer quelques terminologies relatives aux mesures ordinaires et aux mesures de probabilité suivantes :

Mesures	Probabilités
Tribu T Espace mesurable (E, T) Espace mesuré (E, T, m)	Univers μ Espace probabilisable (E, μ) Espace probabilisé (E, μ, Pr)

b) **Univers, issues, événements**

Une issue est un résultat d'une expérience aléatoire. Supposons que l'on jette un dé .Lors qu'il s'immobilisera, il indiquera l'une des six issues suivantes : 1,2, 3, 4,5 ou 6 .Les statisticiens appellent cet ensemble : $\mu = \{1,2,3,4,5,6\}$ l'univers des résultats.

On appelle événement tout sous-ensemble de μ.L'évènement est dit élémentaire s'il ne correspond qu'une seule et unique issue.

Les opérations sur les évènements sont les opérations classiques sur les ensembles. Les plus importantes sont l'intersection, l'union, et le complémentaire.

Considérons une expérience aléatoire dont l'univers est μ . Nous voulons assigner à chaque issue I un nombre réel compris de 0 à 1 :

$0 \leq Pr(I) \leq 1$. On indiquera que l'issue est **impossible** ;

1 indiquera que l'issue est **certaine**

Lorsqu'on considère lors d'une expérience des évènements élémentaires dont les issues sont $I_1, I_2, \ldots I_n$, alors $Pr(I_1) + Pr(I_2) + \ldots . Pr(I_n) = 1$

Dans ces conditions, ces évènements élémentaires sont **équiprobables**, c'est-à-dire ils sont constitués d'un seul élément et ont tous la même probabilité. Du point de vue pratique, la probabilité d'une issue est le rapport entre le nombre des cas favorables et le nombre des cas possibles.

$$\boxed{P(I) = \frac{Nombre\ de\ cas\ favorables}{Nombre\ de\ cas\ possibles}}$$

En tenant compte de la définition de la probabilité d'une issue et des opérations sur les évènements, on en déduit les axiomes du calcul des probabilités et théorèmes suivants :

Axiome 1 : $0 \leq Pr(I) \leq 1$

Axiome 2 : $Pr(\mu) = 1$

Axiome 3 : $Pr(I \cup J) = Pr(I) + Pr(J)$, si $I \cap J = \emptyset$

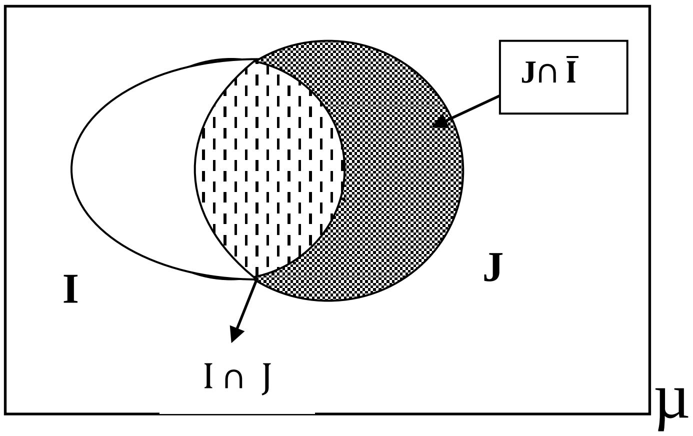
Théorème 1 : $Pr(I \cup J) = Pr(I) + Pr(J) - Pr(I \cap J)$

Théorème 2 : Pr (I^{-})= 1- Pr(I)

Théorème 3 : Pr (∅)= 0

Théorème 4 : Pr ($\bigcup_{i=1}^{n}$ I) =$\sum_{i=1}^{n} \Pr(In)$, si les évènements E_i (ou les issues I_i) sont deux à deux incompatibles.

Théorème 5 : Pr (J∩ $\bar{I}$)= Pr(J) - Pr (J∩ E).



c) Probabilité conditionnelle

Supposons que nous attendions les résultats d'une expérience aléatoire et que nous connaissions la probabilité Pr(I) de l'issue de l'évènement attendu I. Si l'expérience s'est déroulée, nous recevons une information supplémentaire, par exemple l'issue J que F s'est produit, ces renseignement va en général modifier la probabilité de réalisation de l'événement E, on parle d'une probabilité conditionnelle de E ; on parle d'une probabilité conditionnelle de E par F, notée :

Pr(E/F) = $\frac{\mathbf{Pr}(\boldsymbol{E}\cap\boldsymbol{F})}{\mathbf{Pr}(\boldsymbol{F})}$ ou Pr (I/J) = $\frac{\Pr(I\cap J)}{\Pr(J)}$

Exemple 1 : on lance deux dés .Quelle est la probabilité d'obtenir une somme supérieure à t, sachant que l'un des deux dés indique un 2 ?

