MERABET HAYET

INTRODUCTION AU CALCUL DE PROBABILITES ET STATISTIQUE INFERENTIELLE

AVANT-PROPOS

Ce polycopié est un support de cours destiné aux étudiants de statistique appliquée et informatique. Les quatre chapitres sont structurés en deux parties : l'essentiel de la théorie des probabilités et la statistique ʻʻ classiqueʼʼ avec l'estimation et les tests. L'objectif principal est d'initier l'étudiant aux probabilités et à la statistique inférentielle.

La théorie des probabilités fournit des modèles mathématiques permettant l'étude d'expériences dont le résultat ne peut être prévu avec une totale certitude.

En statistique inférentielle, on cherche à connaître l'inconnu. Pour ce faire, on émet des hypothèses ou on estime des grandeurs partiellement inconnues,

- On se fixe des probabilités *a priori* de se tromper dans ses estimations ou son test;
- On cherche (difficile) pour quels choix des paramètres (intervalle ou estimateur) ces probabilités sont atteintes;
- On prend ses décisions à partir des choix faits ci-dessus.
- La démarche statistique n'est pas seulement une auxiliaire des sciences destinée à valider ou non des modèles préétablis, c'est aussi une méthodologie indispensable pour extraire des connaissances à partir de données et un élément essentiel pour la prise de décision.

Table des matières

CHAPITRE IV: TESTS D'HYPOTHESES

IV. 1 - Concepts dans l'élaboration d'un test d'hypothèse

IV. 2 - les tests paramétriques

IV. 2. 1 - Test d'une moyenne

IV. 2. 2 - Test d'une proportion

IV. 2. 3 - Test d'une variance

IV.3 - Les tests non paramétriques

IV. 3.1 - Test du Khi- 2

IV.3.2 - Test d'indépendance de χ^2

IV.3.3 - Test Kolmogorov-Smirnov

INTRODUCTION AUX PROBABILITES ET STATISTIQUE INFERENTIELLE

CHAPITRE I

INITIATION AU CALCUL DE PROBABILITE

I.1- Définitions

I.1.1- Introduction

Avant de faire le calcul de probabilité et définir la probabilité on doit d'abord réaliser dans un premier temps une expérience qui nous donne l'ensemble des résultats possibles qu'on appelle "Ω".

Exemples :

Expériences	Ensemble des résultats t
1) On jette un dé et on écrit la face qui apparaît.	Ω_1 = {1, 2, 3, 4, 5,6}
2) On jette une pièce de monnaie.	Ω_2 = {P (pile), F (face)}
3) Le nombre de personne qui entrent	

dans un bureau de poste de 8^H à 14^H.	Ω_3 = N (infini)
4) On lance un dé deux fois de suite. Le résultat de l'expérience est un couple ordonné des chiffres obtenus.	$\Omega_4 = \{1,2,3,4,5,6\}\times\{1,2,3,4,5,6\}$ $= \{(1,1), (1,2),\ldots, (1,6),(2,1),\ldots, (6,6)\}$
5) On jette une pièce de monnaie trois fois de suite	Ω_5 = {PPP, PPF, PFP, PFF, FPP,…, FFF}

I.1.2 - Les évènements

Ω est appelé ensemble (référentiel) et tout ensemble de Ω est appelé évènement, autrement dit : A évènement $\Leftrightarrow$ A $\subset \Omega$ ou encore A $\in$ P(Ω).

On distingue deux types d'évènements :

- Evénement élémentaires : dans l'exemple 4, l'ensemble $\{$ (5,5)$\}$.
- Evénements composés, dans l'exemple 4, {(1,3), 2,6),…}.

ϕ : est appelé l'événement impossible.

Ω : est appelé l'événement certain.

I.1.3 - Opération sur les évènements

Un événement est une proposition logique relative au résultat de l'expérience. Un événement est réalisé ou non suivant que la proposition est vrai ou fausse. Une fois l'expérience accomplie, nous identifierons un événement à la partie de Ω pour laquelle un événement est réalisé.

a) Complémentaire de A dans t

$\overline{A}=\Omega-A=C_{\Omega}^{A}$

L'événement complémentaire est l'événement contraire de A, si A se réalise alors $\overline{A}$ ne se réalise pas.

Exemple :

Dans l'exemple 4, on appelle A la somme des deux chiffres est supérieure ou égale à 4.

A = {(2,2), (2,3),…, (6,6)}

$\overline{A}$ = {(1,1), (1,2), (2,1)}.

b) **Conjonction** : $A \cap B$

L'événement A inter B est l'événement défini par la réalisation de A et la réalisation de B en même temps.

Exemple :

A : "ensemble de la somme de deux chiffres supérieur ou égale à 4".

B : "la somme inférieure strictement à 6".

$A \cap B$ = {(1,4), (4,1), (1,3), (3,1), (3,2) (2,3), (2,2)}.

c) La disjonction : $A \cup B$

L'événement A union B est défini par la réalisation de au moins l'événement A ou bien l'événement B. dans l'exemple précédent $A \cup B = \Omega$.

d) A—B:

L'événement A—B est défini par la réalisation de A sans la réalisation de B.

e) $A \Delta B$

L'événement $A \Delta B = (A \cup B)-(A \cap B) = (A \cap \overline{B}) \cup (\overline{A} \cap B)$.

f) AĊB (implication)

L'événement défini par, lorsque A se réalise alors B se réalise aussi.

I.1.4- Algèbre des événements

On dit que A est une algèbre d'événement si et seulement si :

- $A \subset o(t)$

- $\forall A \in A$, $C_t^A \in A$

- $\forall$A, B$\in$A, (A$\cup$B) $\in$A

- $\Omega\in$A.

Exemples :

- $P(\Omega)$ est une algèbre d'événements.

-Pour tout évènement A, l'ensemble $\{A, C_t^A, \Omega, \phi\}$ est une algèbre.

Remarque:

1. Si A est une algèbre d'événements, si $(A_i)_{i\in I=1,\ldots,m} \in A \Rightarrow \left(\bigcup_{i=1}^{m} A_i\right) \in A$, autrement dit, il y a stabilité par passage à une union finie.

2. A est une algèbre $\Rightarrow \Omega\in A \Rightarrow C_t^t = \phi\in A$.

3. A est une algèbre, je considère un ensemble fini d'éléments

$(A_i)_{i\in I} \in A \Rightarrow \left(\overline{A_i}\right)_{i\in I_+} \in A \Rightarrow \left(\bigcup_{i\in I}\overline{A_i}\right)\in A \Rightarrow \left(\overline{\bigcap_{i\in I}\overline{A_i}}\right)\in A \Rightarrow \left(\bigcap A_i\right)\in A$, autrement dit, il y a stabilité par passage à une intersection finie.

4. A, B$\in$A $\Rightarrow$A, $\overline{B}$ $\in$ A $\Rightarrow$(A-B)$\in$ A

AΔB=(A-B)$\cup$ (B-A)$\in$ A

I.1.5- Tribu ou $\pm$-Algèbre

Définition

On dit que A est une σ-algèbre d'événements si et seulement si :

- A$\subset$P(Ω)
- $\forall$A$\in$A, $\overline{A}$ $\in$A
- $\forall$i$\in$N, $A_i\in$A$\Rightarrow\left(\bigcup_{i\in N} A_i\right)\in$A.

- $\Omega \in A$,

Remarque: Toute tribu est une Algèbre.

Théorème : L'intersection d'un ensemble de tribu (Algèbre) est encore une tribu (respectivement une Algèbre) sur Ω.

I.1.6 - Système complet d'évènements

On dit qu'un ensemble S est un système complet d'évènements ssi :

- $S \subset P(\Omega)$
- $\left(\bigcup_{A \in S} A \right) = t$
- $\forall A \in S,\ B \in S,\ A \neq B \Rightarrow A \cap B = \phi$

I.1.7- Algèbre (Tribu) engendrée par un ensemble d'évènements

Définition :

Soit A un ensemble d'évènements différent de l'ensemble vide $A \subset P(\Omega)$, on appelle Algèbre (σ-algèbre) engendrée par A la plus petite algèbre (σ-algèbre) qui comprend A.

A algèbre (σ-algèbre) engendrée par A, d'une manière équivalente :

- A algèbre
- $A \subset A$
- A algèbre (σ-algèbre)
- $A \subset A' \Rightarrow A \subset A'$.

On envisage deux cas:

1. A est système complet fini ou dénombrable:

$A = \{a_i,\ i \in I\}$, alors:

L'algèbre engendrée par A(σ-algèbre) si I est infini est formée d'évènements qui s'écrivent sous la forme :

$\bigcup_{i \in J} a_i$, $J \subset I$

2. Ω=R

A est l'ensemble des intervalles de la forme $\left]-\infty, a\right[$ alors la tribu engendrée par A comprend les intervalles fermés et ouverts d'extrémité a et b, les demi-droites de la forme $\left]a, +\infty\right[$.

Les singletons : Tous les ensembles qui s'écrivent sous la forme de l'union sur un ensemble infini dénombrable d'intervalles ou bien de singletons, cette tribu est la tribu de Borel (B_R).

I.1.8 - Espace probabilisable

Définition :

Soit Ω l'ensemble fondamental et A une algèbre ou bien une σ-algèbre d'événement, le couple (Ω, A) est appelé espace probabilisable.

Définition :

Soit le couple (Ω,A) un espace probabilisable , on appelle probabilité sur A toute application P : A $\rightarrow$R+ telle que :

- P(Ω)=1
- $\forall$A$\in$A, $\forall$B$\in$A: A$\cap$B=ϕ: P (A$\cup$B) =P (A) +P (B).

Dans le cas ou A est une σ-algèbre, la 2eme condition est remplacée par :

$(A_i)_{i\in N}$ $\in$A, $P\left(\bigcup_{i\in N} A_i\right)=\sum P(A_i)_{i\in N}$

$\forall$i, j$\in$N; $A_i \cap A_j=\phi$.

Le triplet (Ω, A, P) est appelé espace probabilisé.

Conséquences :

1. P($\overline{A}$)=1-P(A)

En effet:A$\cup\overline{A}$

$P(\Omega) = P(A)+P(\overline{A})$

$1=P(A)+P(\overline{A})$

$P(\overline{A}) =1-P(A)$

2. $A \subset B \Rightarrow B = A \cup (B-A)$

   $P(B) = P(A) + P(A-B)$

   De plus $P(A) \leq P(B)$

3. $A\cup B=A\cup(B-(A\cap B))$

   $P(A\cup B)=P(A)+P(B-(A\cap B)$

   $P(A\cup B)=P(A)+P(B)-P(A\cap B)$.

Exemple:

On jette un dé. Déterminer la probabilité d’avoir un chiffre inférieur à 5.

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

$P(\Omega)=P(\{1\})+ P(\{2\})+ P(\{3\})+ P(\{4\})+ P(\{5\})+ P(\{6\})=1$

$6P(\{i\})=1\Rightarrow P(\{i\}) =1/6$

$A= \{1, 2, 3, 4\}$

$P(A) =4/6=2/3$.

D’une manière générale si A est la σ-algèbre engendrée par le système complet

$S= \{A_i\}$ $i\in I$, $I\subset N$. La probabilité P sur A est définie si on définit toutes les probabilités $p_i =P(A_i) \geq 0$, $\forall i\in I$

$\sum_{i\in I} p_i = \sum_{i\in I} P(A_i) = P(\Omega)=1$

$P(A) = \sum_{j\in J} p_j$, $J\subset I$.

1. Si Ω est fini$\Rightarrow P(A)=\text{card } A/\text{card}\Omega$
2. Si Ω est infini ($\Omega=R$), P sera connue si on connaît la probabilité sur tous les intervalles $P=\left]-\infty, a\right[=F(a)$.

I.1.9 - Probabilités conditionnelles

Définition 1 :

Soit A, B deux évènements définis sur un même espace probabilisé (Ω, A, P) tel que $P(B)\neq 0$. On appelle probabilité de A sachant B ou bien sachant que B s'est réalisé, le nombre :

$P(A/B) = P(A\cap B)/P(B)$.

Conséquences:

L'application:

$A \rightarrow [0,1]$

$A \rightarrow Q(A)=P(A/B)$, est aussi une probabilité.

En effet :

$Q(\Omega)=P(\Omega/B)=P(\Omega\cap B)/P(B)=1$

$\forall A_1, A_2, A_1\cap A_2=\phi$

$Q(A_1\cup A_2)= P(A_1\cup A_2)\cap B)/P(B)=P((A_1\cap B)\cup(A_2\cap B))/P(B)$.

Puisque $(A_1\cap B)\cap(A_2\cap B)=\phi$,on a donc

$Q(A_1\cup A_2)=(P(A_1\cap B)+P(A_2\cap B))/P(B)= Q(A_1)+Q(A_2)$.

Définition 2 : Indépendance d'événements

On dit que A et B sont deux évènements indépendants ssi:

$P(A\cap B) =P(A).P(B)$

Ou encore

$P(A/B)= P(A)$.

Exemple :

On considère l'expérience qui consiste à lancer un dé deux fois, quelle est la probabilité d'avoir une somme inférieur à 5 (strictement) sachant qu'on a obtenu le chiffre 2 au premier lancement.

Réponse :

B= {(2,1), (2,2), ….., (2,6)}

A= {(1,1), (1,2), (1,3), (2,1), (2,2), (3,1)}

P(A/B)= P(A∩B)/P(B)=P({(2,1), (2,2)})/P(B)=(2/36)/(6/36)=1/3.

Définition 3 :

Soit $(A_i)_{i=1,\dots,n}$ un ensemble d'évènements définis sur un même espace (Ω, A, P), on dit que ces évènements sont deux à deux indépendants ssi :

$\forall i\neq j$, $P(A_i \cap A_j)=P(A_i).P(A_j)$.

Définition 4 :

On dit que l'ensemble des évènements $(A_i)_{i=1,\dots,n}$ sont totalement indépendants si:

$P(\cap_{i\in I} A_i) = \pi_{i\in I} P(A_i)$.

Théorème:

Soit ((Ω, A, P) un espace probabilisé tel que :

Si $P\left(\bigcap_{i=1}^{n} A_i\right)\neq 0$ alors :

$$P\left(\bigcap_{i=1}^{n} A_i\right)=P(A_1).\ P(A_2/A_1).\ P(A_3/A_1\cap A_2)\dots\dots\dots P\left(A_n \Big/ \bigcap_{i=1}^{n-1} A_i\right).$$

Démonstration :

$P(A_2/A_1)=P(A_1\cap A_2)/P(A_2)$ On a donc pour n=2: $P(A_1\cap A_2)= P(A_1).P(A_2/ A_1)$.

Supposons que l'égalité est vraie pour n et montrons qu'elle reste vraie pour n+1.

$$P\left(\bigcap_{i=1}^{n} A_i\right)=P\left(A_{n+1}\cap\left(\bigcap_{i=1}^{n} A_i\right)\right)= P\left(\bigcap_{i=1}^{n} A_i\right).\ P\left(A_{n+1} \Big/ \bigcap_{i=1}^{n-1} A_i\right)=P(A_1).P(A_2/A_1)\dots.\ P\left(A_n \Big/ \bigcap_{i=1}^{n-1} A_i\right).$$

$$P\left(A_{n+1} \Big/ \bigcap_{i=1}^{n-1} A_i\right).$$

Théorème des probabilités Totales

Soit $(A_i)_{i=1,\dots,n}$ un système complet d'évènements, alors, pour tout événement $B\in A$, on a

$P(B)=\sum_{i=1}^{n} P(B/A_i)P(A_i)$.

En effet :

$\Omega=\bigcup_{i=1}^{n} A_i$

$B=B\cap\Omega=B\cap(\bigcup_{i=1}^{n} A_i)=\bigcup_{i=1}^{n}(B\cap A_i)$.

$P(B)= \sum_{i=1}^{n}(B\cap A_i)$

$=\sum_{i=1}^{n} P(B/A_i).P(A_i)$.

Conséquence :

I.1.10 - Théorème des probabilités de cause :(Théorème ou formule de Bayes)

$P(A_j/B)=P(B\cap A_j)/P(B)=P(B/A_j).P(A_j)/\sum_{i=1}^{n} P(B/A_i)P(A_i)$.

Exemple :

On considère 3 urnes identiques. Dans la 1ère il y a 20 boules blanches et 5 boules rouges, dans la 2ème il y a 2 boules blanches et 98 boules rouges, dans la 3ème il y a 4 boules blanches et 6 boules rouges. Calculer la probabilité qu'une boule blanche provienne de la 1ère urne.

Réponse :

B : événement "tirer une boule blanche"

A_i : "Faire le tirage dans l'urne N°i, i=1, 2,3".

$P(A_1/B)=P(B/A_1).P(A_1)/\sum_{i=1}^{3} P(B/A_i)P(A_i)=\dfrac{\frac{1}{3}\times\frac{20}{25}}{\frac{1}{3}\times\left(\frac{20}{25}\times\frac{2}{100}\times\frac{4}{10}\right)}=0{,}65$.

$P(B/A_1)=20/25$

$P(B/A_2)=2/100$

$P(B/A_3) = 4/10$.

- **Tirage exhaustif - tirage de Bernoulli**

On considère une urne et 2 façons de tirer les éléments de cette urne, ou bien on tire **sans remise** les éléments (exhaustif) ou bien on tire chaque élément et avant le tirage suivant, on remet le précédent élément, dans ce cas on a le tirage **avec remise** (tirage de Bernoulli).

I.1.11 - Eléments d'analyse combinatoire

1. **Arrangement avec répétition :**

Définition 1 :

Soit un ensemble à n éléments qu'on note A, B,…..On appelle arrangement avec répétition de p éléments choisis parmi n éléments

$A_n^p = n \times n \times n \times \ldots \times n$ (p fois) $= n^p$.