Résolution : on pose : E = la somme des dés est >6

F= un de deux dés indique un 2.

Donc E∩F = la somme des deux dés est >6 et un des deux dés indique un 2.

Pr(E∩ F) = $\frac{4}{36}$ et Pr (F)= $\frac{11}{36}$

Donc , Pr(E/F) = $\frac{\Pr(E\cap F)}{\Pr(F)} = \frac{4/36}{11/36} = \frac{4}{11}$

Du point de vue logique des propositions on a :

$$\Pr[p/q] = \frac{\Pr[p \wedge q]}{\Pr[q]}$$

Par exemple, dans une élection, le candidat A a 0,4 chance d'être élu, B en a 0,3, C 0,2 et D 0,1. Juste avant l'élection, C se retire .Quelles sont maintenant les chances des candidats restants ?

Résolution : soit q la proposition que C ne sera pas élu, c'est-à-dire que A, B ou C gagneront.

Remarquons que $\Pr[q]$= 0,8, d'où il résulte que toutes les autres probabilités sont multipliées par un facteur 1/0,8 = 1,25.

On a:

(A) $\Pr[P_A/q] = \frac{Pr[P_A \wedge q]}{\Pr[q]} = 0{,}4/0.8 = 0{,}5$,Pr [$P_A \wedge$ q] = 0,5

(B): $\Pr[P_B/q] = \frac{\Pr[P_B \wedge q]}{\Pr[q]} = \frac{0{,}3}{0.8} = 0{,}375$, Pr [$P_B \wedge$ q] = 0,375

(D) : Pr $[P_D/q] = \frac{\Pr[P_D \wedge q]}{\Pr[q]} = \frac{0{,}1}{0.8}$ =0,125, Pr [$P_D \wedge$ q] = 0,125

d) Probabilités totales

Considérons une partition E_1, E_2,… E_n de l'ensemble des évènements E, c'est-à-dire .P(E) =1, $E_1 \cup E_2$ … E_n= E, $A_i \cap A_j$ =Φ pour $i \neq j$.

Alors la possibilité d'un évènement $B < E$ vaut :

Pr(B)= Pr(E_1).Pr(B) E_1 + … + Pr(E_n) .Pr

Car Pr(B) = Pr(B∩ E)

= Pr (B∩ E_1+ Pr(B∩ E_2) +… + Pr(B∩ E_n)

= Pr (E_1).Pr (E_1) +… Pr(E_n).Pr (B/E_n)

e) procédure stochastiques finis

Nous appelons processus stochastique une succession d'expériences dont les résultats ils impliquent une certaine part du hasard. Nous envisageons un nombre fini d'issues possibles.

Considérons une succession de trois expériences représentées par la figure suivante :

L'ensemble de tous les résultats possibles pour chacune des expériences est $\{a, b, c, d, e, f\}$.

Remarquons que si la première donne le résultat a, la troisième aura pour résultat $\{c, f\}$

Si la première a pour résultat b, la seconde aura pour issues possibles l'ensemble$\{a, e, d\}$.

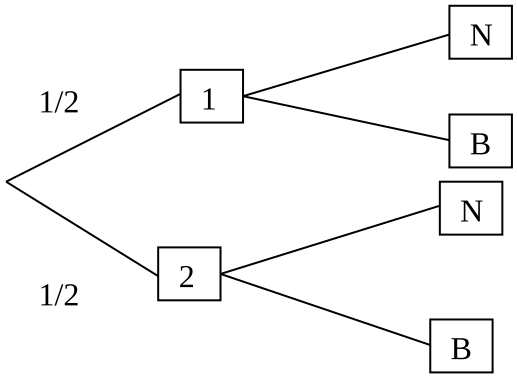
Nous représentons par Pr_a la possibilité que la première expérience donne le résultat a. Nous représentons par Pr_{bdc} la possibilité que d soit le résultat de la seconde expérience sachant que la première a donné le résultat b. Nous représentons par Pr_{bdc} la probabilité que c soit le résultat de la troisième expérience sachant que la première a donné le résultat b et la deuxième d. De plus, nous avons :

$$Pr_a + Pr_b = Pr_{ba} + Pr_{bc} + Pr_{bd} = Pr_{bda} + Pr_{bdc} = 1$$

Pour la succession des résultats b et d par exemple, nous avons

$Pr_b\, Pr_{bd} Pr_{bdc}$+$Pr_b\, Pr_{bd}$ +Pr_{bdc}=$Pr_b\, Pr_{bd}[Pr_{bda}$+$Pr_{bdc}]$ = $Pr_b\, Pr_{bd}$

Considérons l'exemple de deux urnes, dont la première contient deux boules noires et trois boules blanches ; on choisit une urne au hasard, puis au hasard encore, on en tire une boule. Quelle est la probabilité pour que ce soit une boule blanche ?



Le choix de la boule blanche nous conduit à la somme des trajets :

$$Pr_1 Pr_{1B} + Pr_2 Pr_{2B} = \frac{1}{2} * \frac{2}{5} + \frac{1}{2} * \frac{1}{3} = \frac{1}{2}\left(\frac{3}{5} + \frac{1}{3}\right) = \frac{1}{2} * \frac{9+5}{15} = \frac{1}{2} * \frac{14}{15} = \frac{7}{15}$$

f) Théorème de Bayes

- Considérons une partition $A_1, A_2, \ldots A_n$ de l'ensemble des événements μ .soit B un événement quelconque on a : A_i

$$\Pr(A_i/B) = \frac{(A_i \cap B)}{\sum_{j=1}^{n} \Pr(A_j \cap B)} = \frac{Pr(A_i \cap B)}{Pr(A_1 \cap B) + Pr(A_2 \cap B) + \cdots Pr(A_n \cap B)}$$

Notons que pour démontrer ce théorème, on se sert des résultats des possibilités conditionnelles et totales :

- $\Pr(A_i \cap B) = \Pr(B).\Pr(B/A_i)$
- $\Pr(B) = \Pr(A_1 \cap B) + (A_2 \cap B) + \ldots \Pr(A_n \cap B) = \Pr(A_1).\Pr(B/A_1) + \ldots + \Pr(A_n).\Pr(B/A_n)$.

En d'autres termes,

$$\Pr(A_i/B) = \frac{Pr(A_i).Pr(B/A_i)}{Pr(A_1) + Pr(B/A_1) + \cdots Pr(A_n).\Pr(B/A_n)}$$

- Dans la logique des propositions on peut avoir un ensemble complet $\{p_1, p_2, \ldots p_n\}$des propositions de μ et une proposition quelconque q de **.** on a :

  $$P[P_i/q] = \frac{Pr[p_i \wedge q]}{[Pr[p_1 \wedge q] + Pr[p_2 \wedge q] + \cdots Pr[p_n \wedge q]} = \frac{Pr[p_i].Pr[q/p_i]}{Pr[p_1].Pr[q/p_1] + \cdots + Pr[p_n].\mathrm{pr}[q/p_n]}$$

  32

Par exemple, dans un laboratoire, on a fait les constats suivants :

- ✓ Si une souris porte l'anticorps A, alors 2 fois sur 5 elle porte l'anticorps B ;
- ✓ Si une souris ne porte pas l'anticorps A, alors 4 fois sur 5 elle ne porte pas l'anticorps B.

Calculer la probabilité que, si une souris porte l'anticorps B, alors elle porte aussi l'anticorps A.

Pour la partition $\{A, \overline{A}\}$, la formule de Bayes donne :

$\Pr(B) = \Pr(A).\Pr(B/A) + \Pr(\overline{A}.)\Pr(B/\overline{A})$

$$= \frac{1}{2} * \frac{2}{5} + \frac{1}{2} * \frac{1}{5} = \frac{3}{10}$$

D'où,

$$\Pr(A_i/B) = \frac{\Pr(A).\Pr(B/A)}{\Pr(B)} = \frac{\frac{1}{2} * \frac{2}{5}}{\frac{3}{10}} = \frac{2}{3}$$

$$\Pr(A_i/B) = \frac{\Pr(A).\Pr(B/A)}{\Pr(B)} = \frac{\frac{1}{2} * \frac{1}{5}}{\frac{3}{10}} = \frac{1}{3} \ ;$$

Interprétation

Probabilité à priori : $\Pr(A) = \frac{2}{3}$, $\Pr(\bar{A}) = \frac{1}{2}$;

Apports d'informations : $\Pr(B/A) = \frac{2}{5}$, $\Pr(B/\bar{A}) = = \frac{1}{5}$,

Probabilité à priori pour les porteuses de l'anticorps de l'anticorps B :

$\Pr(A/B) = \frac{2}{3}$ et $\Pr(\bar{A} \ /B) = = \frac{1}{3}$.

Probabilité des causes

Les 2/3 des souris porteuses de l'anticorps B portent aussi l'anticorps A, ce qui dénote une nette incidence de A sur B.

Considérons l'exemple suivant :

Un étudiant doit choisir l'un des cours suivants : mathématiques, physique, chimie et botanique. En se fondant sur l'intérêt qu'il leur a manifesté son professeur pense les matières ont respectivement les probabilités 0.4, 0.3, 0.2 et 0.1 d'être choisies.

Le professeur ne sait pas quel a été le choix effectué, mais il sait que l'étudiant a reçu la note A.

Il estime d'autre pat que les probabilités d'avoir un A sont 0.1 pour les mathématiques, 0.2 pour la physique, 0.3 pour la chimie, 0.9 pour la botanique.

Comment le professeur peut-il réviser des estimations initiales des probabilités des différents choix ?