Dans chaque arrangement de p éléments, chaque élément peut apparaître plusieurs fois.

Exemple :

Quel est le nombre de numéros de téléphone qu'on peut avoir si chaque numéro comporte 7 chiffres.

$A_{10}^7 = 10^7$.

2- Arrangement sans répétition

Définition :

On appelle arrangement sans répétition de p éléments une opération ordonnée et sans répétition de p éléments choisis parmi les n éléments.

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

Exemple :

On veut élire dans une société le bureau de direction qui comporte le directeur, le sous-directeur et le secrétaire principal. Combien de bureaux peut-on obtenir si on dispose d'une liste de 5 candidats ?

Réponse :

A_5^3=5.4.3=60 bureaux.

3- Permutation sans répétition

Définition :

On appelle permutation sans répétition de n éléments une opération ordonnée de l'ensemble de n éléments :

P_n= A_n^n=n(n-1)….(n-n+1)=n!

Exemple:

Combien de place possible peut-on avoir pour 10 invités autour d'une table qui a la forme $\amalg$?

Réponse :

P_{10}=10 !

4- Permutation avec répétition

Définition :

On considère un ensemble de n éléments formés de ∇ groupes (a,a,…,a) (α fois), (b,b,…,b) (β fois),…., (s,s,….,s) (σ fois).

n= $\alpha+\beta+\ldots+\sigma$.

On appelle permutation avec répétition de n éléments une opération ordonnée d'un ensemble à n éléments. Le nombre de permutations possibles :

$$\frac{n!}{\alpha!\beta!\ldots\sigma!}.$$

Exemple :

Quelle est le nombre de numéros possibles à l'aide de 5 chiffres qui sont 2,2,3,4,4 ?

Réponse :

Le nombre de chiffres= $\frac{5!}{2!1!...2!}$ =30.

5- Combinaison sans répétition

Définition :

Soit un ensemble de n éléments, on appelle combinaison sans répétition de p éléments choisis parmi n éléments, une opération non ordonnée et sans répétition de p éléments choisis parmi n éléments.

C_n^p : le nombre de combinaisons sans répétition de p éléments choisis parmi n éléments.

A partir d'une combinaison quelconque, on peut engendrer un arrangement sans répétition en échangeant les lettres qui existent, de façon générale, à chaque combinaison C_n^p, on peut engendrer :

a.c.d c.a.d

a.c.d

a.d.c d.a.c

p ! arrangements sans répétition.

$$A_n^p = p! C_n^p \Rightarrow C_n^p = \frac{A_n^p}{p!} = \frac{n(n-1)....(n-p+1)}{p!} = \frac{n!}{p!(n-p)!}.$$

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

Exemple :

Quel est le nombre de listes possibles à l'aide de n noms et chaque liste comporte s noms.

Réponse :

C_n^s

6- Combinaison avec répétition

Définition :

On appelle une combinaison avec répétition de p éléments choisis parmi n éléments, une opération non ordonnée et dans laquelle il y a possibilité de répéter p éléments parmi n.

$K_n^s = C_{n-p+1}^p$.

Application :

On considère dans une urne N boules tel qu'on dispose de N_1 boules de la couleur N°1, N_2 boules de la couleur N°2,……N_k boules de couleur k. Evidemment $\sum_{i=1}^{k} N_i = N$. On appelle P_k le pourcentage de boules de couleur k.

$P_k = \frac{Nk}{N}$.

On tire n boules de l'urne et on note par n_k le nombre de boules qui ont la couleur k. Cet échantillon est noté $n=(n_1*n_2*n_3*\ldots n_k)$.

1. Dans le cas d'un tirage exhaustif

On a C_N^n échantillons possibles.

Quel est le nombre de choix possibles ?

Nous avons $C_{N_1}^{n_1}$ choix possibles de n_1 boules parmi N_1.

Nous avons $C_{N_2}^{n_2}$ choix possibles de n_2 boules parmi N_2.

Nous avons $C_{N_k}^{n_k}$ choix possibles de n_k boules parmi N_k.

P ({n_1*n_2*n_3*…n_k})= $\frac{C_{N_1}^{n_1}.C_{N_2}^{n_2}.......C_{N_k}^{n_k}}{C_N^n}$.

Exemple :

Dans une urne on a 3 boules rouges, 2 boules noirs, 2 boules blanches. On fait un tirage sans remise de 2 boules. Quelle est la probabilité que ces 2 boules soient toutes les 2 rouges ?

Réponse :

P{(2*0*0)}= $\frac{C_3^2.C_2^0.C_2^0}{C_7^2}$ =3/21≈0,143.

2. Tirage avec remise

La probabilité P_k de tirer une boule de couleur k ? $P_k = \frac{Nk}{N}$, et d'une façon générale la probabilité d'avoir (n_1*n_2*n_3*…n_k)est égale à $P_1^{n_1}$, $P_2^{n_2}$,…. $P_k^{n_k}$, et cette probabilité ne change pas quelque soit l'ordre mais le nombre d'arrangement de n_1*n_2*n_3*…n_k n'est autre que le nombre d'arrangements avec répétition de n boules tels que on a n_1 de la 1^{ere} couleur , n_2 de la couleur 2, ….., n_k de la couleur k.

P({n})= $\frac{n!\,P_1^{n_1}.P_2^{n_2}....P_k^{n_k}}{n_1!\,n_2!.....n_k!}$.

Application :

Dans l'exemple précédent, quelle est la probabilité d'obtenir 2 boules rouges ?

Réponse :

P(2*0*0)= $\frac{2!}{2!0!.0!}\left(\frac{3}{7}\right)^2\left(\frac{2}{7}\right)^0\left(\frac{2}{7}\right)^0=\frac{9}{49}\approx 0{,}184.$.

CHAPITRE II

VARIABLES ALEATOIRES REELLES

Introduction :

La variable aléatoire réelle est une grandeur numérique attachée aux résultats d'une expérience aléatoire.

Exemple :

La somme des points affichés par deux dés ou la durée de vie d'une ampoule électrique.

II 1 - Définitions

Soit (Ω, A, P) un espace probabilisé dans le but de mettre une probabilité sur l'espace ($\Re$,B_R), on choisi une fonction X : $\Omega \rightarrow \Re$, tel que $\forall$I intervalle de $\Re$, (X^{-1} (I))$\in$A. X est appelée alors "variable aléatoire sur $\Re$ " et la probabilité P_X définie sur la tribu B_R est appelée "la loi de probabilité de X". Les valeurs de la variables seront désignées par des lettres minuscules (x) que l'on appelle les réalisations de X et l'ensemble des réalisations possibles de X c'est l'ensemble d'arrivée X(Ω).

1. Si X(Ω) a un nombre fini ou bien infini dénombrable on dit que X est une variable aléatoire discrète.

Exemple :

On jette une pièce de monnaie 3 fois. Le nombre de 'pile' obtenu est une variable aléatoire qui prend des valeurs : 0, 1, 2,3.

Ω={PPP, PPF,PFP, FFP, FPF, FFF,….}

{X=2}= {$\omega\in\Omega$; X(ω)=2}={PPF, PFP,FPP}

X est une variable aléatoire réelle (v.a.r.) finie et discrète.

Dans un intervalle de temps T, le nombre d'appels téléphoniques dans un bureau de poste est une variable aléatoire réelle qui prend les valeurs : 1, 2,3,…, X$\in$N, X (Ω) (est infini) et est une v.a. discrète dénombrable.

2. Si X (Ω) est une partie de $\Re$ ou bien $\Re$ tout entier, on dit que X est continue.

Exemple :

L'intervalle de temps qui sépare l'entrée de deux étudiants à la bibliothèque est une variable aléatoire positive qui prend ces valeurs dans $\Re^+$.

II.2 - Variables aléatoires discrètes

Définition :

Si X : $\Omega \rightarrow R$, X est une v.a. discrète si $X(\Omega)= \{x_1, ..., x_n)$ alors X est finie ou bien infinie dénombrable.

$\{X=x_i\}=X^{-1}(\{x_i\})=\{\omega\in\Omega ; X(\omega)=x_i\}$.

Alors tous les ensembles $\{X=x_i\}$ forment un système complet sur Ω.

$P_i= P (\{X=x_i\}) \geq 0$

$\sum P_i = 1$.

Dans ce cas l'ensemble formé des couples (x_i, P_i) s'appelle : **loi de probabilité** de la variable aléatoire de X.

Si X (Ω) est dénombrable alors : X $(\Omega)=\{x_i ; i\in N\}$; $\sum_{i=1}^{\infty} P_i = 1$.

Exemple :

On jette simultanément deux dés. On veut connaître la somme des deux chiffres affichés.

$\Omega=\{(i,j) ; i=1,...6, j=1,...6\}$

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

X : $\Omega \rightarrow R$

$(i,j) \rightarrow i+j$

$X(\Omega)=\{2,3,....,12\}$

$P_X(\{X=2\})=P(X^{-1}\{2\})=P((1,1))=1/36$.

$P_X(\{3\})= P(X^{-1}\{3\})=P((1,2),(2,1))=2/36.$

Propriétés:

- Si X $(\Omega)< \infty$ est finie :

1. **Espérance mathématique moyenne**. On appelle moyenne de X et on note

   $E(X)=\overline{X}=\sum_{i=1}^{n} x_i P_i$.

2. **Moments d'ordre r :**

   $r\in N$, $m_r(X)= \sum_{i=1}^{n} x_i^r P_i$.

   Remarque:

   $m_1= E(X)$

   $m_2= \sum_{i=1}^{n} x_i^2 p_i = E(X^2)$

3. **Moments centrés d'ordre r :**

   $\mu_r(X)= \sum_{i=1}^{n} \left(x_i - E(X)\right)^r p_i$

Remarquons que :

   $\mu_1=0.$

4. **Variance d'écart-type**

   $V(X)= \mu_2(X)=E((X-E(X))^2)$

   $= \sum_{i=1}^{n} \left(x_i - E(X)\right)^2 p_i$

   $= \sum_{i=1}^{n} \left(x_i^2 - 2x_i E(X) + E(X)^2\right) p_i$

   $=E(X^2)-2E(X)^2 + E(X)^2$

   $=E(X^2)-(E(X))^2$

$\sigma(X)=\sqrt{V(X)}$ est **l'écart-type.**

Généralisation :

Si $X(\Omega)$ est infini , on définit la moyenne, les moments centrés et les moments d'ordre r à condition que les séries soient convergentes.

$X(\Omega)$ est infini dénombrable.

Dans ce cas on écrit :.

$$\overline{X} = \sum_{i=1}^{\infty} x_i p_i$$

$$m_r = \sum_{i=1}^{\infty} x_i^r p_i$$

$$\mu_r = \sum_{i=1}^{\infty} \left(x_i - E(X)\right)^2 p_i \, .$$

Remarques :

1- Soit f une fonction définie sur $X(\Omega)$.

$$\Omega \xrightarrow{X} X(\Omega) \xrightarrow{f} \to \Re$$

On définit l'espérance mathématique par :

$$E(f(x)) = \sum_{x_i \in X(\Omega)}^{\infty} f(x_i)\ P_i$$

2- $E.(f(x)+g(x))=E(f(x))+E(g(x))$

3- $E(\lambda f(x))= \lambda E(f(x)),\ \lambda \in \Re$

On dit que l'espérance est un opérateur linéaire. En particulier :

$E(ax+b)=aE(x)+b$

$V(\lambda x)=E((\lambda x-E(\lambda x))^2$

$=E(\lambda^2((x-E(X))^2$

$=\lambda^2 V(X)$

$V(c)=E((c-E(c))^2=0$, $(E(c)=c)$.

11. 3 - Exemples de lois discrètes usuelles:

3.1. Lois uniforme sur {1,2,…,n}

X (Ω)={1,2,...,n}

P(X=1)=P(X=2)=..........=P(X=n)=1/n

E(X)= $\sum_{k=1}^{n} k\frac{1}{n} = \frac{1}{n}\frac{(n+1)n}{2} = \frac{n+1}{2}$

V(X)=E(X^2)-E(X))^2

$= \sum_{k=1}^{n} \frac{1}{n}.k^2 - \left(\frac{n+1}{2}\right)^2$

$= \frac{(n+1)(2n+1)}{6} - \frac{(n+1)^2}{4}$

$= \frac{n^2-1}{12}.$

3.2 - Loi de Bernoulli de Paramètre p

Définition :

C'est la loi d'une variable aléatoire qui prend deux valeurs, la valeur 1 avec la probabilité p et la valeur o avec la probabilité (1-p) ; elle correspond à une expérience à 2 résultats : le succès avec la probabilité p et l'échec avec la probabilité (1-p).

X (Ω)={0,1} ; $\Omega \xrightarrow{X}$ {0,1}

P(X=1)=p, P(X=0)=1-p

E(X)= $\sum x_i P_i$

=1.p+0.(1-p)=p

V(X)=E(X^2)-E(X))^2

$= \sum x_i^2 p_i - p^2$

=1.p+0.(1-p)-p^2

$=p(1-p)$.

3.3 - Loi binomiale de paramètres n et p :B(n, p)

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

$P(X=k)= C_n^k\, p^k(1-p)^{n-k}$

$\forall$ k=0,…,n et $0<p<1$

En réalité la variable X suit la loi binomiale B (n, p), de paramètres n et p, est la somme de n variables aléatoires indépendantes qui ont la même loi de Bernoulli de paramètre p.

Si on a n évènements $A_1, A_2, \ldots, A_n$ tels que : $P(A_1)= P(A_2)=\ldots=P(A_n)=p$.

$(X=k)\Leftrightarrow\exists$ k évènements A_i réalisés avec succès.

$P(X=k)= C_n^k\, p^k(1-p)^{n-k}$

$E(X)= \sum_{k=0}^{n} k\, C_n^k\, p^k (1-p)^{n-k} =n.p$

$V(X)=E(X^2)-E(X))^2= np(1-p)$.

3.4 - Loi hypergéométrique H(N, M, n)

Définition :

Soit une population statistique qui comporte N éléments parmi lesquels M possèdent une certaine propriété. On tire d'une manière aléatoire et sans remise n éléments, on appelle X la v.a. qui représente le nombre d'éléments qui possèdent la propriété dans l'échantillon tiré.

$C_M^k =0$ si $k>0$, $P(X=k)= \frac{C_M^k . C_{N-M}^{n-k}}{C_N^n}$, k=0 ,…,n

$E(X)=n.\frac{M}{N}$

$V(X)=n.\ \frac{M}{N} . \frac{N-M}{N} . \frac{N-n}{N-1}$.

Conséquence:

$N \to \infty$, $\frac{M}{N} \to p$; $0<p<1$

$H(N, M, n) \to B(n, p)$.

3.5 - Loi de Poisson P(ª)

Définition :

Elle correspond à une variable aléatoire X définie sur N avec :

$X \sim P(\lambda) \Leftrightarrow P(X=n) = e^{-\lambda}\frac{\lambda^n}{n!}$, $\forall n \in N$; $\lambda>0$

$E(X) = \sum_{n=0}^{\infty} n e^{-\lambda}\frac{\lambda^n}{n!} = e^{-\lambda}\sum_{n=1}^{\infty}\frac{\lambda^{n-1}}{(n-1)!} = \lambda$

$V(X)=E(X^2)-E(X))^2=1.\lambda=\lambda$.

Théorème d'approximation d'une loi binomiale par une loi de Poisson :

Soit une suite de variables aléatoires (X_n) de loi binomiale, pour n suffisamment grand ($n\to\infty$) ($n>20$) et $p<0,1$ et $np\to\lambda$ alors $(X_n) \xrightarrow{n\to\infty} X$ où $X\sim P(\lambda)$.

11. 4 - Vecteurs aléatoires discrets

Définition :

Soit (Ω,A,P) un espace probabilisé. On appelle vecteur aléatoire discret une application Ω dans $\Re^n$,

$Z : \Omega \to \Re^n$

$\omega \to (X_1(\omega), X_2(\omega), \ldots, X_n(\omega))$

$\omega \to (x_1, x_2, \ldots, x_n)$

Toute application partielle qui à ω associe x_i est une variable discrète.

Cas particulier : n=2

Soit (X,Y) un couple de v.a., sa loi est définie par le triplet (x_i, y_j, P_{ij}) tels que :

$x_i \in X(\Omega), y_j \in Y(\Omega)$, $P_{ij} = P(X=x_i, Y=y_j)$.

y\x	**x_1**	**x_2.........x_i..........**
y_1	P_{11}	P_{21}.......P_{i1}..........$P_{.1}$
y_2	P_{12}	P_{22}.......P_{i2}..........$P_{.2}$
.		.
.		.
.		.
y_j	P_{1j}	P_{2j}.......P_{ij}...........$P_{.j}$
.		.
.		.
.	$.P_{1.}$	$P_{2.}$.......$P_{i.}$...........1

$P(X=x_i)= P(X=x_i\cap\Omega)=P(X=x_i\cap(\bigcup_j Y=y_j\ )$

$= P(\bigcup_j X=x_j \cap Y=y_j\ )$

$=\sum_j P_{ij}$

La loi marginale de X est donnée par : $P(X=x_i)= P_{i.}= \sum_j P_{ij}$.

La loi marginale de Y est donnée par : $P(Y=y_j)=P_{.j}=\sum_i P_{ij}$.

Définition :

On définit la loi de probabilité conditionnelle de X sachant que $Y=y_j$ par :

$(x_i, \frac{P_{ij}}{P_{.j}})$ quand $X/Y=y_j$.