En utilisant le théorème de Bayes on obtient :

Pr (Il ait choisi les maths/il a obtenu la note A)

$$= \frac{(0.4)*(0.1)}{(0.4)*(0.1)+(0.3)*(0.2)+(0.2)(0.3)+(0.1)*(0.9)} = \frac{4}{25}$$

NB : On ne peut pas appliquer le théorème de Bayes sans connaitre les probabilités initiales des propositions.

Si Pr [P_1]= Pr [P_2] = Pr [P_3] = Pr [P_4] , alors

$$\text{Pr}\,[P_i /q] = \frac{Pr[q/p_i]}{Pr[q/p_1]+Pr[q/p_2]+Pr[q/p_3]+ Pr[q/p_4]}\ ,$$

CHAPITRE 3 : LES LOIS DE PROBABILITES

3.1 VARIABLES ALEATOIRES

Nous définissons une variable aléatoire comme une fonction réelle des éléments d'un espace E . Nous désignerons par une lettre majuscule (X, Y, W,…). Chaque valeur particulière d'une valeur aléatoire X est désignée par une lettre minuscule (par exemple x). En d'autres termes,

X : E→ $\mathbb{R}$,x→ $X(x)$.

A tout élément se résultant d'une variable aléatoire on peut associer une probabilité.

Pr(X= x) = p(x),

satisfaisant

Pr(x= $-\infty$) = 0 = Pr(x= $+\infty$).

Une variable alternative peut être discrète ***continue*** ou ***mixte***. Notons (qu'une variable aléatoire X est mixte lorsque certaines de ces valeurs sont discrètes et d'autres continues. A partir d'un même espace E on peut aussi définir 2,3,…,n variables aléatoires $x_1, x_2, \ldots x_n$ dont la probabilité associée est telle que :

$Pr(X_1= x_1, X_2= x_2 ,\ldots,X_n=x_n) = p(x_1,x_2,\ldots,x_n)$,

avec $Pr(x_1 = -\infty, \ldots, x_n= -\infty) = 0 = Pr(x_1 = +\infty, \ldots, x_n= +\infty)$.

Dans une expérience aléatoire, une variable x est dite « variable aléatoire » lorsqu'on peut associer aux termes x_1, $x_1 \ldots$, xn qui représente les probabilités $Pr_1, Pr_2 , \ldots , Pr_n$ .Dans ce cas, on a :

X	X_1	X_2	…	X_n
Pr	Pr_1	Pr_2	…	Pr_n

Avec comme conditions :

$\sum_i Pr_i = Pr_1 + Pr_2 + \ldots Pr_n = 1$

Une variable aléatoire peut être discrète ou continue. Dans le cas discret, on note :

$Pr(x=x_i) = Pr(x=\lambda_1 , x=\lambda_2, \ldots Pr(x=\lambda_n)$

= [Pr(x=λ_1) ,Pr(x=λ_2), … Pr(x=λ_n)]

= (Pr_1+ Pr_2 + … Pr_n)

Tandis que dans le cas continu, on a :

Pr(x=λ)=p(λ),$-\infty < x < +\infty$,

Avec$\int_{-\infty}^{+\infty} p(\lambda)$ =1

La condition $\int_{-\infty}^{+\infty} p(\lambda) = 1$

On peut aussi parler des expériences à plusieurs variables aléatoires discrètes continues.

Une expérience à deux variables aléatoires discrètes X et Y peut se représenter comme suit :

x/y	**y_1**	**y_2**	**…. y_n**
x_1	**(x_1,y_1)**	**(x_1,y_2)**	**(x_1,y_n)**
x_2	**(x_2,y_1)**	**(x_2,y_2)**	**(x_2,y_1)**
.	.	.	.
.	.	.	.
.	.	.	.
.	.	.	.
.	.	.	.
x_n	**(x_n,y_1)**	**(x_1,y_1)**	**(x_n,y_n)**

Et à chaque couple (x_i,y_j)avec i et j=1,…,n on peut associer une probabilité

Pr (X=x_i, Y=y_j) = P_{ij} ,

Avec comme condition :

$$\sum_{i,j} P_{ij}=1.$$

Dans le cas de deux variables aléatoires continues X et Y

on peut écrire :

Pr(X=x ,Y=y) = P(x,y),

Avec comme condition :

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} p(x,y)dxdy = 1.$$

Donc p(x,y) est une densité de probabilité.

On peut donc procéder de la même manière dans le cas d'une expérience à 3, 4,5,…,n variables aléatoires discrètes ou continues.

Lorsqu'on considère trois variables aléatoires continues X,Y et Z la densité de probabilité peut s'écrire :

P(X=x,Y=y,Z=z) = p(x,y,z), avec comme condition

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} p(x,y,z)dxdydz = 1$$

3.2 FONCTION DE REPARTITION-ESPERANCE MATH-VARIANCE

a. Fonction de répartition

Nous appelons fonction de répartition d'une variable aléatoire X, notée F_x , une fonction des probabilités cumulées définie par :

F_X :E→[0,1],x→F_X(x) = Pr(X $\leq x$).

Elle a pour propriétés :

(1) $F_x(-\infty) = 0$

(2) $F_x(+\infty) = 1$

(3) $0 \leq F_x(x) \leq 1$

(4) $F_x(x_1) \leq F_x(x_2)$ si $x_1 < x_2$

(5) Pr $(x_1 < X < x_2) = F_x(x_2) = F_x(x_1)$

(6) $F_x(x^+) = F_x(x)$, avec $x^+ = x+\varepsilon$ si $\varepsilon \rightarrow 0$.

Elle est donc définie par

$F_X(x) = \Pr(X \leq x), \forall x \in \mathbb{R}$

- ✓ **S**oit une expérience aléatoire de variable aléatoire x dont les résultats possibles sont des nombres $x_1,x_2,\ldots x_n$ dont les probabilités respectives sont : $\Pr(X=x_1) = P_1,\ldots,\Pr(X=x_n) = P_n$.

b. Esperance mathematique

On appelle espérance mathématique de la variable aléatoire (ou valeur du résultat) la valeur moyenne du résultat :

$E(X) = \sum_{i=1}^{n} x_i\, p_i = x_1 p_1 + \cdots + x_n p_n$

Remarquons que l'espérance mathématique de gain est souvent considérée comme la valeur du jeu pour le joueur.

$$\text{Si } E(x) = \begin{cases} > 0, si\ le\ jeu\ est\ favorable \\ = 0, \text{si le jeu est équitable} \\ < 0, si\ le\ jeu\ est\ défavorable \end{cases}$$

Par exemple quelle est l'espérance mathématique du nombre de succès dans le cas de quatre expériences indépendantes ayant la probabilité 1/3 de réussir ?

Résolution

$E(X) = \sum_{i=0}^{4} C_4^i \left(\frac{1}{3}\right)^i \left(1 - \frac{1}{3}\right)^{4-i} = \frac{4}{3}$

c.Variance

- ✓ On appelle variance de la variable aléatoire X n, la différence entre la moyenne des carrés et le carré de la moyenne, c'est-à-dire :

$V(X) = E(X^2) – (E(x))^2$

$= (\sum_{i=0}^{n} x_i^2\, p_i) – (\sum_{i=1}^{n} x_i\, p_i)^2$

d. Ecart-type

- ✓ *On appelle écart-type de la variable aléatoire X la racine carrée de la variance, c'est-à-dire :*

$\sigma_X = \sqrt{V(x)}$

Il sied de remarquer qu'en cas de variables continues, si

(1) $E(X) = \int_{-\infty}^{+\infty} xp(x)dx$

(2) $V(X) = \int_{-\infty}^{+\infty} x^2 p(x)dx – (\int_{-\infty}^{+\infty} xp(x)dx)^2$

alors on a :

$$\sigma_X \sqrt{\int_{-\infty}^{+\infty} x^2 p(x)dx - (\int_{-\infty}^{+\infty} xp(x)dx)^2}$$

De la même manière, on peut définir l'espérance mathématique, la variance te l'écart-type dans le cas de 2, 3,4,… variables aléatoires et des probabilités conditionnelles.

Dans le cas de n variables aléatoires $X_1,\ldots,X_n$ on a :

$(x_1,\ldots, x_n) = \Pr (X_1 \leq x_1,\ldots,X_n \leq x_1)$

$F_{x_1,\ldots,x_n}$.

Avec comme propriétés :

(1) $F_{x_1,\ldots,x_n}(-\infty,\ldots,+\infty\ )=0$

$F_{x_1,\ldots,x_n}(-\infty,x_2,\ldots,x_n\ )=0$

$F_{x_1,\ldots,x_n}(x_1,\ldots,x_{n-1},-\infty\ )=0$

(2) $F_{x_1,\ldots,x_n}$

(3) $0 \leq F_{x_1,\ldots,x_n} \leq 1$

(4) $F_{x_1,\ldots,x_n}(x_1,\ldots,x_n\ )$ est une fonction non décroissante pour toutes les variables $x_1,.. x_n$.