En effet :

$P(X=x_i/Y=y_j)=\dfrac{P(X=x_i \cap Y=y_j)}{P(Y=y_j)}=\dfrac{P_{ij}}{P_{.j}}$.

Définition :

On dit que X et Y sont deux v.a. discrètes indépendantes ssi :

$\forall x_i \in X(\Omega)$ et $\forall y_j \in Y(\Omega)$

$P(X=x_i \cap Y=y_j)= P(X=x_i).P(Y=y_j)$

En d'autres termes :

$\forall i, j$ on a: $P_{ij}=P_{i\,.}.P_{.\,j}$.

Théorème:

Si X et Y sont indépendantes alors E(X. Y)=E(X).E(Y).

En effet:

$$E(X\,Y)=\sum_{i,j} x_i y_j P(X=x_i, Y=y_j)$$

$$=\sum_{i,j} x_i y_j P(X=x_i)P(Y=y_j)$$

$$=\sum_{i} x_i P(X=x_i)\sum_{j} y_j P(Y=y_j)$$

$$=E(X).E(Y).$$

Propriété:

$$E(X+Y)= \sum_{i,j}(x_i+y_j)P(X=x_i, Y=y_j)$$

$$=\sum_{i}\left(\sum_{j}(x_i+y_j)P(X=x_i, Y=y_j)\right)$$

$$\sum_{i}\left(x_i\sum_{j}P(X=x_i, Y=y_j)+\sum_{j}y_j P(X=x_i, Y=y_j)\right)$$

$$=\sum_{i} x_i P(X=x_i)+\sum_{j} y_j \sum_{i} P(X=x_i, Y=y_j)$$

$$=\sum_{i} x_i P_i+\sum_{j} y_j \sum_{i} P(X=x_i, Y=y_j)$$

$=E(X)+E(Y)$

D'une manière plus gnérale:

$$E\left(\sum_{i=1}^{n} x_i\right) = \sum_{i=1}^{n} E(X_i)$$

$$E\left(\sum_{i=1}^{n} C_i x_i\right) = \sum_{i=1}^{n} C_i E(X_i).$$

ff.**4** - **Variables aléatoires continues**

Définition :

On dit que X est une variable aléatoire continue si X (Ω) est un intervalle de $\Re$ ou bien $\Re$ tout entier.

$X : \Omega \rightarrow \Re$

Définition :

On dit que X est une v.a. absolument continue s'il existe une fonction positive et continue telle que :

$\forall x \in \Re$, $P(X<x) = \int_{-\infty}^{x} f(t)dt$

f est appelée **densité de probabilité.**

$\forall x \in \Re$, $F_x(x) = P(X<x) = \int_{-\infty}^{x} f(t)dt$

On a donc:

$$\int_{-\infty}^{+\infty} f(t)dt = 1$$

On définit l'espérance de X si elle existe par :

$$E(X) = \int_{-\infty}^{\infty} x\, f(x)\, dx$$

De la même manière :

$V(X)=\sigma^2= \int_{-\infty}^{+\infty}(x\text{-}E(x))^2 f(x)dx$

$=E((X\text{-}E(X))^2)$

$=E(X^2)\text{-}E(X))^2$

On définit les moments centrés d'ordre k par :

$\mu_k= E((x\text{-}E(X))^k)$

$= \int_{-\infty}^{+\infty}\left(x - E(X)^k\right)f(x)dx$

D'une manière générale on définit :

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

ff. **5 - Exemples de lois continues usuelles :**

1 - La loi uniforme sur [a, b]

Elle est définie par :

$f(x)= \frac{1}{b-a}$ si $x \in [a,b]$

$=0$ sinon

On a bien

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

2 - loi de Laplace-Gauss ou normale

X suit la loi normale de moyenne m et de variance σ^2 :

On note :$X \sim N(m, \sigma^2)$

$E(X)=m$ et $V(X)= \sigma^2$,

$\forall x \in \Re$, $f(x)= \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-m)^2}{2\sigma^2}}$

Remarque :

Posons :

Y= $\frac{X\text{-}m}{\sigma}$ ~N(0,1) . Cette loi est appelée loi normale centrée réduite et elle est tabulée.

On a donc :

X=σY+m

E(Y)= $\frac{1}{\sigma}$ E(x-m)=0

V(Y)= $\frac{1}{\sigma^2}$ V(x-m)=1.

La densité de probabilité est donnée pour tout y de $\Re$ par :

f(y)= $\frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}$.

Théorème de Bienaymé - Tchebychev

Soit X une variable aléatoire de variance finie alors pour tout a>0 :

P(|X-E(X)|>a)$\leq \frac{\text{var } X}{a^2}$

En effet:

- Cas où E(X)=0

P(|X|>a)= $\int_{|t|>a}^{+\infty} f(t)dt \leq \int_{|t|>a} \frac{t^2}{a^2} f(t)dt \leq \int_{-\infty}^{+\infty} \frac{t^2}{a^2} tdt \leq \frac{1}{a^2} E(X^2) = \frac{VarX}{a^2}$.

- Cas où E(X)$\neq$0

P(|X-E(X)|>a)=P(|Y|>a) $\leq \frac{E\left((X\text{-}E(X))^2\right)}{a^2} = \frac{VarX}{a^2}$.

ff. **6 - Fonction de répartition d'une variable aléatoire**

Soit X une variable aléatoire définie sur (Ω, A, P). On appelle fonction de répartition de X une fonction définie sur $\Re$ dans l'intervalle [0,1] par :

Fx : $\Re \rightarrow$[0,1]

$x \to F_x(x)=P(X<x)$

Propriétés :

1. La fonction de répartition F_x est une fonction croissante :

En effet : Pour $x_1<x_2$:

$\{x<x_2\}=\{X<x_1\}\cup\{x_1\leq X<x_2\}$

$P(X<x_2)=P(X<x_1)+P(x_1\leq X<x_2)$

$F_x(x_2)= Fx(x_1)+P(x_1\leq X<x_2)\geq 0$

$P(x_1\leq X<x_2)=F_x(x_2)-F_x(x_1)$

$F_x(x_2)>F_x(x_1)$.

2. $\lim\limits_{x\to-\infty} F_x(x)=\lim\limits_{x\to-\infty} P(X<x)=0$

$\lim\limits_{x\to+\infty} F_x(x)=\lim\limits_{x\to+\infty} P(X<x)=1$

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

3 - F_x est continue à gauche

Soit $(\varepsilon_n)_{n\in N}$ une suite décroissante et qui tend vers 0 :

$$\lim_{n\to+\infty} F_x(x-\varepsilon_n)-F_x(x)=\lim_{n\to+\infty} P\left([x\text{-}\varepsilon_n,x[\right)=P\left(\bigcap_{n\in N}[x-\varepsilon_n,x[\right)$$

$$=P(\phi)=0.$$

Remarque :

$$\lim_{\varepsilon_n\to 0} F_x(x+\varepsilon_n)-F_x(x)=P\left(\bigcap_{n\in N}[x,\, x+\varepsilon_n[\right)=P_x(\{x\})\geq 0$$

4 - Si F_x est continue, alors :

$$P(X=x)=\lim_{k\to 0} P(x\leq X<x+k)$$

$$=\lim_{k\to 0} F(x+k)-F(x)=0.$$

5 **- Si X est absolument continue alors**

F'(X)=f(x) et on a:

P(x_1<X<x_2)=F(x_2)-F(x_1)= $\int_{x_1}^{x_2} f(t)dt$

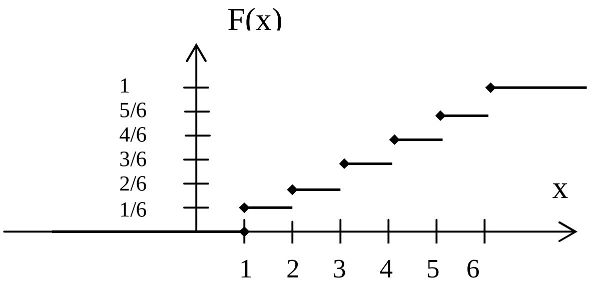
Exemples :

1. **Si X est une variable aléatoire discrète, dans ce cas :**

$F_x(x)$= P(X<x)=P$\left(\bigcup_{x_i<x}(X=x_i)\right)=\sum_{x_i<x}P(X=x_i)$.

Application

Si on lance un dé. X représente la variable du chiffre affiché⇔X (Ω)= {1,2,…16}, P_i=1/6.



X ≤ 1⇔F(x)= 0 et 1 ≤ x_2⇔F(x)= 1/6.

2- **Si X est une v.a. absolument continue**

a) Soit une variable aléatoire X : X~N(m,σ^2)

$$\Rightarrow f_X(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2\sigma^2}(x-m)^2}$$

Y= $\frac{X\text{-}m}{\sigma}$ ⇒F_Y(y)=P(Y<y)

=P(X<σy+m)

=F_X(σy+m)

F_Y (y)=σf_x(σy+m)

f_Y (y)= $\sigma.\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2\sigma^2}(\sigma y)^2}$

$f_Y(y) = \frac{1}{\sqrt{2\pi}} e^{\frac{-y^2}{2}}$.

Donc : X~N(0,1)

b) Soit une variable aléatoire X : X~N(0,1) $\Rightarrow f_X(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-y^2}{2}}$ $\forall x \in \Re$. Trouver la loi de $Y=e^X$

$F_Y(y)=P(Y<y)$

Si $y \leq 0$:

$\Rightarrow F_Y(y)= P(Y<y)= P(e^x<y)=0$

Si $y>0$:

$\Rightarrow F_Y(y)=P(Y<y)= P(e^x<y)$

$=P(X<\log y)$

$F_Y(y)= F_X(\log y)$

$F_Y(y)=\frac{1}{y} .F_X(\log y)= \frac{1}{y} . \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\log y)^2}$.

3 - Fonction de répartition d'un couple de variables aléatoires

Définition :

Soit(X,Y) un couple de variable aléatoire défini sur (Ω,A,P) dans $\Re^2$. On appelle fonction de répartition du couple (X,Y) la fonction définie par :

$(x,y) \in \Re^2$, $F(x,y)=P(X<x, Y<y)$

$=P(]-\infty,x[,]-\infty,y[)$.

Propriétés :

F est une fonction de $\Re^2$ dans [0,1].

- F est croissante par rapport à x et à y.
- F est continue à gauche par rapport à chaque variable.
- $\lim\limits_{x,y\to\infty} F_{(X,Y)}(x,y)=1$, $\lim\limits_{x,y\to-\infty} F_{(x,y)}(x,y)=0$
- S'il existe une fonction f défini de $\Re^2$ dans $\Re^+$:

$F_{(X,Y)}(x,y)=P(X<x, Y<y)=\int_{-\infty}^{x}\int_{-\infty}^{y} f(u,v)\,du\,dv$.

On dit alors que le couple (x,y) admet une densité de probabilité, et :

$\int_{-\infty}^{x}\int_{-\infty}^{y} f(u,v)\,du\,dv =1$

De plus $\frac{\partial^2 F}{\partial x \partial y}=f$.

Lois marginales :

Soit X,Y un vecteur de densité f,on définit la loi marginale de X et Y par :

$f_X(x)=\int_{\Re} f_{(X,Y)}(x,y)\,dy$

$f_Y(y)=\int_{\Re} f_{(X,Y)}(x,y)\,dx$.

Théorème :

- X et Y sont indépendantes ssi $F_{(X,Y)}(x,y)= F_X(x).\ F_Y(y)$.

Dans le cas où les v.a. sont absolument continues, X et Y sont indépendantes ssi :

$f_{(X,Y)}(x,y)= f_X(x).f_Y(y)$.

Autrement dit la densité du couple est égale au produit des densités marginales

Définition :

La densité conditionnelle de X/Y :

$f_{X/(Y=y)}(x)=f(x/y)= \frac{f_{(X,Y)}(x,y)}{f_Y(y)}$

$f_{Y/X=}\frac{f_{(X,Y)}(x,y)}{f_X(x)}$.

Fonction de répartition d'une fonction de deux variables aléatoires

Envisageons les deux cas

a) cas discret :

$Z=\varphi(X,Y)\Rightarrow G_Z(z)=P(Z<z)=P(\varphi(X,Y)<z)$

$$= \sum_{\{x_i, y_{i.}, \ddot{} (x_i, y_i) < z\}} P(X = x_i, Y = y_i)$$

b) cas continu :

Z=φ(X,Y)⇒G_Z(z)= P(Z<z)= $\displaystyle\int_{\{(x,y);\ddot{} (x,y)<z\}} \int f(x,y)dxdy$

Exemple : Somme de 2 variables aléatoires indépendantes : Z=X+Y.

K_Z(z)= P(Z<z)=P(X+Y<z)

$$= \iint_{\{(x_i, y_{i.}) \in \Re^2, (x+y)<z\}} f(x,y)dxdy$$

Si X, Y sont indépendantes, alors :

K(z)= $\displaystyle\iint_{\{(x,y);(x+y)<z\}} f(x)g(y)dx\, dy$

$$= \int_{-\infty}^{+\infty} f(x)dx \int_{-\infty}^{z-x} g(y)dy$$

$$= \int_{-\infty}^{+\infty} f(x) G_Y(z-x)dx$$

$$= \int_{-\infty}^{+\infty} g(y) F_Z(z-y)dy .$$

On déduit aussi :

- **Formule du produit de convolution :**

k(z)= $\displaystyle\int_{-\infty}^{+\infty} f(x)g(z\text{-}x)dx = \int_{-\infty}^{+\infty} g(y)f(z\text{-}y)dy$.

Théorème :

Soi (X,Y) un couple de variables aléatoires dans $\Re^2$. On pose :

(U,V)=φ(X,Y)

U=r_1(X,Y) et V=r_2(X,Y)

X=s_1(U,V) et Y=s_2(U,V)

Les s_i sont les transformés inverses de r_i. Supposons que les s_i soient continûment par rapport à chaque variable ;

J : Déterminant Jacobien

$$J = \begin{vmatrix} \partial x/\partial u & \partial x/\partial v \\ \partial y/\partial u & \partial y/\partial v \end{vmatrix} = D(x,y)/D(u,v)$$

Dans ce cas la loi de probabilité du vecteur(U,V):

$g(u,v)=| J| f(s_1(u,v)s_2(u,v))$.

Exemple:

Soit X, Y deux variables aléatoires indépendantes de même loi normale centrée réduite:

$f(x,y)= \frac{1}{2®} e^{-\frac{1}{2}\left(x^2+y^2\right)}$

U=X, V=Y/X, v=y/x $\Rightarrow$ y=uv

$$J = \begin{vmatrix} 1 & 0 \\ v & u \end{vmatrix} = u$$

$\Rightarrow$ $g(u,v)=| u| \frac{1}{2\pi} e^{-\frac{1}{2}\left(u^2+u^2v^2\right)}$

u=x, v=x+y, y=v-u

$$\Rightarrow \begin{vmatrix} 1 & 0 \\ -1 & 1 \end{vmatrix} = |1|$$

$k(v)= \int_{-\infty}^{+\infty} f(u,v-u)du$

$= \int_{-\infty}^{+\infty} f(u)g(v-u)du$.

Coefficient de correlation

- covariance de X, Y :

On défini : $Cov(X,Y)=E[(X-E(X))(Y-E(Y)]$

$=E(XY-XE(Y)-YE(X)+E(X)E(Y))$

$=E(XY)-E(X)E(Y)-E(X)E(Y)+E(X)E(Y)$

$=E(XY)-E(X)E(Y)$.

Si $X=Y \Rightarrow \quad Cov(X,Y)=Var(X)$.

$Var(X+Y)=VarX+VarY+2\,cov(X,Y)$.

Définition :

On appelle coefficient de corrélation $\rho=\frac{cov(X,Y)}{\sigma(x)\sigma(Y)}$, $-1\leq\rho\leq1$.

On démontre que $|\rho|\leq1$

Remarque :

Si X et Y sont indépendantes alors $\rho=0$.

La réciproque est généralement fausse. $\rho=0$ ssi X et Y sont indépendantes et de même loi normale.

Contre-exemple :

Soit $X=\{1,0,1\}$avec les probabilités respectives, ¼, ½, ¼ :Posons $Y=X^2$

On a $E(X)=0$,

$E(Y)=0$

$E(XY)= E(X^3)=0$

$\Rightarrow \quad cov(X,Y)=0$ et donc $\rho=0$ et pourtant X et Y ne sont pas indépendantes.

Théorème :

Image d'une variable aléatoire X :

Soit φ une fonction bijective et monotone :

$Y=\varphi(X) \Rightarrow \qquad X=\varphi^{-1}(Y)$.

Supposons φ croissante :

$F_X(x)=P(X<x)=P(\varphi^{-1}(Y)<x)=P(Y<\varphi(x))$

$F_X(x)=P(Y<\varphi(x))=F_Y(\varphi(x))$.

La dérivée $f_x(x)=f_x(\varphi(x)).\varphi'(x)$.

$$\Rightarrow \qquad f_Y(y)=\frac{f_x\left(\varphi^{-1}(y)\right)}{\varphi'\left(\varphi^{-1}(y)\right)} \qquad (1)$$

Supposons φ décroissante

$F_X(x)=P(X<x)=P(\varphi^{-1}(Y)<x)=P(Y\geq\varphi(x))$

$F_X(x)=1-P(Y<x)=1-F_Y(\varphi(x))$

$f_x(x)=-f_Y(\varphi(x)).\varphi'(x)$

$$\Rightarrow \qquad f_Y(y)=\frac{-f_x\left(\varphi^{-1}(y)\right)}{\varphi'\left(\varphi^{-1}(y)\right)} \qquad (2)$$

(1) et (2) $\qquad \Leftrightarrow \qquad f_Y(y)=\dfrac{f_x\left(\varphi^{-1}(y)\right)}{\left|\varphi'\left(\varphi^{-1}(y)\right)\right|}$.

ff. **7 - Fonction génératrice** :

Définition :

Soit X une variable aléatoire réelle, la fonction génératrice des moments est une fonction réelle notée $M_X(t)$ définie pour tout réel t par :

$M_X(t)=E(e^{tX})$

$M_X(t)=\sum_{x_i} e^{tx_i} P(X=x_i)$ dans le cas discret.