(5) $F_{x,y}\ (x_2, y_2) + F_{x,y}\ (x_1, y_1) - F_{x,y}\ (x_1, y_2) - F_{x,y}\ (x_2, y_1) = \Pr(x_1 < x \leq x_2, y_1 < y \leq y_2\ ) \geq 0$

(6) $F_{x,y}\ (x,+\infty) = F_{x_1}(x)$ et $F_{x,y}\ (+\infty,y) = F_y\ (y)$

Fonction densité de probabilité

La fonction densité de probabilité, notée $f_x(x)$, est définie comme étant la fonction dérivée de la fonction de répartition, c'est-à-dire :

$$f_x(x) = \frac{dF_x(x)}{dx}$$

Ses propriétés sont :

(1) $0 \leq f_x(x)$, pour tout x,

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

(3) $F_X(x) = \int_{-\infty}^{x} f_x\ (\varepsilon)d\varepsilon$

(4) $\Pr\ (x_1 < x \leq x_2) = \int_{x_1}^{x_2} f_x\ (x)dx\ \partial$

Dans le cas de n variables $x_1,\ldots, x_n$ une telle fonction est définie par :

$F_{x_1,\ldots,x_n}(x_1,\ldots,x_n) = \frac{\partial^n F_{x_1,\ldots,x_n}(x_1,\ldots,x_n)}{\partial x_1 \partial x_n,\ldots,\partial x_n}$ et ses propriétés sont :

(1) $0 \leq F_{x_1,\ldots,x_n}(x_1,\ldots,x_n)$

(2) $\int_{-\infty}^{+\infty} \ldots \int_{-\infty}^{+\infty} f_{x_1,\ldots x_n}(x_1,\ldots x_n) d_{x_1} \ldots d_{x_n} = 1$

(3) $F_{x_1,\ldots,x_n}(x_1,\ldots,x_n) = \int_{-\infty}^{x_1} \ldots \int_{-\infty}^{x_n} f_{x_1,\ldots x_n}(\varepsilon_1,\ldots,\varepsilon_n) d\varepsilon_1 \ldots d\varepsilon_n$

(4) $\Pr(x_{11} < x_1 \leq x_{12}, \ldots x_{n1} < x_n \leq x_{n2})$

$= \int_{x_{11}}^{x_{12}} \ldots \int_{x_{n1}}^{x_{n2}} f_{x_1,\ldots x_n}(x_1,\ldots x_n) d_{x_1} \ldots d_{x_n}$

e.Chaines de Markov

On réalise une succession d'expériences ayant les propriétés suivantes :

- ✓ Chaque expérience possède un nombre fini d'issues possibles $a_1, a_2, \ldots, a_r$, dont l'une a_j se réalise ;
- ✓ La propriété de cette issue particulière n'est pas nécessairement dépendante des issues des expériences antérieures, mais dépend surtout du résultat de l'expérience qui la précède immédiatement ;
- ✓ Il existe des nombres p_{ij} représentant la probabilité qu'une expérience aboutisse en a_j sachant que la précédente a aboutit en a_i .

Les issues $a_1, a_2, \ldots, a_r$ sont appelées ***les états du processus*** et les nombres P_{ij} sont appelés ***probabilités de transition*** . Un tel processus est appelé ***une chaine de Markov***.

Les probabilités de transition sont représentées par la matrice :

$$P = \begin{pmatrix} p_{11} & \cdots & p_{1r} \\ \vdots & \ddots & \vdots \\ p_{r1} & \cdots & p_{r1} \end{pmatrix}$$

f. probabilités conditionnelles

On sait que :

$\Pr(A/B) = \frac{\Pr(A \cap B)}{\Pr(B)}$

Si A est identifié comme un évènement $\{X \leq x\}$ pour une variable aléatoire X, la fonction de répartition de probabilité conditionnelle est définie par :

$F_X(x/B) = \Pr(X \leq x/B) = \frac{\Pr(X \leq x \cap B)}{\Pr(B)}$,

Avec $\{X \leq x \cap B\} = \{X \leq x\} \cap B$;

Ses propriétés sont :

(1) $F_x(-\infty/B) = 0$
(2) $F_x(+\infty/B) = 1$
(3) $0 \leq F_x(x/B) \leq 1$
(4) $F_x(x_1/B) \leq F_x(x_2/B)$ si $x_1 < x_2$
(5) $\Pr(x_1 < X \leq x_2/B) = F_x(x_2/B) - F_x(x_1/B)$
(6) $F_x(x^+/B) = F_x(x/B)$

S'agissant de la fonction densité de probabilité conditionnelle résultant de $Fx(x\backslash B)$ on a :

$f_x(x \backslash B) = \frac{dF_x(x\backslash B)}{dx}$ avec

(1) $f_x(x\backslash B) \geq 0$
(2) $\int_{-\infty}^{+\infty} f_x(x\backslash B)dx = 1$
(3) $F_x(x\backslash B) = \int_{-\infty}^{+\infty} f_x(\varepsilon \backslash B)d\varepsilon$
(4) $\Pr(x_1 < X \leq x_2/B) - \int_{x_1}^{x_2} f_x(x \backslash B)dx$

3. ECHANTILLONNAGE-ESTIMATION

3.1. Echantillonnage

Soit $x_1, x_2, \dots, x_n$ des variables aléatoires indépendantes et de même loi. On dit que $(x_1, x_2, \dots, x_n)$ est un échantillon de taille n ou une n-échantillon de la variable X.