$M_X(t)=\int_{-\infty}^{+\infty} e^{tx}\, f(x)\,dx$ dans le cas continu.

Exemples:

1. Soit X la loi de Bernoulli de paramètre p : X~B(p).

Alors :

$M_X(t)=(1-p)+p\ e^t$.

2. Si X suit une loi de poisson de paramètre α : $X \sim P(\alpha)$.

$$M_X(t)=\sum_x e^{tx} e^{-\alpha} \cdot \frac{\alpha^x}{x!} = = e^{-\alpha}\sum_{x\geq 0}\frac{(\alpha . e^t)^x}{x!} = e^{-\alpha}. e^{\alpha e^t} = e^{\alpha(e^t-1)}.$$

3. Si X suit la loi de densité f définie par :

$f_X(x)=x/2$ si $0\leq x<2$

=0 sinon

$$MX(t)=\int_0^2 e^{tx}\frac{x}{2}dx=\frac{1+2te^{2t}-e^{2t}}{2t^2}.$$

- **Détermination des moments** :

Les moments sont obtenus par la dérivation de $M_X(x)$ par rapport à t et faire ensuite t=0 :

$$M'_X(t)=\frac{d}{dt}E\left(e^{tX}\right)=E\left(\frac{d}{dt}\left(e^{tX}\right)\right)\Rightarrow E(X.e^{tX})\Big|_{t=0}=E(X).$$

$$M_X''(t)=\frac{d}{dt}\left(E\left(Xe^{tX}\right)\right)\Rightarrow E\left(X^2e^{tX}\right)\Big|_{t=0}=E\left(X^2\right),.............$$

$$....,\ M_X^n(t)=E\left(X^ne^{tX}\right)\Rightarrow E\left(X^ne^{tX}\right)\Big|_{t=0}=E\left(X^n\right).$$

Remarque:

Le résultat est aussi obtenu lorsqu'on considère le développement de Mac laurin de la fonction e^{tX}.

$e^{tX}= 1+EX+\frac{(EX)^2}{2!}+.....$ $\Rightarrow$ $M_X(t)=E(e^{tX})=1+tE(X)+.....$

Propriétés de la fonction génératrice:

1. $M_{aX+b}(t)=E(e^{t(aX+b)})=e^{tb}.M_X(at)$.

2. Si X_1, X_2 sont deux variables aléatoires indépendantes, alors :

$$M_{X_1+X_2}(t)=E\left(e^{t\left(X_1+X_2\right)}\right)=E\left(e^{tX_1}.e^{tX_2}\right)=M_{X_1}(t).M_{X_2}(t).$$

2. Les fonctions génératrices des moments sont uniques autrement dit : Si deux v.a. ont la même fonction génératrice, alors elles ont la même loi.

ff.**8** - **Lois continues usuelles**

1- Loi exponentielle :

X suit une loi exponentielle de paramètre β, $\beta>0$, sa densité est définie par :

f(x)=0 si xY0

= $\frac{1}{\beta}e^{-\frac{x}{\beta}}$ **si xñ0.**

Remarque :

Si un événement se réalise suivant la loi de Poisson, le temps entre deux réalisations consécutifs est une loi exponentielle :

$E(X_{expo.})= \frac{1}{E\left(X_{Poisson}\right)}$.

La moyenne de la loi exponentielle est égale à l'inverse de la moyenne de la loi de Poisson.

* $F(X)= \int_{-\infty}^{X} f(t)dt=\int_{0}^{X}\frac{1}{\beta}e^{-t/\beta}dt$

$\Leftrightarrow$ F(X)=0 si x<0

=$1-e^{-x/\beta}$ si $x\geq0$.

* $M_X(t)=E(e^{tx})= \int_{0}^{+\infty} e^{tx}\frac{1}{\beta}e^{-x/\beta}dx$

$$=\frac{1}{\beta}\left(\frac{e^{x\left(t-\frac{1}{\beta}\right)}}{z-\frac{1}{\beta}}\right)\Bigg|_0^{+\infty}=\frac{-1}{\beta\left(t-\frac{1}{\beta}\right)}=\frac{1}{1-\beta t}$$ à condition que $(\beta t-1)<0\Rightarrow t<1/\beta$.

$M'_X(t) = \frac{-\beta}{(1-\beta t)^2}$ $\Leftrightarrow$ $M'_X(0) = \beta = E(X)$.

2 - Loi gamma

2.1. Fonction Gamma:

$$\Gamma(\alpha) = \int_0^{+\infty} x^{\alpha-1} e^{-x} dx \text{ , } (\alpha>0)$$

Propriétés :

- $\alpha>0$: $\Gamma(\alpha) = \int_0^{\alpha} x^{-\alpha-1} e^{-x} dx$.
- $\Gamma(1) = 1$
- $\alpha>1$: $\Gamma(\alpha) = (\alpha-1)\,\Gamma(\alpha-1)$
- $\forall n \in N^*$: $\Gamma(n) = (n-1)!$
- $\alpha=1/2$: $\Gamma(1/2) = \sqrt{®}$.

2. 2 - Loi gamma

Une variable aléatoire X suit la loi de Gamma de paramètre $\alpha>0$ et $\beta>0$ si sa densité s'écrit sous la forme :

$$f_X(x) = \frac{1}{T(\alpha)\beta^{\alpha}} . x^{\alpha-1} e^{-\frac{x}{\beta}} \quad \text{si } x \geq 0.$$

$$= 0 \quad \text{si } x<0.$$

α : paramètre de forme

β : paramètre d'échelle.

Propriétés :

1. Pour $\alpha=1$, on retrouve la loi exponentielle de paramètre β.

2. Si on a α variables aléatoires indépendantes selon la loi exponentielle de paramètre β, alors $\sum_{i=1}^{\alpha} X_i$ suit la loi Gamma de paramètre (α,β): X= $\sum_{i=1}^{n} X_i$ ~G(~,)

3. La somme de n variables indépendantes de loi Gamma est encore une variable de loi Gamma.

4. La fonction de répartition $F_X(x)$ est donnée par :

$$\frac{1}{T(\alpha)\beta^{\alpha}}.\int_{x}^{\infty} t^{\sim-1} e^{-\frac{t}{\beta}} dt, x \geq 0$$

=0 si x<0.

Si α est un entier positif, la fonction de répartition devient :

F(x)=1- $\sum_{k=0}^{\alpha-1}\left(\frac{x}{\beta}\right).\frac{e^{-x/\beta}}{k!}$, x≥0,

et pour $\theta=1/\beta$

F(x)=1- $\sum_{k=0}^{\alpha-1}\left(\frac{\theta x}{k!}\right)e^{-\theta x}$

5. $M_X(t)=E(e^{tx})= \int_{0}^{+\infty} e^{tx}.\frac{1}{T(\alpha).\beta^{\alpha}}.x^{\alpha-1}e^{-x/\beta}dx$

$$= \frac{1}{T(\alpha)\beta^{\alpha}}\int_{0}^{\infty} x^{\alpha-1}e^{-x\left(\frac{1}{\beta}-t\right)}dx .$$

On pose: (

$\left(\frac{1}{\beta}-t\right)x = U$ et pour $\left(\frac{1}{\beta}\right)$>t on a :

$$M_X(t)= \frac{1}{T(\alpha)\beta^{\alpha}}\frac{1}{\left(\frac{1}{\beta}-t\right)^{\alpha-1}}\int_{0}^{+\infty} U^{\alpha-1}e^{-U}\frac{dU}{\frac{1}{\beta}-t}$$

$$= \frac{\beta^{\alpha}}{\beta^{\alpha}(1-\beta t)^{\alpha}} \quad \Rightarrow \quad M_X(t)= \frac{1}{(1-\beta t)^{\alpha}} .$$

6. Lorsque α devient assez grand, on approche la loi Gamma par la normale.

Loi de KHI-2 à n degré de liberté χ_n^2 **:**

Définition :

C'est le cas particulier de la loi Gamma lorsque α=n/2 et β=2 :

F(x,n)= $\frac{1}{T(n/2)2^{n/2}}.x^{\frac{n}{2}-1}e^{-\frac{x}{2}}$, x≥0

=0,   sinon.

Si Z_1, Z_2, …Z_n sont des variables aléatoires indépendantes et identiquement distribuées selon la loi normale N(0,1) alors $\sum_{i=1}^{n} Z_i^2 \sim \chi_n^2$

Propriétés:

- E(χ_n^2**)** =n

- Var(χ_n^2**)** =2n

- Z~χ_n^2 ,T~χ_m^2 et Z indépendante de T alors Z+T~χ_{n+m}^2 .

- Si n$\to\infty$, la loi de KHI-2 est approchée par la normale si X+Y suit la loi Gamma.

ff.**9 - Convergence des suites de variables aléatoires** :

1. Différents types de convergences

a) Convergence en probabilité :

Définition 1 :

Soit $(X_n)_{n\in N}$ une suite de variables aléatoires réelles. On dit que $(X_n)_{n\in N}$ converge en probabilité vers la variable X et on note $X_n \xrightarrow{P} X$, ssi :

$\forall\varepsilon\in 0$, $\lim_{n\to+\infty}$ P(ω ;|$X_n(\omega)$-$X(\omega)$|>ε)=0

ou bien

$\forall\varepsilon$>0 et$\forall$ ¢ >0, $\exists n_0\in N$; $\forall n>n_0$ P(|$X_n(\omega)$-$X(\omega)$|>ε)< ¢

Remarque :

Une condition suffisante pour qu'une suite de variables aléatoires tende en probabilité vers une constante a est $E(X_n)=a$ et $V(X_n) \underset{n\to+\infty}{\longrightarrow} 0$.

Définition 2 :

La suite X_n est de Cauchy en probabilité si $|X_m - X_p|_{(m,p)\in N} \overset{P}{\underset{n,p\to+\infty}{\longrightarrow}} 0$.

Proposition :

Si une suite$(X_n)_{n\in N}$ converge vers X en probabilité, alors cette suite est de Cauchy en probabilité.

b) Convergence presque sûre :

Définition :

On dit que X_n est une suite qui converge presque sûrement vers X et on note : $Xn \overset{P.S.}{\longrightarrow} X$ ssi :

$$P(\omega\,;\ \lim_{n\to+\infty} X_n(\omega)=X(\omega))=1.$$

Proposition :

1. Si (X_n) converge presque sûrement, alors cette suite converge en probabilité.

2. Si (X_n) converge vers X en probabilité, alors on peut extraire une sous suite $X(j)_{j\in N}$ qui tend presque sûrement vers X.

c) Convergence en loi

Définition :

La suite (X_n) de variables aléatoires réelles définie sur un même espace de probabilité (Ω,A,P) est dite convergente en loi vers la variable aléatoire X si pour tout x réel où la fonction de répartition F de X est continue on a :

$$\lim_{n\to+\infty} \mathbf{F_n(x)} = \underset{X_n \overset{L}{\underset{n\to+\infty}{\longrightarrow}} X}{F(x)},$$

F_n: fonction de répartition de X_n. On note :

$X_n \xrightarrow{L} X$

Résumé :

Convergence en probabilité

$\Downarrow$ $\Uparrow$ (limite est constante) $\Uparrow$ $\Uparrow$ (une sous suite)

Convergence en loi **convergence presque sûrement**

Théorème limite central :

Soit (X_n) une suite de variables aléatoires indépendantes de même loi de moyenne m et d'écart type σ. Alors :

$S_n = \sum_{i=1}^{n} X_i \Rightarrow E(S_n) = n.m$

$\Rightarrow Var(S_n) = n.\sigma^2$.

$\sigma(S_n) = \sigma\sqrt{n}$.

La suite (Y_n) définie par :

$Y_n = \frac{S_n - E(S_n)}{\sigma(S_n)} = \frac{\sum_{i=1}^{n} X_i - nm}{\sigma\sqrt{n}} \xrightarrow[n\to+\infty]{L} X$ où $X \sim N(0,1)$.

Lois des grands nombres :

1. Loi faible des grands nombres :

Soit (X_n) une suite de variable aléatoire définie sur un même espace probabilisé (Ω, A, P) indépendantes et de moyenne commune m, alors :

$Z_n = \frac{S_n}{n} = \frac{\sum_{i=1}^{n} X_i}{n} \xrightarrow[n\to+\infty]{P} m$

2. Loi forte des grands nombres :

Soit (X_n) une suite de variable aléatoire et indépendantes de moyenne m et de même variance σ^2, alors la suite (Z_n) converge presque sûrement vers m quand $n \to +\infty$:

$$Z_n = = \frac{\sum_{i=1}^{n} X_i}{n} \xrightarrow[n \to +\infty]{P.S} m$$

Application :

1. Monter : $\lim_{n \to +\infty} P\left(\frac{\sum_{i=1}^{n} X_i - n}{\sqrt{n}} \leq 0 \right) = \frac{1}{2}$ si $X_1, \ldots X_n$ sont n variables aléatoires indépendantes identiquement distribuées selon la loi de poisson de paramètre $\lambda=1$.

D'après le théorème limite central (T.L.C.) :

$$\left(\frac{\sum_{i=1}^{n} X_i - n}{\sqrt{n}} \xrightarrow{L} X \right)$$; $X \sim N(0,1)$, $P(X \leq 0) \to 1/2$

2. Soit $X_1, \ldots X_n$ n variables aléatoires indépendantes identiquement distribuées selon la loi uniforme sur [0,1]. Calculer :

$\lim_{n \to +\infty} E\left(\frac{\sum_{i=1}^{n} X_i - \frac{n}{2}}{\sqrt{\frac{n}{12}}} \right) = 0$, car $E(X_i)=1/2$, $Var(X_i)=1/12$.

CHAPITRE III

ECHANTILLONNAGE ET ESTIMATION DE PARAMETRES

Introduction:

On se propose d'étudier les concepts de l'inférence statistique c'est-à-dire présenter des principes qui vont permettre sur la base de résultats d'échantillons, d'estimer les valeurs des paramètres d'une population avec un certain niveau de confiance ou encore vérifier certaines hypothèses statistiques posées sur les mêmes valeurs des paramètres d'une population. Les problèmes traités en inférence sont de deux types :

1. L'estimation de paramètres
2. Les tests d'hypothèse.

III. 1 - Les Sondages : échantillonnage d'une population

Un étude statistique sur tous les éléments d'une population est souvent physiquement irréalisable et s'avère dans plusieurs cas très chère. Alors comment obtenir certaines indications fiables sur diverses caractéristiques d'une population sans examiner tous les éléments ? C'est à cette question que nous allons tenter de répondre tout en nous limitant qu'à certains aspects pratiques de la statistique que sont les sondages ou échantillonnages.

Les sondages sont utilisés dans de nombreux secteurs par exemple les sondages d'opinion publique pour connaître avec une certaine précision l'opinion des gens sur divers sujets politiques, économiques ou autres.

Les échantillonnages visent donc à découvrir des renseignements au sujet d'une population particulière (ensemble d'unités statistiques d'individus ou d'éléments satisfaisant à une définition commune). Pour obtenir ces renseignements, il s'agit de prélever un échantillon représentatif de la population.

Définition :

Un échantillon représente une partie de la population et il est constitué d'un groupe d'unités statistiques tirées de la population préalablement définie.

Echantillon aléatoire :

III. 2 - Principe de la méthode d'échantillonnage aléatoire :

Soit une population de N unités statistiques (objets-individus) sur laquelle on veut prélever un échantillon de taille n. Nous supposons que l'on dispose d'une liste de toutes les unités qui constituent la population sans omission ni répétition. Cette liste constitue ce que l'on appelle : **la base de sondage**. Une façon de construire l'échantillon consiste à attribuer à chaque unité statistique de la population un numéro unique et à prélever ensuite par tirage au sort n numéros constituant ainsi l'échantillon ; cette façon de procéder s'appelle l'échantillonnage aléatoire et l'échantillon ainsi constitué échantillon aléatoire.

Remarque :

Lorsque chaque sous ensemble de n unités de la population parmi N unités de la population a la même probabilité d'être choisi, nous somme alors en présence d'un échantillon aléatoire simple.

A partir d'une base de sondage, on peut construire un échantillon aléatoire simple. Le tirage peut s'effectuer comme suit :

- **Tirage sans remise** (tirage exhaustif)

La base de sondage varie donc à chaque tirage.

- **Tirage avec remise** : (tirage indépendant)

La composition de la base ce sondage reste inchangée, chaque unité peut être désignée plus d'une fois.

III. 3 - Construction de l'échantillon à l'aide d'une table de nombres aléatoires.

La table est constituée des nombres 0,1,2,…,8,9 et chacune de ces valeurs a la même probabilité d'apparition. On choisit un point d'entrée dans la table et un itinéraire de lecture. On peut lire les nombres en ligne, en colonne, de haut en bas etc….

III.4 - Construction de l'échantillon par tirage systématique :

Lorsque la population est très grande (la base de sondage comporte un très grand nombre d'individus dont la numérotation est très laborieuse ou presque impossible à faire), il devient impossible de construire un échantillon par tirage au sort, on utilise alors la méthode de tirage systématique ; elle consiste en prélever les individus régulièrement espacés suivant un pas choisi.

La mise en œuvre de cette méthode de sondage se présente comme suit.

Soit N la taille de population et n la taille de l 'échantillon.

a) On calcul le rapport n/N, l'inverse de ce rapport définit le pas K=N/n, c'est l'intervalle fixe entre deux tirages.
b) On choisit de façon aléatoire le 1er individu dans la table des nombres aléatoires dont le numéro doit se situer entre 1 et K.
c) Si a est le 1er numéro choisi, l'échantillon de taille n sera composé des individus, a, a+K, a+2K,…a+(n-1)K.

III.5 - Fluctuations d'échantillonnage d'une moyenne

Pour être en mesure d'effectuer des énoncés en probabilité sur les valeurs que peut prendre la moyenne $\overline{X}$ ou d'estimer la moyenne μ de la population par intervalle de confiance ou encore effectuer un test d'hypothèse sur la moyenne μ, il faut connaître les propriétés de la distribution (loi) d'échantillonnage de $\overline{X}$.

Loi d'échantillonnage de $\overline{X}$

Définition :

La loi des différentes valeurs que peut prendre la moyenne d'échantillon $\overline{X}$ obtenue de tous les échantillons possibles de même taille est appelée loi d'échantillonnage de $\overline{X}$.

Remarques :

1. D'une façon générale, la distribution d'échantillonnage caractérise les fluctuations d'échantillonnage de toute statistique (moyenne, proportion, variance, …) calculée sur les échantillons possibles de même taille.
2. Par **la statistique**, on entend toute mesure calculée à partir des données d'échantillons, d'autre part, un paramètre est une valeur spécifique et fixe pour une population donnée de cette population.

- **Propriétés mathématiques de** $\overline{X}$: moyenne d'échantillon :

Soit un échantillon de taille n (tirage avec remise) dont les éléments possèdent un caractère mesurable X suivant une distribution de probabilité de moyenne E(X)= μ et Var(X)= σ^2. En prélevant au hasard n échantillons de taille n de cette population, on crée une suite de n variables aléatoires indépendantes $X_1,\ldots, X_n$ dont chacune a la même loi que X.

1. La moyenne d'échantillon $\overline{X}=\frac{\sum_{i=1}^{n} X_i}{n}$ est une variable aléatoire telle que

$$E(\overline{X})=\frac{E(\sum_{i=1}^{n} X_i)}{n}=\frac{\sum_{i=1}^{n} E(X_i)}{n}=\frac{\sum_{i=1}^{n}\mu}{n}=\mu$$, car les v.a. X_i ont la même loi que X.

2. $\mathrm{Var}(\overline{X})=\mathrm{Var}\frac{1}{n}\sum_{i=1}^{n} X_i=\frac{1}{n^2}\sum_{i=1}^{n}\mathrm{Var}X_i=\frac{n\sigma^2}{n^2}=\frac{\sigma^2}{n}$.

Théorème central limite :

$$\frac{\frac{\sum_{i=1}^{n} X_i}{n}-\frac{n\mu}{n}}{\frac{\sigma\sqrt{n}}{n}}\to N(0,1)$$

$\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}\to N(0,1)$, $\overline{X}$ tend à se rapprocher d'une loi normale de moyenne μ=E($\overline{X}$) et de variance $\frac{\sigma^2}{n}$ dés que n devient grand ($n\geq 30$).

Conclusion :

D'après les résultats empiriques :

1. La distribution de la moyenne $\overline{X}$ est approximativement normale dés que ($n\geq 30$).
2. Si la population possède une distribution pratiquement symétrique, il semble qu'un échantillon d'au moins 15 observations sont convenables pour que la loi de la moyenne soit approximativement normale.

3. Si la loi de X est normale, la loi de $\overline{X}$ est normale quelque soit la taille de l'échantillon.

Propriétés de la distribution d'échantillonnage de $\overline{X}$

On prélève au hasard un échantillon de taille n.

a) Population normale et variance σ^2 connues.

1. La loi de $\overline{X}$ est normale
2. E($\overline{X}$)=μ et var ($\overline{X}$)=$\frac{\sigma^2}{n}$
3. $\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \rightarrow N(0,1)$.

b) La loi de la population ainsi que la variance σ^2 inconnues. ($n \geq 30$)

On utilise les résultas du théorème central limite.

1. $\overline{X}$ est approximativement normale.
2. E($\overline{X}$)=μ
3. L'écart type de $\overline{X}$: $S(\overline{X})=\sqrt{\frac{S^2}{n}}=\frac{S}{\sqrt{n}}$,

où

$S^2=\frac{\sum(X_i-\overline{X})^2}{n-1}$ qui est une bonne estimation de σ^2 et $\frac{\overline{X}-\mu}{S(\overline{X})}=\frac{\overline{X}-\mu}{\frac{S}{\sqrt{n}}}$ suit la loi N(0,1).

III.6 - Estimation des paramètres : Objectif fondamental de l'échantillonnage d'une population.

Introduction :

Un aspect important de l'inférence statistique est celui d'obtenir à partir de l'échantillonnage d'une population, des estimations fiables de certains paramètres de cette population. Les paramètres que nous allons estimer sont la moyenne μ dans le cas d'un caractère mesurable et la proportion p dans le cas d'un caractère dénombrable.

Ces estimations peuvent s'exprimer soit par une seule valeur (estimation ponctuelle) soit par un intervalle (estimation par intervalle). Puisque l'échantillon ne donne qu'une information partielle, ces estimations seront accompagnées d'une certaine marge d'erreur.

III.6 - 1 Estimation ponctuelle :

Définition : Estimer un paramètre, c'est chercher une valeur approchée en se basant sur les résultats obtenus d'un échantillon.

Lorsque (un paramètre), une caractéristique d'une population est estimée par un seul nombre, déduit des résultats de l'échantillon, ce nombre est appelé **une estimation ponctuelle** du paramètre. L'estimation ponctuelle se fait à l'aide d'un estimateur. Cet estimateur est fonction des observations de l'échantillon. L'estimation est la valeur numérique que prend l'estimateur selon les données de l'échantillon.

Remarque : **L'estimateur** est une variable aléatoire dépendant des observations d'un échantillon aléatoire.

Exemple :

$\overline{X}$: est un estimateur ponctuel de μ : (moyenne de la population).

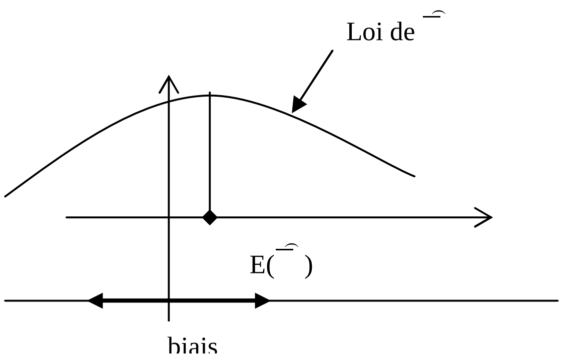
$\widehat{P}$: Estimateur ponctuel de p (la proportion d'individus dans la population échantillonnée présentant un certain caractère qualificatif).

III.6. 2- Propriétés des estimateurs ponctuels :

1. Estimateur non biaisé (sans biais)

Notons θ un paramètre de valeur inconnue d'une population, $\widehat{\theta}$ l'estimateur de θ est dit sans biais si E ($\widehat{\theta}$)= θ.

Biais = E($\widehat{\theta}$)-θ.



On sait que E($\overline{X}$)=μ ,donc $\overline{X}$ est un estimateur sans biais de μ.

$S^2=\frac{\sum\left(X_i-\overline{X}\right)^2}{n-1}$ est une estimation sans biais de σ^2.

Remarque : La médiane est un estimateur biaisé de μ si la population est asymétrique.

2. Estimateur efficace :

Un estimateur sans biais est efficace si sa variance est la plus faible parmi les variances des autres estimateurs sans biais.

Si $\widehat{\theta}_1$ et $\widehat{\theta}_2$ sont deux estimateurs sans biais du paramètre θ.

$\widehat{\theta}_1$ est plus efficace que $\widehat{\theta}_2$ $\Leftrightarrow$ E($\widehat{\theta}_1$)=E($\widehat{\theta}_2$), Var($\widehat{\theta}_1$)< Var($\widehat{\theta}_2$).

3. Estimateur convergent :

Un estimateur $\widehat{\theta}$ est convergent si sa loi tend à se concentrer autour de la valeur inconnue θ, i.e. Var $\widehat{\theta} \underset{n\to+\infty}{\longrightarrow}$ 0, n : taille de l'échantillon.

Exemple : $\overline{X}$ est un estimateur convergent puisque

Var($\overline{X}$)= $\frac{\pm^{2}}{n} \underset{n\to+\infty}{\longrightarrow}$ 0,

Remarque :

Un estimateur sans biais et convergent est dit absolument correct.

III.6. 3- Méthodes d'estimation :

Pour trouver ces estimateurs, on peut utiliser, la méthode des moindres carrés, la méthode du maximum de vraisemblance ou bien la méthode des moments.

1 - Construction d'un estimateur par la méthode des moindres carrés :

Le principe de cette méthode consiste à minimiser la somme des carrés des écarts entre les valeurs observées et les estimations obtenues.

Exemple :

Soit à estimer μ.

$Q=\sum_{i=1}^{n}\left(x_i-\widehat{\mu}\right)^2$, $\widehat{\mu}^2$ = carré des écarts.

On veut minimiser Q :

$\frac{dQ}{d\widehat{\mu}}=0$ et $\frac{d^2Q}{d\widehat{\mu}^2}>0$.

$\frac{dQ}{d\widehat{\mu}}=-2\sum_{i=1}^{n}\left(x_i-\widehat{\mu}\right)=0 \quad \Rightarrow \quad \sum_{i=1}^{n}x_i-\sum_{i=1}^{n}\widehat{\mu}=\sum_{i=1}^{n}x_i-n\widehat{\mu}=0$

$\widehat{\mu}=\bar{x}=\frac{\sum_{i=1}^{n}x_i}{n}$.

$\frac{d^2Q}{d\widehat{\mu}^2}=-2\sum_{i=1}^{n}(-1)=2n>0$.

L'estimateur est donc $\overline{X}$.

2 - Méthode du maximum de vraisemblance :

Soit X une v.a. Notons $f(x,\theta)$ sa densité, θ est le paramètre inconnu. On dispose d'un échantillon de taille n. Les observations $x_1, x_2,\ldots, x_n$ fournies par l'échantillon sont considérées comme les valeurs prises par n variables aléatoires indépendantes $X_1, X_2,\ldots, X_n$, issues de la même distribution de probabilité que X : $f(x_1,\theta)$, $f(x_2,\theta),\ldots, f(x_n,\theta)$.On associe aux valeurs observées $x_1, x_2,\ldots, x_n$, la fonction de vraisemblance L de l'échantillon résultant du produit de probabilité $f(x_i,\theta)$:

$L(x_1, x_2,\ldots, x_n)= f(x_1, \theta)\times f(x_2,\theta)\times\ldots\times f(x_n,\theta)$.

Cette fonction correspond à la probabilité jointe d'obtenir l'échantillon de valeurs $x_1, x_2,\ldots, x_n$ issu d'une population dont la distribution est connue mais non le paramètreθ. On veut déterminer une valeur θ qui maximise cette probabilité, en tenant compte des valeurs obtenues dans l'échantillon. L'estimation de θ par la méthode du maximum de vraisemblance consiste à déterminer la valeur $\widehat{\theta}$ qui maximise la fonction de vraisemblance L. Les conditions requises pour assurer que $\widehat{\theta}$ maximise L sont :

$\frac{dL}{d\theta}=0$, $\frac{d^2L}{d\theta^2}<0$.

Il est parfois plus commode de maximiser le logarithme naturel de L par rapport à θ puisque la fonction logarithme est strictement croissante et la fonction de vraisemblance comporte fréquemment des expressions de la forme d'un produit, d'une puissance…les conditions deviennent :

$\frac{d\ln L}{d\theta}=0$, $\frac{d^2\ln L}{d\theta^2}<0$.

Ln L(x_1, x_2,…, θ) est une fonction croissante et aura sa valeur maximum par la même valeur de θ qu'aurait la fonction de vraisemblance L. Cette méthode d'estimation permet de construire des estimateurs ayant la plupart des propriétés déjà traités dans ce chapitre.

Exemple :

Soit à estimer le paramètre λ d'une loi de Poisson :

$$L(x_1,....x_n)=e^{-a}\frac{a^{x_1}}{x_1!}....e^{-a}\frac{a^{x_n}}{x_n!}=\frac{1}{\prod_{i=1}^{n}x_i!}e^{-na}a^{\sum_{i=1}^{n}x_i}$$

$$\frac{dLogL}{da}=-n+\frac{\sum_{i=1}^{n}x_i}{a}=0\Rightarrow a=\frac{\sum_{i=1}^{n}x_i}{n}$$

$$\frac{d^2LogL}{da^2}\langle 0$$

3- Méthode des moments :

Soit à estimer les paramètres θ_1, θ_2,…, θ_k. On peut les estimer comme solution du système de k équations par rapport à k :

$$E(X^k)=\frac{\sum_{i=1}^{n}X_i^k}{n}.$$

Exemple :

Si X suit la loi B(n,θ), E(X)= nθ, $-=\frac{E(X)}{n}et\hat{-}=\frac{\sum_{i=1}^{n}x_i}{n^2}$.

III- 6. 4 - Estimation par intervalle de confiance :

Problème :

L'estimation ponctuelle ne fournit aucune information concernant la précision des estimations i.e. elles ne tiennent pas compte de l'erreur possible dans l'estimation, erreur attribuable aux fluctuations d'échantillonnage. Quelle confiance avons nous en une valeur unique. On ne peut répondre à cette question en considérant uniquement l'estimation ponctuelle obtenue des résultats de l'échantillon. Il faut lui associer un intervalle qui permet d'englober avec une certaine fiabilité, la vraie valeur du paramètre correspondant.

Définition :

L'estimation par intervalle d'un paramètre inconnu θ consiste à calculer, à partir d'un estimateur $\hat{\theta}$, un intervalle dans lequel il est vraisemblable que la valeur correspondante du paramètre s'y trouve. L'intervalle de confiance est défini par deux limites auxquelles est associée une certaine probabilité, fixée à l'avance et aussi élevée qu'on le désire de contenir la valeur vraie du paramètre.

$P(L_I \leq \theta \leq L_S) = P(\hat{\theta} - k \leq \theta \leq \hat{\theta} + k) = 1-\alpha$

L_I : limite inférieure de l'intervalle de confiance

L_S : limite supérieure de l'intervalle de confiance.

$1-\alpha$: la probabilité associée à l'intervalle d'encadrer la vrai valeur du paramètre θ.

k : quantité qui tient compte des fluctuations d'échantillonnage de l'estimateur $\hat{\theta}$ et de la probabilité $1-\alpha$.

Remarques :

1. α : probabilité complémentaire et correspondant au risque qu'a l'intervalle de ne pas contenir la vraie valeur du paramètre. Si on affirme que θ est compris dans l'intervalle $[L_I, L_S]$ on ne se trompera que 100α sur 100.

2. $(1-\alpha)$: lorsqu'elle est exprimée en pourcentage s'appelle **le niveau de confiance** de l'intervalle.

3. Le niveau de confiance est associé à l'intervalle et non au paramètreθ. Pour calculer cet intervalle on doit connaître la loi d'échantillonnage de l'estimateur correspondant.

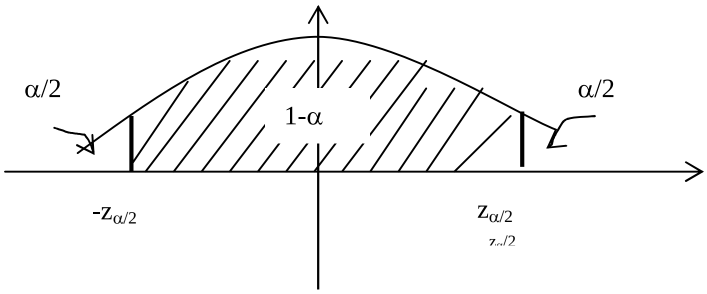
1-Estimation d'une moyenne par intervalle de confiance :

On construit l'intervalle tel que P($\overline{X}$ -k≤μ≤ $\overline{X}$ +k)=1-α. Après avoir prélevé l'échantillon et calculé l'estimation, les limites seront $\bar{x}$ -k≤μ≤ $\bar{x}$ +k.

Si la population est normale alors $\overline{X}$ ~N$\left(\mu, \frac{\sigma^2}{n}\right)$ mais si la loi de la population (caractère mesurable) est inconnue ou bien si la variance est inconnue, on utilise dans le cas n≥30 le théorème central limite et donc $\overline{X}$ suit approximativement deloi normale.

$Z = \frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$ ou bien selon le cas $Z = \frac{\overline{X}-\mu}{\frac{S}{\sqrt{n}}}$.

P($\overline{X}$ -k≤μ≤ $\overline{X}$ +k)=1-α ⇔ P(- $z_{\sim/2} \leq \frac{\overline{X}-«}{\frac{\sigma}{\sqrt{n}}} \leq$+ $z_{\alpha/2}$)=1-α.