$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ est appelé moyenne de l'échantillon ou moyenne empirique.

L'échantillon doit être :

- Représentatif, c'est-à-dire que nnnnnnnnnnnnnnn
- Non exhaustif, c'est-à-dire que nnnnnnnnnnLa ta
- La taille N de la population doit etre suffisamment importante devant celle de l'echantillon pour se ramener au prelevement au hasard.

3.2. Distribution

Supposons que $x_1, x_2, \dots, x_n$ représentent les résultats obtenus à partir d'échantillon aléatoire de taille n et d'une population normalement distribué de moyenne μ et d'écart-type σ.

On sait déjà que la distribution des moyennes échantillonnales est normalement distribuée et l'écart-type de cette distribution est donné par :

<u>1er cas</u> :*l'écart-type σ de la population est connue etla taille N de la population est infinie*

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

<u>2ème cas</u> : *l'écart-type σ de la population est connue etla taille N de la population est finie*

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$$

Par conséquent, la variable aléatoire

$$\hat{z} = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}}$$

suit une distribution normale centrée réduite ; donc cette variable est une cote z.

<u>3ème cas</u> : *l'écart-type de la population est inconnue etla taille N de la population est infinie*

Il faut le determiner a l'aide de l'ecart-type s de l'echantillon et obtenir :

$$\sigma_{\bar{x}} = \frac{s}{\sqrt{n}}$$

<u>4ème cas</u> : *l'écart-type de la population est inconnue etla taille N de la population est finie*

il faut le déterminer a l'aide de l'écart-type s de l'échantillon et obtenir :

$\sigma_{\bar{x}} = \frac{s}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$,

avec$s = \sqrt{\frac{\sum_i (x_i - \bar{x})^2}{n-1}}$

Par conséquent, la variable aléatoire

$$t = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}}$$

ne se distribue pas normalement, mais selon une distribution particulière appelée t de Student.

3. 3. Estimation

3.3.1.Estimation ponctuelle

L'estimation d'un paramètre τ est une fonction des observations (aléatoire, par conséquent) qui est une évaluationdeτ. Il est :

- **Sans biais** si sa moyenne est égale aτ ;
- **Convergent** (ou **constant**) s'il tend vers τ quand le nombre d'observations tend vers l'infini.

Si le paramètre τ est la moyenne ou la variance d'une variable aléatoire X, alors on a des estimateurs très simples : la moyenneempirique (observée) et la variance empirique.

En particulier, l'estimateur d'une proposition$p$, qui est la moyenne d'une variable de Bernoulli $\mathcal{B}(p)$, est la proposition observéenotéep_o.

On peut de la même manière procéder à l'estimateur sans biais de σ^2 (***à travailler comme T.P. en groupes***).

Dans les autres cas, on utilise une méthodetrèsgénéraleappelée le **maximum de vraisemblance** (***à travailler comme T.P. en groupes***).

Onécrit donc la probabilité des observations comme fonction du (ou des) paramètres τ et on estime τ par la (ou les) valeur (s) maximum de cette probabilité.

3.3.2.Estimation par intervalle de confiance

La confiance est la probabilité avec laquelle l'intervalle couvre la vraie valeur du paramètre.on veut que cette probabilité soit proche de 1. On la note 1 –α, avec α petit. En général, α est de l'ordre de 0,05 ou moindre.

Pour un proposition p : elle est obtenue en prenant la proposition observée p_o plus ou moins un terme qui dépend de la confiance 1 – α que l'on veut pouvoir accorder à l'intervalle.

$$\hat{p} \in \left]p_o - p_o q_o n z_{1-\frac{\alpha}{2}} ; p_o + p_o q_o n z_{1-\frac{\alpha}{2}}\right[,$$

avec$\hat{p}$ désignant l'estimateur de p.

Dans cette expression, $z_{1-\alpha}$est le 1 – α quantile de la loi normale $\mathcal{N}(0;1)$:

Pr(Z ≤ $z_{1-\alpha}$) = 1 – α

La confiance est donc la probabilité avec laquelle l'intervalle couvre la vraie valeur de p.

De même pour la moyenneμ de la population, l'intervalle de confiance est donnée par:

$\mu \in]\bar{x} - \mathcal{Z}\sigma_{\bar{x}} ; \bar{x} + \mathcal{Z}\sigma_{\bar{x}}[$.

Avec $\bar{x}$ la moyenne de l'échantillon, $\mathcal{Z}$ la valeur de z qui correspond au niveau de confiance désiré et $\sigma_{\bar{x}}$l'erreur- type de la moyenne.

- $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$, si*l'écart-type σ de la population est connue etla taille N de la population est infinie ;*
- $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$ *, si l'écart-type σ de la population est connue etla taille N de la population est finie ;*
- $\sigma_{\bar{x}} = \frac{s}{\sqrt{n}}$, si *l'écart-type de la population est inconnue etla taille N de la population est infinie ;* il faut l'estimerà l'aide de l'ecart-type s de l'echantillon ;
- $\sigma_{\bar{x}} = \frac{s}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$, si *l'écart-type de la population est inconnue etla taille N de la population est finie ;* il faut le déterminer a l'aide de l'écart-type s de l'échantillon.

Pour un niveau de confiance de 99 % , Z vaut 2,58 ; pour un niveau de confiance de 95 % , Z vaut 1,96, et pour un niveau de confiance de 90 % , Z vaut 1,64.

Par exemple, supposons qu'on obtienne, pour un echantillon de cinq etudiants de Licence 2 EcoPo- Lubumbashi, les quotients intellectuels (QI) suivants :

75, 88, 112, 85, 80

Et on veut construire un intervalle avec un niveau de confiance de 95 % . on suppose que $\sigma = 15$ et on sait que $\bar{x} = \frac{75+88+112+85+80}{5}$ = 88. Puisque N est inconnue et n= 5, on a :

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{15}{\sqrt{5}} = \frac{15}{2,2361} = 6,71$$

Or, pour un niveau de confiance(NC) de 95 %, la valeur de Z vaut 1,64. Donc

$$\mu \in \left]\bar{x} - Z\sigma_{\bar{x}} ; \bar{x} + Z\sigma_{\bar{x}}\right[= \left]74,85;\ 101,15\right[$$

En particulier, dans le cas d'usage du de Student, on :

$\mu \in \left]\bar{x} - t_{\alpha/2}\sigma_{\bar{x}};\ \bar{x} + t_{\alpha/2}\sigma_{\bar{x}}\right[$,

avec$t_{\alpha/2}$ la valeur de t qui correspond a la probabilité $\alpha/2$ et $\sigma_{\bar{x}}$ l'estimé de erreur-type de la moyenne.

Notons aussi que la valeur de t associee a un niveau de confiance donné se lit dans la table des t a (n -1) degré de liberté , en abrégé :

dl = n – 1.

Donc $\frac{\alpha}{2} = \frac{1-NC}{2}$.

3.4.Contrôle de fabrication

On appelle qualité d'un lot la loi de probabiliténotéeω, la proportion des piecesdefectueuses qu' il contient.

Pr$(\omega) = \sum_{k=0}^{c} C_n^k \, \omega^k (1-\omega)^{n-k}$

$$\alpha = \sum_{k=0}^{c} C_n^k \, \omega_o^k 1 - \omega_o{}^{n-k}$$

$$1 - \beta = \sum_{k=0}^{c} C_n^k \, \omega^k (1-\omega)^{n-k}$$

3.5. Applications

4. TESTS D'HYPOTHESES

4.1 Types d'hypothèses

Les tests d'hypothèses constituent un autre aspect important de l'inférence. En statistiques, un test est une procédure de décision entre deux hypothèses. Il s'agit d'une démarche consistant à rejeter ou ne pas rejeter une hypothèses statistique, appelée hypothèse nulle, en fonction d'un jeu de données(échantillon).

Il s'agit d'une statistique inférentielle :à partir de calculs réalisésur des donnéesobservées, nous émettons des conclusions sur la population, en leur rattachant des risques de se tromper.

L'hypothèse nulle notéeH_o, est celle que l'on considère vraie a priori. Le but du test est de decider si cette hypothèse a priori est crédible.

L'hypothese alternative, noteeH_1est l'hypothèsecomplémentaire a l'hpothesenulle.

Ces deux hypothèses ne sont pas symétriques. H_1est choisie uniquement par défaut si H_o n'est pas crédible. Le choix de H_o et H_1 est imposé par le test qu'on utilise et ne relève donc pas de l'utilisateur.

4.2 Ecriture deshypothèses

Les tests peuvent être : unilatéral inferieur, unilatéral supérieur ou multilatéral.

a) Test unilatéral inferieur

Pour une loi $L_o$complétementspécifiée par exemple, les hypothèses d'un test unilatéral inferieur peuvent s'écrire :

$H_o \equiv L \geq L_o$et$H_1 \equiv L < L_o$

b) Test unilatéralsupérieur

Pour une loi $L_o$complétementspécifiée par exemple, les hypothèsesd'un test unilatéralsupérieurpeuvent s'ecrire :

$H_o \equiv L \leq L_o$et$H_1 \equiv L > L_o$

c) Test multilatéral

Pour une loi $L_o$complètementspécifiée par exemple, les hypothèsesd'un test multilatéralpeuvent s'écrire :

$H_o \equiv L = L_o$et$H_1 \equiv L \neq L_o$

Apres l'émission des hypothèses nulle et alternative, l'on doit définir la " statistique de test ".

4.3 Méthodes de tests d'hypothèses

4.4 Applications