$- z_{\sim/2} \leq \frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \leq + z_{\sim/2}$ ⇔ $- z_{\alpha/2}.\frac{\sigma}{\sqrt{n}} \leq \overline{X}-\mu \leq z_{\alpha/2}.\frac{\sigma}{\sqrt{n}}$

$- \overline{X} - z_{\alpha/2}.\frac{\sigma}{\sqrt{n}} \leq \mu \leq - \overline{X} + z_{\alpha/2}.\frac{\sigma}{\sqrt{n}}$; $\overline{X} \underbrace{- z_{\alpha/2}\frac{\sigma}{\sqrt{n}}}_{-k} \leq \mu \leq \overline{X} \underbrace{- z_{\alpha/2}\frac{\sigma}{\sqrt{n}}}_{k}$

P($\overline{X} - z_{\alpha/2}.\frac{\sigma}{\sqrt{n}} \leq \mu \leq \overline{X} + z_{\alpha/2}.\frac{\sigma}{\sqrt{n}}$)=1-α

qui est de la forme P($\overline{X}$ -k≤μ≤ $\overline{X}$ +k)=1-α ; k=$z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$.

Remarque :

Dans le cas d'un échantillon n≥30 provenant d'une population de variance inconnue, On estime σ^2 par S^2 et l'intervalle de confiance ayant un niveau de confiance 100(1-α)% d'encadrer la vraie valeur μ s'écrit donc :

$$\bar{x} - z_{\alpha/2}\frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + z_{\alpha/2}\frac{s}{\sqrt{n}}$$

Marge d'erreur associée à l'estimation et taille d'échantillon pour ne pas dépasser cette marge :

Elle est définie par $|\bar{x} - \mu|$, au niveau 1-α.

a) Si la population est normale et σ connue :

$$|\bar{x} - \mu| \leq z_{\alpha/2}.\frac{\sigma}{\sqrt{n}}$$

$$- z_{\alpha/2}.\frac{\sigma}{\sqrt{n}} \leq \bar{x} - \mu \leq$$

On remarque que plus le niveau de confiance est élevé, plus la marge est grande). Si on appelle E cette marge, on a $\sqrt{n} = \left(\frac{z_{\alpha/2}.\pm}{E}\right)$.

b) Si n≥30 et σ inconnue : $|\bar{x} - \mu| \leq z_{\alpha/2}.\frac{s}{\sqrt{n}}$.

$z_{\alpha/2}.\frac{s}{\sqrt{n}}$ =E (erreur maximale probable).

$$\sqrt{n} = \left(\frac{z_{\alpha/2}.s}{E}\right)$$

- **Loi de Student**

Lorsque l'échantillonnage s'effectue à partir d'une population normale de variance inconnue et que la taille de l'échantillon est petite (n<30) ; l'estimation de σ^2 par S^2 n'est plus fiable. Elle varie trop d'échantillon en échantillon. Dans ce cas l'écart réduit

$T = \dfrac{\overline{X} - \mu}{S/\sqrt{n}}$ ne suit plus la loi normale, mais plutôt suit la loi de Student à ν=n-1 (ν : degré de liberté).

- **Propriétés de la loi de Student :**

Cette loi est définie sur]-∞, +∞[,symétrique par rapport à l’origine et tabulée.

$\sum (x_i - \overline{x}) = 0 \Rightarrow$ On ne calcule pas n écart mais (n-1). On perd un degré de liberté.

* Construction de l’intervalle de confiance :

$\overline{x} - t_{\alpha/2,\neg} \dfrac{s}{\sqrt{n}} \leq \mu \leq \overline{x} + t_{\alpha/2,\neg} \dfrac{s}{\sqrt{n}}$, ν=n-1.

Exemples :

2 - Estimation d’une proportion p par intervalle de confiance :

$\hat{P}$: Estimateur de p qui est la proportion possédant un certain caractère qualificatif.

$\hat{p}$: Valeur observée de $\hat{P}$, pour n.$\hat{p}$ ≥5 et n (1-$\hat{p}$)≥5 ,on déduit que :

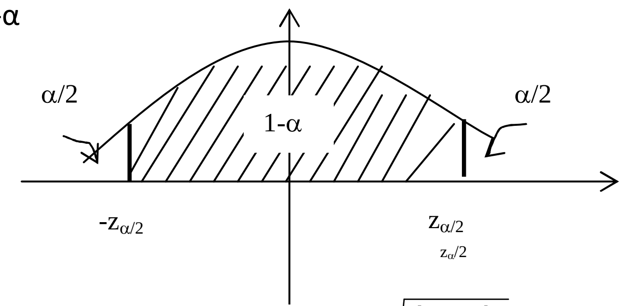
- $\hat{P}$ est approximativement normale :

- E($\hat{p}$)=p et σ($\hat{p}$)=$\sqrt{\dfrac{p(1-p)}{n}}$

$Z = \dfrac{\hat{P} - p}{\sqrt{\dfrac{p(1-p)}{n}}}$ suit la loi N(0,1) et l’intervalle de confiance pour p est :

$\hat{p} - z_{\alpha/2}\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}} \leq p \leq \hat{p} + z_{\alpha/2}\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}$.

$P(-z_{\alpha/2} < Z < +z_{\alpha/2}) = 1-\alpha$



La marge d’erreur est donnée par E=$|p$ $_{,2}\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}$.

CHAPITRE IV

TESTS D'HYPOTHESES

IV. 1- Concepts dans l'élaboration d'un test d'hypothèse

Définitions :

1. **Une hypothèse statistique** est un énoncé (une affirmation) concernant les caractéristiques (valeurs de paramètres, forme de la distribution des observations) d'une population.
2. **Un test d'hypothèse** (ou test statistique) est une démarche qui a pour but de fournir une règle de décision permettant sur la base de résultats d'échantillon de faire un choix entre deux hypothèses statistiques. Les hypothèses statistiques qui sont envisagées a priori s'appelle l'hypothèse nulle H_0 et l'hypothèse alternative H_1.

H_0 : L'hypothèse selon laquelle on fixe a priori un paramètre de la population à une valeur particulière.

H_1 : n'importe quelle autre hypothèse qui diffère de l'hypothèse H_0.

Remarque :

C'est l'hypothèse nulle qui est soumise au test et tout au long du travail, on suppose que cette hypothèse est vraie.

3. Seuil de signification d'un test d'hypothèse :

Le risque, α, de rejeter à tort l'hypothèse nulle H_0 alors qu'elle est vraie s'appelle le seuil de signification du test.

α = P(rejeter H_0/ H_0 vrai)= P(choisir H_1/H_0 vraie),

à ce seuil α, on fait correspondre par la loi d'échantillonnage de la statistique une région de rejet de H_0 (région critique). La surface (l'aire) de cette région est égale à α. La région de rejet est donc constituée d'un ensemble de valeurs de la statistique qui conduisent au rejet H_0.

Résumé :

Pour pouvoir faire un test on considère une statistique qui est une variable aléatoire dont la valeur observée sera utilisée pour décider du rejet ou non rejet de l'hypothèse H_0, et la loi de cette statistique est déterminée en supposant toujours que H_0 est vraie.

- **Risques de 1ère espèce et de 2ème espèce :**

Définition :

- α=P (rejeter H_0/H_0 vraie) est la probabilité de commettre une erreur de 1ère espèce.

- β=P (ne pas rejeter H_0/H_1 vraie) est la probabilité de commettre une erreur de 2ère espèce.

La puissance d'un test est la valeur (1-β) et c'est la probabilité de rejeter H_0 lorsque l'hypothèse vraie est H_1. Plus β est petit plus le test est puissant.

		La décision est	Probabilité de prendre cette décision avant l'expérience	La décision est	Probabilité de prendre cette décision avant l'expérience
Conclusion du test	Ne pas rejeter H_0	bonne	1-α	fausse	β
	Rejeter H_0	fausse	α	bonne	1-β

Le risque de 1ère espèce est le risque du producteur ou du fournisseur et le risque de 2ème espèce est le risque du consommateur.

IV. 2 - Les tests paramétriques

IV. 2.1- Test concernant la moyenne «.

Supposant que la moyenne μ de notre caractère est égale à une valeur fixe μ_0 ($\mu=\mu_0$), quelles sont les régions de rejet et non rejet de cette hypothèse H_0?

a) Test bilatéral

$H_0 : \mu=\mu_0$

$H_1 : \mu<\mu_0$ ou $\mu>\mu_0$

$H_1=\mu\neq\mu_0$

$\mu < \mu_0$	$\mu = \mu_0$	$\mu > \mu_0$
région de rejet H_0	région de non rejet de H_0	région de rejet H_0

$\bar{x}_{c_1}$ « $_0$ $\bar{x}_{c_2}$

Si à la suite des résultats de l'échantillon, la valeur de $\overline{X}$ se situe dans l'intervalle $\bar{x}_{C_1} < \overline{X} < \bar{x}_{C_2}$, on ne pourra pas rejeter H_0 au seuil choisi que si $\overline{X} > \bar{x}_{C_2}$ ou bien $\overline{X} < \bar{x}_{C_1}$.

b) **Test unilatéral**

Lorsque on s'intéresse au changement de la moyenne μ dans une seule direction, on fait un test unilatéral, par exemple :

$H_0 : \mu = \mu_0$

$H_1 : \mu < \mu_0$

$\mu < \mu_0$	$\mu \geq \mu_0$
région de rejet H_0	région de non rejet de H_0
	$\bar{x}_C$ μ_0 $\bar{x}$

Application :

Supposons que l'on s'intéresse au test sur la moyenne et nous déterminons la règle de décision : $H_0 : \mu = \mu_0$ et $H_1 : \mu \neq \mu_0$. On suppose de plus que la population est de loi normale et de variance σ^2 connue. On prélève un échantillon de taille n et on considère que $\overline{X}$ comme statistique convenable pour ce test : $\overline{X}$ ~N (μ_0, $\frac{\sigma^2}{n}$) Pour un seuil de signification fixé à α au départ :

P (rejet H_0/H_0 vraie) = α, P (non rejet H_0/H_0 vraie)= 1-α

P ($\overline{X} \le \overline{x}_{C_1}$ /H_0 vraie)+ P($\overline{X} > \overline{x}_{C_2}$ /H_0 vraie)= α

P (non rejet H_0/H_0 vraie) = P($\overline{x}_{C_1} < \overline{X} < \overline{x}_{C_2}$)=1-$\alpha$

$$p\left(\frac{\overline{x}_{C_1}-\mu_0}{\sigma/\sqrt{n}} < \frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}} < \frac{\overline{x}_{C_2}-\mu_0}{\sigma/\sqrt{n}}\right)=1-\alpha$$

p($-z_{\alpha/2} < Z < z_{\alpha/2}$)=1-$\alpha$

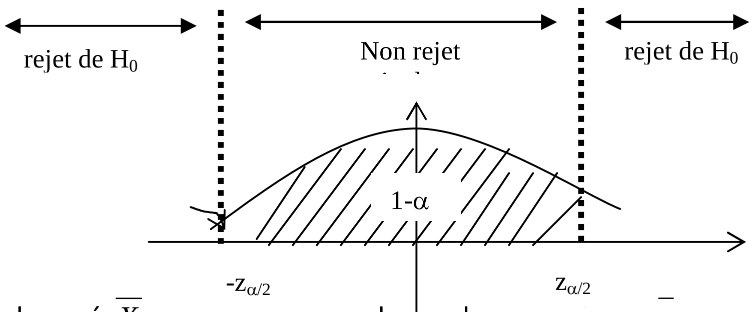
d'où :

$\overline{x}_{C_1} = \mu_0 - z_{\alpha/2}.\ \sigma/\sqrt{n}$

$\overline{x}_{C_2} = \mu_0 + z_{\alpha/2}.\ \sigma/\sqrt{n}$

Conclusion:

On va rejeter l'hypothèse H_0, si $\overline{X}$ est supérieur à $\overline{x}_{C_2}$ ou bien $\overline{X} < \overline{x}_{C_1}$. On rejette si Z<-$z_{\alpha/2}$ ou bien Z>+$z_{\alpha/2}$.



Ainsi si l'écart observé $\overline{X}$ - μ_0 est plus grand en valeur absolue à $\overline{x}_{C_2}$ - μ_0 ou bien $\overline{x}_{C_1}$ - μ_0 , nous dirons que $\overline{x}$ - μ_0 est statistiquement significative au seuil α.

Tableaux : synthèse des tests sur une moyenne

Conditions d'application : Echantillon prélevé au hasard d'une population normale de variance connue. - Hypothèse nulle : H_0 : $\mu=\mu_0$ - Seuil de signification : α $Z=\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}$ ~N(0,1)	
Hypothèses alternatives	Règles de décisions
H_1 : $\mu\neq\mu_0$ H_1 : $\mu>\mu_0$ H_1 : $\mu<\mu_0$	Rejeter H_0 si : $Z<-z_{\alpha/2}$ ou $Z>z_{\alpha/2}$ Rejeter H_0 si : $Z>z_\alpha$ Rejeter H_0 si : $Z<-z_\alpha$

Conditions d'application : Echantillon de grande taille $n\geq30$ et prélevé au hasard. - Hypothèse nulle : H_0 : $\mu=\mu_0$ - au seuil de signification : α -sous H_0 vraie $Z=\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}$ ~N(0,1) où $S^2=\frac{\sum\left(\overline{x_i}-\overline{x}\right)^2}{n-1}$	
Hypothèses alternatives	Règles de décisions

<table>
<tr><td>$H_1 : \mu \neq \mu_0$

$H_1 : \mu > \mu_0$

$H_1 : \mu < \mu_0$</td><td>Rejeter H_0 si : $Z < -z_{\alpha/2}$ ou $Z > z_{\alpha/2}$

Rejeter H_0 si : $Z > z_{\alpha}$

Rejeter H_0 si : $Z < -z_{\alpha}$</td></tr>
<tr><td colspan="2">Conditions d’application : Echantillon de taille n<30 et prélevé au hasard.

- Hypothèse nulle : $H_0 : \mu = \mu_0$

- au seuil de signification : α

sous H_0 vraie

$T = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}$ ~Student à (n-1) d.d.l</td></tr>
<tr><td>Hypothèses alternatives</td><td>Règles de décisions</td></tr>
<tr><td>$H_1 : \mu \neq \mu_0$

$H_1 : \mu > \mu_0$

$H_1 : \mu < \mu_0$</td><td>Rejeter H_0 si : $t > t_{\alpha/2}$, n-1 ou $t < -t_{\alpha/2}$, n-1

Rejeter H_0 si :$t > t_{\alpha}$, n-1

Rejeter H_0 si :$t < -t_{\alpha}$, n-1</td></tr>
</table>

Exemple :

Une machine fabrique des pièces dont la longueur suit une loi normale de paramètre μ et σ inconnus. Pour tester l’hypothèse nulle H_0 : « μ = 100 cm » contre l’hypothèse H_1 : « $\mu \neq$ 100 cm », au risque 0,05, on prélève :

1. un échantillon aléatoire (avec remise) de taille n = 10 et on obtient m =99 cm et s = 2 cm. Doit-on rejeter H_0 ?

2. un échantillon aléatoire (avec remise) de taille n = 50 et on obtient m = 99 cm et s = 2 cm. Doit-on rejeter H_0 ?

Réponse :

1. σ inconnu et n=10 <30

$T = \frac{\overline{X} - \mu_0}{s/\sqrt{n}}$ ~T_{n-1} et t=-1.58

Pour α=0.05,$t_{0.95\,;9}$=2.26.On garde H_0

2. σ inconnu et n=50 > 30.

$$Z = \frac{\overline{X} - \text{«}_0}{s/\sqrt{n}} \sim N(0,1) \text{ et } z=-3.53$$

Pour α=0.05,$z_{\alpha/2}$=1.96. On rejette H_0.

IV. 2.2 - test d'une proportion :

Les conditions d'applications :

On considère un échantillon de grande taille prélevé au hasard d'une population binomiale de sorte que np≥5 et n(1-p)≥5 (n>30 est dans bien des cas suffisant).

- Hypothèse nulle : H_0 : $p=p_0$

- au seuil de signification : α

- et sous H_0 vraie

$$Z = \frac{\hat{P} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \sim N(0,1)$$

Hypothèses alternatives :

H_1 : $p \neq p_0$ rejeter H_0 si $Z<-z_{\alpha/2}$ ou $Z> z_{\alpha/2}$

si - $\hat{p} < \hat{p}_{c_1} = p_0 - z_{\alpha/2}\sqrt{\frac{p_0(1-p_0)}{n}}$ ou $\hat{p} > \hat{p}_{c_1} = p_0 + Z_{\alpha/2}\sqrt{\frac{p_0(1-p_0)}{n}}$

H_1 : $p>p_0$ rejeter H_0 si $Z>z_\alpha \Leftrightarrow \hat{P} > \hat{p}_c = p_0 + z_\alpha\sqrt{\frac{p_0(1-p_0)}{n}}$

H_1 : $p<p_0$ rejeter H_0 si $Z<-z_\alpha \Leftrightarrow \hat{P} < \hat{p}_{c_1} = p_0 - z_\alpha\sqrt{\frac{p_0(1-p_0)}{n}}$.

Exemple:

D'après une étude sur le comportement du consommateur, il semble que 2 consommateurs sur 5 sont influencés par la marque de commerce lors de l'achat d'un bien. La directrice du marketing d'un grand magasin à rayons a interrogé 200 consommateurs choisis au hasard afin de connaître leur comportement d'achat. Sur ces 200, 65 se disent influencés par la marque de commerce.

Est-ce que ce sondage permet de supporter, au seuil α=0,01, les conclusions de l'étude sur le comportement du consommateur ?

P_0=2/5= 0.4 ;

$$\hat{p} = \frac{65}{200} = 0.325; np_0 = 80\rangle 5; n(1-p_0) = 120\rangle 5$$
$$H_0 : p = 0.4$$
$$H_1 : p\rangle 0.4$$
$$z = -2.165; z_{\sim} = 2.33$$

Oui les consommateurs sont influencés par la marque du commerce

IV. 2.3 - Test concernant la variance

Rappel :

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)} \Leftrightarrow \frac{\sum (X_i - \overline{X})}{\sigma^2} \sim \chi^2_{(n-1)}.$$

Loi de Student : $T = \dfrac{\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}}{\sqrt{\dfrac{(n-1)S^2}{\sigma^2}/n-1}} = \dfrac{\overline{X}-\mu}{s/\sqrt{n}}$

Exemple introductif :

La durée de vie d'un composant électronique est une v.a. normale de variance σ^2=2500 dans un échantillon aléatoire de taille n=16.

Quelle est la probabilité que la variance soit comprise entre 1210 et 4166 ?

Solution :

P (1210 $\leq S^2 \leq$ 4166) = P$\left(\frac{15\times 1210}{2500} \leq \frac{(n-1)S^2}{\pm^2} \leq \frac{15\times 4166}{2500}\right) = P\left(x_c^2 \leq x_{15}^2 \leq x_d^2\right) = 0{,}90.$

- **Synthèse sur test pour la variance dans une population normale**

Conditions d'application

On considère un échantillon d'application de taille n prélevé au hasard d'une population normale. L'hypothèse nulle H_0 : $\sigma^2 = \sigma_0^2$ et au seuil de signification α sous H_0 vraie. On considère la statistique :

$$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2}$$

Hypothèses alternatives	règles de décision
$H_1 : \sigma^2 \neq \sigma_0^2$	rejeter H_0 si $\chi 2 > \chi^2_{\alpha/2,\, n-1}$ ou $\chi 2 < \chi^2_{1-\frac{\alpha}{2},n-1}$
$H_1 : \sigma^2 > \sigma_0^2$	rejeter H_0 si $\chi 2 > \chi^2_{\alpha,\, n-1}$
$H_1 : \sigma^2 < \sigma_0^2$	rejeter H_0 si $\chi 2 < \chi^2_{1-\alpha,\, n-1}$

Exemple :

Un laboratoire utilise une technique pour doser la teneur en calcium d'un produit laitier et on se demande si le nouveau procédé sera adopté. Le dosage est assuré avec un écart-type de 8 milligrammes et le nouveau procédé sera adopté s'il y'a réduction de la dispersion. Sur un échantillon de 10 valeurs on a obtenu $s^2 = \frac{130}{9}$. Peut-on adopter le nouveau procédé ou non pour α=0,05 ?

Réponse :

$$\frac{(n-1)s^2}{\sigma^2} = \frac{130}{64} = 2.03$$

$$\chi^2_{9;0.05} = 3.32$$

On rejette H_0. On adoptera donc le nouveau procédé de fabrication reconnu plus précis que l'ancien.

IV. 4 -- Test d'égalité de deux moyennes de deux échantillons :

1- On désire comparer deux populations E_1,et E_2 par rapport à un même caractère C. On considère le premier échantillon $X_1, X_2,\ldots\ldots, X_{n1}$ dans la première population selon la loi normale $N(\mu,\ \sigma_1^2\ )$ avec σ_1^2 connue, et soit $Y_1, Y_2,\ldots\ldots, Y_{n2}$ un échantillon de ce caractère dans la $2^{ème}$ population selon la loi normale $N(\gamma, \sigma_2^2)$où σ_2^2 est aussi connue.

On veut faire un test $H_0 : \mu=\gamma$

$$\overline{X} \sim N\left(\mu, \frac{\sigma_1^2}{n_1}\right)$$

$$\overline{Y} \sim N\left(\gamma, \frac{\sigma_2^2}{n_2}\right)$$

$$\overline{X} - \overline{Y} \sim N(\text{« -¥}, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})$$

La statistique de test :

$$Z= \frac{\overline{X} - \overline{Y} - 0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0, 1)$$

Pour $H_1 : \mu\neq\gamma$ il y a rejet de H_0 si : $z<-z_{\alpha/2}$ ou $z>z_{\alpha/2}$

Pour $H_1 : \mu>\gamma$ il y a rejet de H_0 si : $z>z_{\alpha}$

Pour $H_1 : \mu<\gamma$ il y a rejet de H_0 si : $z<-z_{\alpha}$.

2- Si les échantillons sont prélevés au hasard et indépendamment et de tailles respectives ≥30 ,

$$Z= \frac{\overline{X} - \overline{Y} - 0}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \sim N(0, 1)$$

On retrouve la même conclusion que le cas précédent.

3 - Echantillons de petite taille (n_1 <30, n_2<30) prélevés au hasard et indépendamment de populations normales de variances inconnues mais supposées égales à une valeur commune.

On veut faire un test H_0 : $\mu=\gamma$

$$T = \frac{\overline{X} - \overline{Y}}{\sqrt{\frac{\sum \left(X_i - \overline{X}\right)^2 - \sum \left(Y_i - \overline{Y}\right)^2}{n_1 + n_2 - 2}} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

Pour $n = n_1 + n_2 - 1$

Pour H_1 : $\mu\neq\gamma$ il y a rejet de H_0 si : Rejeter H_0 si : $t>t_{\alpha/2}$, n-1 ou $t<-t_{\alpha/2}$, n-1

Pour H_1 : $\mu>\gamma$ il y a rejet de H_0 si :$t >t_{\alpha}$, n-1

Pour H_1 : $\mu<\gamma$ il y a rejet de H_0 si : si :$t <-t_{\alpha}$, n-1

IV. 3 - Tests non paramétriques

Introduction

Les tests non paramétriques sont considérés lorsque on ignore la loi de la variable aléatoire ou d'un phénomène aléatoire et concernent donc la forme de la loi de la v.a. du phénomène considéré dans la population ou bien n'importe quelle caractère qualitatif non-paramétrique de la population.

Les tests d'ajustement auront but de vérifier qu'un échantillon provient ou non d'une v.a. d'une loi connue. Parmi ces tests nous considérons le test de Ki-2, χ^2.

IV.3.1 - Le test de ¡^{2}

Définition : Les fréquences

Les fréquences vont être les effectifs absolus suivant lesquelles un caractère donné apparaît dans des classes d'une population donnée, ils seront appelés les fréquences observées et les fréquences déduites d'une loi théorique sont appelées les fréquences théoriques.

- **Principe du test**

1- La loi d'Hypothèse est complètement spécifiée

On suppose que les résultats de l'expérience appartiennent à l'une des classes A_1, A_2,……, A_k .

$\Sigma n_i = n$

On choisit la statistique $\chi^2_{cal} = \sum_{i=1}^{k} \frac{(n_i - np_i)^2}{np_i}$

p_i = probabilité théorique

L'hypothèse H_0 de processus peut être ajustée par la loi théorique et pour accepter cette hypothèse au seuil α on doit tester si :

$$p\left(\chi^2_{cal} > \chi^2_{k-1}\right) \leq \alpha$$

χ^2_{k-1} : Valeur théorique de la variable χ^2 donnée par la table.

En général si $P\left(\chi^2_{cal} < \chi^2_{k-1}\right) \leq 1-\alpha$ alors on accepte H_0.

Exemple :

Dans un jeu de pile ou face, on a observé qu'en jetant 60 fois la pièce, on observe 24 fois pile, on a donc :

$n_1=24$, $n_2=36$

Théoriquement l'expérience du jeu suit la loi binomiale de paramètre p=1/2 et n=60.

$p_1=1/2 \Rightarrow p_1=n'_1/n \Rightarrow n'_1=60.1/2=30$

$p_2=1/2 \Rightarrow p_2=n'_2/n \Rightarrow n'_2=n.p_2=30$.

$$\Rightarrow \chi^2_{cal} = \sum \frac{(n_i - np_i)^2}{np_i} = \frac{(24-30)^2}{30} + \frac{(36-30)^2}{30} = 2{,}4 \Rightarrow \chi^2_{th} = \chi^2_1 = 3{,}84 \text{ pour } \alpha=0.05$$

L'hypothèse est acceptée selon laquelle on a une probabilité uniforme, la pièce n'est pas truquée.

Remarques :

1) dans le cas continu :

Soit F la fonction de répartition de cette distribution, dans ce cas on divise la droite réelle en k intervalles $e_1, e_2, \ldots\ldots, e_{k-1}$ entre $-\infty$ et $+\infty$ et on déduit

$P([e_{i-1}, e_i[ ) = F(e_i)F(e_{i-1}) = p_i$

n_i= nombre d'individus dans $[e_{i-1}, e_i[$.

2) Le test ne peut être appliqué que pour les effectifs théoriques >5 ($np_i>5$). Dans le cas contraire on est obligé de fusionner les classes.

3) Le test est basé sur une propriété asymptotique, si n>30 on approche la loi χ^2 par la loi normale.

2- La distribution d'Hypothèse n'est pas complètement spécifiée

Dans ce cas, il faut d'abord dans un premier temps, estimer les paramètres de la loi qui sont inconnus.

On estime la moyenne et on estime la variance.

$$\bar{x} = \frac{\sum_{i=1}^{k} n_i x_i}{n}, \quad \sigma^2 = \sum \frac{n_i \left(x_i - \bar{x}\right)^2}{n}$$

A ce moment là on aura à comparer $p(\chi^2_{\text{cal}} > \chi^2_{k-1-r}) \leq \alpha$.

r : nombre de paramètres estimés.

Exemple 1

Dans une bibliothèque universitaire, on a effectué une étude sur l'affluence des usagers de terminaux donnant accès à une banque de données. On a effectué un relevé sur deux journées (considérées comme étant des journées de pointe) du nombre d'arrivées d'usagers dans un intervalle de 2 minutes. Le groupement des observations a donné lieu à la distribution ci-dessous.

Nombre d’arrivées par intervalle de 2 minutes	Fréquences
0	9
1	15
2	18
3	11
4	6
5	1

Est-ce que ce relevé permet de supporter, au seuil de signification ~**=0,05**, l’hypothèse selon laquelle le nombre d’arrivées par intervalle de 2 minutes se comporte d’après une loi de Poisson ?

Solution :

H_0=le nombre d’arrivées par intervalle de 2 mn est distribué selon une loi de Poisson de paramètre $\hat{a} = \frac{113}{60} = 1.88$.

x_i	p_i
0	0.1526
1	0.2869
2	0.2697
3	0.1690
4	0.079

5	0.0299

k=5 ; χ^2_{cal}=0.5321; $\chi^2_{0.05;3}$=7.81. On garde H_0

Rappels

La loi d'hypothèse H_0 n'est pas complètement spécifiée.

- loi de Poisson : $\hat{\lambda} = \frac{\sum_{i=1}^{k} n_i x_i}{n}$
- loi Binomiale : $\hat{p} = \frac{\sum_{i=1}^{k} n_i x_i}{n}$
- loi normale : $\text{«} = \frac{\sum_{i=1}^{n} n_i x_i}{n}, \quad \text{£}^2 = \frac{\sum_{i=1}^{k} n_i (x_i - \overline{x})^2}{n - 1}$

Exemple 2 :

Dans une usine fabricant des tubes de verres, on effectue un control visuel sur des échantillons de 20 tubes. La répartition des nombres d'échantillons avec 0 tubes défectueux, avec 1 tube défectueux,……, 6 tubes défectueux, on observe au total 80 échantillons de taille 20 sur une période de deux semaines.

Les lois de χ^2 et de Kolmogorov sont tabulées et la loi de χ^2 se fait pour différent lois de variables aléatoires.

Nombre de tubes défectueux	Nombre d'échantillons
0	13
1	21
2	19
3	12
4	9
5	4
6	2

La variable que l'on considère est le nombre de tubes défectueux. Est-ce que les observations se comportent d'après une loi binomiale au seuil de signification α ?

$n = \Sigma n_i = 80$

$$\hat{x} = \frac{\sum n_i x_i}{n} = \frac{163}{80} = 2{,}03$$

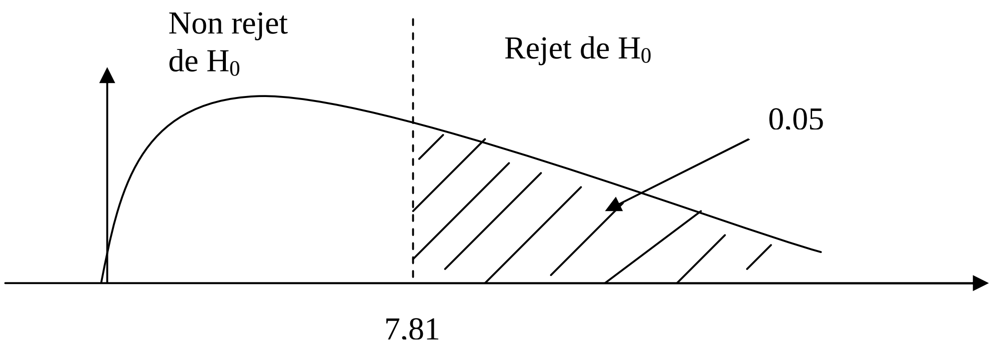
On estime $\hat{p}$ par $\frac{2{,}03}{20} = 0{,}101$.

Nombre de défectueux	n_i	p_i	80 p_i	n_i-80p_i
0	13	0,121	9,72	3,27
1	21	0,270	21,61	-0,616
2	19	0,285	22,81	-3,81
3	12	0,190	15,20	-3,20
4	9	0,0898	7,18	4,35
5	4	0,0319	2,55	
6	2	0,0089	0,712	
7 plus	0	0,0025	0,2	

H_0, le processus peut-il être ajusté par une loi binomiale de n=20 et $\hat{p}$ =0,101 ?

$$\chi^2_{Cal} = \sum_{i=1}^{5} \frac{(n_i - Np_i)^2}{Np_i} = 4{,}21$$

$$\chi^2_{0{,}0{,}05} = 7{,}81$$



Conclusion :

On garde alors H_0

Exemple 3:

Grâce à un programme sur ordinateur on simule 500 valeurs prises par une variable aléatoire X suivant une loi normale de moyenne 109 et d'écart type 0,5. Les valeurs obtenues se répartissent de la façon suivante :

CLASSE	EFFECTIF	CLASSE	EFFECTIF
[107,8 ; 108[	5	[109 ; 109,2[	75
[108 ; 108,2[	5	[109,2 ; 109,4[	85
[108,2 ; 108,4[	30	[109,4 ; 109,6[	55
[108,4 ; 108,6[	40	[109,6 ; 109,8[	30
[108,6 ; 108,8[	70	[109,8 ; 110[	10
[108,8 ; 109[	85	[110 ; 110,2[	10

Peut-on considérer au risque de $1^{ère}$ espèce de 5%, puis de 1%, que le programme est correct ?

Réponse :

Le programme est correct si $X \sim N(109 ; (0.5)^2)$.

H_0 : « $X \sim N(109 ; (0.5)^2)$ ».On pose $U = \frac{X - 109}{0.5} \sim N(0 ;1)$.

$\chi^2_{cal} = 22.1$

$\chi^2_{0.95;11} = 19.7$; $\chi^2_{0.99;11} = 24.7$

Conclusion : 19.7<22.1<24.7.

On rejette H_0 au risque 5% mais pas au risque1%.

IV.3.2 - Test d'indépendance de χ^2.

Soit X, Y, deux variables aléatoires et on veut tester leur indépendance sur la base de n observations. La fréquence absolue est notée par k_{ij} pour les valeurs $X=x_i$, $Y= y_i$ $1\leq i \leq s$ et $1\leq j \leq r$.

S'il y a indépendance alors,

$p_{ij} = p_{i.}\times p_{.j}$ et $p_{i\bullet} = \frac{k_{i\bullet}}{n} = \frac{\sum_j k_{ij}}{n}$, $p_{\bullet j} = \frac{k_{\bullet j}}{n} = \frac{\sum_i k_{ij}}{n}$

On compare les fréquences relatives observées $\frac{k_{ij}}{n}$ aux fréquences $p_{i.}$ et $p_{.j}$, pour cela on utilise l'approximation :

$$\chi^2_{Cal} = \sum_{i,j} \frac{\left[k_{ij} - n\left(p_{i\bullet} \times p_{\bullet j}\right)\right]^2}{np_{i\bullet} \times p_{\bullet j}} = n\left(\sum_{i,j} \frac{k_{ij}^2}{k_{i\bullet}k_{\bullet j}} - 1\right).$$

Le test de χ^2 aura pour degré de liberté (r-1)(s-1). On doit pouvoir faire l'approximation $np_{i.}p_{.j} \geq 5$.

Exemple

Pour comparer l'efficacité de deux médicaments agissant sur la même maladie, mais aux prix très différents, la Sécurité Sociale a effectué une enquête sur les guérisons obtenues en suivant chacun des traitements. Les résultats sont rassemblés dans le tableau suivant:

	Médicament cher	Médicament bon-marché	
Guérisons	44	156	200
Non-guérisons	6	44	50
	50	200	250

Peut- on dire que la guérison est indépendante du prix du médicament au risque 5%?

Réponse :

H_0 **:** « Il y a indépendance entre la guérison et le prix du médicament »

$\chi^2_{\text{Cal}} = 2.5$ ¡ ${}^{2}_{the} = 3.84.$

On ne rejette pas H_0.

Application :

Un expérimentateur intéressé par la prédominance visuelle et l'habilité manuelle. On choisit 400 sujets.

	Œil gauche	Les deux yeux	Œil droit
Main gauche	$37=k_{11}$	$61=k_{12}$	$27=k_{13}$
Les deux mains	$30=k_{21}$	$27=k_{22}$	$18=k_{23}$
Main droite	$61=k_{31}$	$104=k_{32}$	$35=k_{33}$

n=400

$\chi^2_{\text{Cal}} = 6{,}14, \quad \chi^2_4$ pour $\alpha = 0{,}05 \Rightarrow \chi^2_{0,0,05} = 9{,}49$

$\chi^2_{\text{Cal}} < \chi^2_{4,0,05}$

On accepte l'hypothèse qu'il y'a indépendance entre la prédominance visuelle et l'habilité manuelle.

IV.3.3 - Test Kolmogorv-Smirnov

Ce test est basé sur la comparaison de la fonction de répartition empirique (observée) de l'échantillon qu'on note F_n et la fonction de répartition théorique F de la population.

Principe du test

On considère la distance maximum $D_n=\max|F_n(x)-F(x)|$.

Le but est de tester l'hypothèse selon laquelle l'échantillon $x_1, \ldots, x_n$ provient d'une population de fonction de répartition F qu'on suppose continue. On définit la fonction de répartition empirique de l'échantillon :

$$F_n(x) = \begin{cases} 0 \text{ si } x < x_{(i)} \\ \dfrac{i}{n} \text{ si } x_i < x \leq x_{(i+1)} \\ 1 \text{ si } x > x_{(n)} \end{cases} \quad x_i \text{ échantillon ordonné, } x_1 < x_2 < .... < x_n$$

Fonction de répartition : c'est la fonction qui donne la proportion.

Le test se base sur la statistique de Kolmogorov :

$$D_n = \sup_{x \in \Re} |F_n(x) - F(x)|$$

La loi de D_n ne dépend pas de F, au niveau du seuil de significationα, on accepte l'hypothèse si p ($Dn \leq d_\alpha$)= 1-α (d_α lue à partir de tableau) et on rejette l'hypothèse dans le cas contraire autrement dit si on a p($Dn > d_\alpha$)=α.

Exemple :

Peut-on considérer les valeurs suivantes comme celles d'un échantillon d'une loi uniforme sur [0,1] :0.8-0.4-0.25-0.70-0.60-0.20-0.5-0.30-0.15-0.10-0.65-0.95-0.45-0.85-0.55.

Réponse :

H_0 **: « $F=F_0$:** loi uniforme sur[0,1] »

$F_n(x)$={0,1/15,2/15,…..14/15}, D_n= Sup | $\mathbf{F_n(x)-F_0(x)}$|=Sup | $\mathbf{F_n(x)-x}$ |=0.1 et d_α=0.34

On accepte alors l'hypothèse H_0.

Exercices d'application

Série d'exercices N°1

Exercice 1

On jette en même temps une pièce de monnaie et un dé. Ecrire l'ensemble des résultats possibles Ω.

On appelle *A* l'événement « obtenir pile » et *B* l'événement « obtenir un nombre inférieur strictement à 3 ». Ecrire les événements suivants : $\overline{A}$, $\overline{B}$, $A \cap B$, *A*-*B*, $A \oplus B$.

Exercice 2

On jette deux pièces de monnaie. Quelle est la probabilité d'obtenir deux fois pile ?

Exercice 3

Montrer que si *A* et *B* sont deux événements indépendants, alors A et $\overline{B}$ ainsi que $\overline{A}$ et $\overline{B}$ sont indépendants.

Exercice 4

On lance un dé deux fois. On appelle les événements :

A « le premier chiffre affiché est pair »,

B « le deuxième chiffre affiché est impair »,

C « les deux chiffres apparus sont de même parité ».

Montrer que *A*, *B*, *C* sont deux à deux indépendants mais ne sont pas mutuellement indépendants .

Exercice 5

Trois coups sont tirés sur une même cible, les probabilités d'atteindre la cible lors du premier, du second, et du troisième coup sont respectivement p_1=0,4 ; p_2=0,5, p_3=0,7 .

Calculer la probabilité pour que :

1/ Un de ces trois coups exactement atteigne la cible.

2/ Au moins un coup atteigne la cible.

Exercice 6

Soit une population composée de 48% d'hommes et de 52% de femmes. On suppose que 3% d'hommes et 0,5% de femmes sont atteints d'une certaine maladie. On choisit une personne au hasard.

1/ Quelle est la probabilité que cette personne soit malade ?

2/ Sachant que cette personne est malade, quelle est la probabilité que ce soit un homme ?

Exercice 7

Un laboratoire a mis au point un alcool.-test, les premiers résultats sont les suivants :

- 2% des personnes contrôlées par la police sont réellement en état d'ébriété,
- 95 fois sur 100, l'alcool-test s'est révélé positif, alors qu'une personne était réellement en état d'ébriété.
- 95 sur 100 l'alcool-test s'est révélé négatif, alors qu'une personne n'était pas en état d'ébriété.

Quelle est la probabilité qu'une personne soit réellement en état d'ébriété lorsque l'alcool-test est positif ?

Exercice 8

A l'aide de mot JACINTHE, combien peut-on former :

a) de mots de six lettres ?
b) de mots de six lettres dont la deuxième et la dernière sont des voyelles et les quatre autres des consonnes ?
c) de mots de six lettres comportant deux voyelles et quatre consonnes ?

Exercice 9

Dans une course comportant 18 partants, quelle est la probabilité pour un joueur jouant au hasard de toucher le tiercé dans l'ordre ? Dans le désordre ?

Exercice 10

Une boite contient 41 transistors, dont 5 défectueux. On prend 3 au hasard. Quelle est la probabilité que tous les trois soient défectueux ? Que deux au moins soient défectueux ?

Exercice 11

Soit une population constituée des nombres 1, 2, 3, 4, 5. Quelle est la probabilité de choisir un échantillon ordonné de taille n=2 (tirage sans remise) et calculer la probabilité de choisir un échantillon non ordonné de taille 2 (tirage sans remise).

Série d'exercices N°2

Exercice 1 :

Soit X une variable aléatoire qui prend les valeurs entières entre 0 et 10 avec les probabilités

$p_k = P(X=k) = a\,(10-k).k$

1/ Calculer a.

2/ Calculer E(X) et V(X).

On rappelle que si $S_k = 1^k + 2^k + \dots + n^k$,

$S_1 = \frac{n(n+1)}{2}$, $S_2 = \frac{n(n+1)(2n+1)}{6}$; $S_3 = (S_1)^2$ et $S_4 = \frac{n(n+1)(2n+1)(3n^2+3n-1)}{30}$

Exercice 2

Dans une urne, il y a 10 boules numérotées de 0 à 9. On tire deux boules au hasard en même temps et on rappelle X la variable aléatoire " le plus petit des nombres portés par les deux boules".

Quelle est la loi de X ? Calculer sa moyenne et sa variance.

Exercice 3

On lance un dé jusqu'à l'obtention du chiffre 5. Soit X la v. a. qui représente le nombre de jets nécessaires. Trouver la loi de X. Calculer E(X), V(X).

Exercice 4

Soient X_1, X_2, X_3, et X_4, quatre variables aléatoires de Bernoulli de paramètre *p* et indépendantes. On pose :

$Y = X_1 + 2X_2 + 4X_3 + 8X_4$.

Calculer $P(Y<8)$, $P\left(Y=6 / Y<8\right)$, $P(4 \leq Y < 8)$.

Exercice 5

Calculer E(X) et V(X) si $X \sim P(\lambda)$.

Exercice 6

On suppose que le nombre de personnes se présentant à une station service en une période de 15 minutes est une v. a. X de distribution

x_i	0	1	2	3	4
$P(X=x_i)$	0,1	0,2	0,4	0,2	0,1

De plus, la probabilité pour qu'une personne se présentant à la station, demande du super est 0,4. Soit Y la v. a. "nombre de personnes ayant demandé du super dans la période de 15 minutes.

1/ Quelle est la loi conditionnelle de Y pour $X=x_i$?

2/ En déduire la loi du couple (X,Y) puis la loi de Y.

3/ Calculer la probabilité des événements suivants :

[Y=1], [X≥2 et Y=0]

Exercice 7

Soient X et Y deux v. a. réelles indépendantes telles que X~P(λ) et Y~P(β). Quelle est la loi de X+Y et trouver P($X = k / X + Y = n$).

Série d'exercices N°3

Exercice 1

Soit *f* la fonction réelle définie par :

f(x)= $k(4x-x^2)$ si $x \in]0,4[$

=0 ailleurs

1/ Déterminer *k* pour que *f* soit la densité de probabilité d'une v.a. réelle absolument continue X.

2/ Déterminer la fonction de répartition de X et en déduire $P(1<X<2)$ et $P\left(X>3 / X>2\right)$.

3/ Calculer E(X) et V(X).

4/ Déterminer la densité de la v.a. Y= $\sqrt{X}$.

Exercice 2

Soient X, Y deux variables indépendantes de fonction de répartition F et G. On pose :

Z= Max (X, Y) et T = min (X,Y). Déterminer la fonction de répartition de Z et T à l'aide de F et G.

Exercice 3

On suppose que le temps qu'il faut en minutes à un métro pour parcourir le trajet entre la station M_1 et la station M_2 (et vice versa) est distribué uniformément (loi uniforme) sur l'intervalle [8,12] .

a- Préciser la loi de probabilité X régissant la variable aléatoire : « temps qu'il faut en minutes pour parcourir le trajet entre la station M_1 et la station M_2 ».

b- Déterminer la fonction de répartition de X.

c- Quelle est la probabilité que le métro effectue le trajet entre la station M_1 et la station M_2 en moins de 9.5 minutes.

Exercice 4

Supposons que la durée de vie d'un dispositif est distribuée selon une loi exponentielle de paramètre . Supposons que le dispositif fonctionne depuis t_0 heures : quelle est la probabilité qu'il fonctionne au moins pour une autre période de *t* heures.

Exercice 5

Soient X_1, X_2,…., X_n v.a.i.i.d. de $U_{[0,1]}$. Monter à l'aide du théorème central limite que $\lim_{n\to\infty} E\left(\dfrac{\sum_{i=1}^{n} X_i - n/2}{\sqrt{n/12}}\right)=0$.

Série d'exercices N°4

Exercice 1 :

Supposons qu'une population consiste en 5 contenants (numérotés de 1 à 5) et que le poids respectif de chacun est :

x_1=332g ; x_2=336g ; x_3=340g ; x_4=344g ; x_5=348g.

1/ Calculer la moyenne et la variance du caractère poids en grammes de cette population.

Exercice 2

Une entreprise fabrique une marque de céréales dans un contenant de 340g incluant les céréales, le procédé indique une dispersion de 10g (±=10). On a établi que le poids des contenants est distribué selon la loi normale. Pour vérifier si le procédé de remplissage se maintient autour de 340g en moyenne, on prélève de temps en temps un échantillon de 16 contenants.

1/ Déterminer les caractéristiques de la distribution d'échantillonnage du poids $\overline{X}$.

2/ Quelle est la probabilité que le poids moyen d'un échantillon se situe entre 335g et 345g ?

3/ dans 99%, entre quelles valeurs va varier la moyenne échantillonnage autour de la moyenne 340g.

2/ Quel est l'effet sur la loi de $\overline{X}$ si on choisit les échantillons de taille n=4 ; n=8 ; n=16 ; n=64 ?

Exercice 3

Montrer que $S^2 = \frac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}{n-1}$ est un estimateur sans biais de ±².

Exercice 4

Soient X_1, X_2 et X_3 un échantillon aléatoire prélevé d'une population infinie avec $E(X_i)$=« et var(X_i)=±².

Montrer que : $\hat{\text{«}} = \frac{X_1 + 2X_2 + 3X_3}{6}$ est un estimateur sans biais de « mais moins efficace que $\overline{X} = \frac{\sum_{i=1}^{3} X_i}{3}$.

<u>**Série d'exercices N°5**</u>

Exercice 1

La durée de vie d'un certain bien économique peut être représentée par une variable aléatoire absolument continue *X* de loi exponentielle de paramètre $k \in R_+^*$.

Le paramètre positif *k* est inconnu, on se propose de l'estimer à partir d'un échantillon de *n* biens indépendants dont on mesure les durées de vie. On obtient ainsi une suite $(x_1, x_2, \ldots, x_n)$ qui est une réalisation de *n*-uple $(X_1, X_2, \ldots, X_n)$ où les X_i sont indépendants et suivent tous la loi définie plus haut.

1°/ Déterminer par la méthode du maximum de vraisemblance un estimateur $\hat{k}_n$ du paramètre inconnu *k.*

1°/ L'estimateur $\hat{k}_n$ est-il sans biais ?

3°/ L'estimateur $\hat{k}_n$ est-il convergent ?

Exercice 2

Un laboratoire indépendant a vérifié la résistance à l'éclatement (en kg/cm^2) d'un réservoir à essence d'un certain fabricant. Des essais ont permis de considérer que la résistance à l'éclatement est distribuée normalement avec une variance de 9. Sur un échantillon de 10 réservoirs, les résultats conduisent à une résistance moyenne à l'éclatement de $219 kg/cm^2$. Estimer par intervalle de confiance la résistance moyenne à l'éclatement de ce type de réservoir avec un niveau de confiance de 95%.

Exercice 3

On a déterminé pour un échantillon de 36 lampes fluorescentes de puissance nominale 40W provenant d'une production de 2000 unités, le voltage de démarrage de chacune. Le voltage moyen obtenu est 93,4 volts avec un écart-type de 0,96 volts. Estimer par intervalle de confiance le voltage moyen de démarrage avec les niveaux 90% ; 95% ; et 99%.

Exercice 4

Dans un atelier mécanique, on a vérifié le diamètre de tiges tournées sur un tour automatique. Le diamètre de tiges (tournées sur un tour automatique) peut fluctuer selon le réglage du tour. 20 tiges prélevées au hasard ont été mesurées avec un micromètre de précision. Les résultats sont présentés dans le tableau suivant :

Diamètre en mm de tiges tournées			
39,5	39.9	39,7	40,0
40,6	41,5	39,1	39,9
38,4	40,0	42,6	41,1
37,8	38,5	40,0	41,3
39,4	41,2	38,4	10,8

En supposant que le diamètre des tiges est distribué selon une loi normale, estimer par intervalle de confiance le diamètre moyen de tiges de cette fabrication avec le niveau de confiance de 90%.

Série d’exercices N°6

Exercice 1

Le responsable du procédé de fabrication de tiges métalliques suggère d’introduire un nouvel alliage dans le procédé de fabrication des tiges. Avant la modification, les tiges avaient une résistance moyenne à la rupture de 50 kg/cm^2. Sur un échantillon de 40 tiges, on a obtenu une résistance moyenne de 54,5 kg/cm^2 et un écart-type de 2,4. Est-ce que l’écart observé dans la résistance moyenne à la rupture avant et après est suffisamment élevé pour conclure au seuil de signification ~ = 0,01 qu’il y a une augmentation significative de la résistance moyenne.

Exercice 2

Une entreprise fournit à un client des supports métalliques. Le client exige une longueur moyenne de 70 mm. On suppose que la dispersion de la fabrication est égale à σ = 3 mm et que la longueur suit une loi normale. On veut vérifier si le procédé opère à 70 mm. Sur un échantillon de 25 supports on a obtenu une longueur moyenne de 69 mm. Doit-on conclure au seuil de signification α = 0,05 que la machine est déréglée ?

Exercice 3

Un laboratoire utilise une technique pour doser la teneur en calcium d’un produit laitier et on se demande si le nouveau procédé sera adopté. Le dosage est assuré avec un écart-type de 8 milligrammes et le nouveau procédé sera adopté s’il y’a réduction de la dispersion. Sur un échantillon de 10 valeurs on a obtenu $s^2=\frac{130}{9}$. Peut-on adopter le nouveau procédé ou non pour α=0,05 ?

Exercice 4

Une entreprise fabrique des transistors. Un contrôle régulier est effectué pour détecter les transistors défectueux. Le processus de fabrication produit habituellement 2% de transistors défectueux. Un récent contrôle de 300 transistors donne 11 transistors défectueux. Peut-on conclure au seuil de signification α = 0,01 que ce pourcentage est anormalement élevé ?

Exercice 5

Dans une étude préparatoire qui vise à adapter le poste de conduite de camions aux dimensions des conducteurs, on a mesuré dans un laboratoire de recherche, la taillez d'un certain nombre de conducteurs de camions. Les tailles mesurées sur 80 conducteurs sont compilées dans la distribution suivante :

Tailles en cm	Nombres de conducteurs
$150 \leq X < 155$	2
$155 \leq X < 160$	2
$160 \leq X < 165$	14
$165 \leq X < 170$	20
$170 \leq X < 175$	19
$175 \leq X < 180$	15
$180 \leq X < 185$	6
$185 \leq X < 190$	2

Est-ce que l'ajustement d'une loi normale à la distribution observée des tailles des conducteurs peut être satisfaisant pour $\alpha = 0{,}01$? On donne $\overline{x} = 170{,}6875$ et $s = 7{,}3418$

®®®®®®®®®®®®®®®®®®®®®®®®

BIBLIOGRAPHIE

1- J. Genet, G. Pupion et M. Repussard : Probabilités , Statistiques et Sondages. Cours moderne et exercices corrigés, Vuibert (1974).

2- G. Baillargeon : Méthodes statistiques de l'ingenieur, 3ème édition, Volume 1, Les éditions SMG (1990).

3- J. Bass : Eléments de calcul des probabilités, 3ème édition, Masson et Cie éditions (1974).

4- K. khaldi : Méthodes Statistiques et probabilités ; Edition Casbah (2009).

Cette bibliographie ne se veut pas représentative des centaines de livres couvrant le domaine des probabilités et statistiques. Il s'agit simplement de la liste des documents qui ont été utilisés lors de la réalisation de ce polycopié.