Rappels d'analyse combinatoire

1°) Permutations

Le nombre de bijections qu'il est possible de faire pour un ensemble de n éléments est :
$P_n = n!$

Exemples

- $P_3 = 3! = 3 \cdot 2 \cdot 1 = 6$
- Répartition sur 6 postes de travail de 6 ouvriers : $P_6 = 6! = 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 720$

2°) Arrangements

C'est le nombre de suites de p objets qu'il est possible d'obtenir à partir de n objets en tenant compte de l'ordre. Si p=2, par exemple le couple (a,b) est différent du couple (b,a).

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

Propriétés : si n=p on doit retrouver le nombre de permutations :

$$A_n^n = \frac{n!}{(0)!} = n!$$

Cela est possible si l'on pose 0! = 1

Exemples

- Arrangement 2 parmi trois : $A_3^2 = \frac{3!}{(3-2)!} = 3 \cdot 2 = 6$
- Répartition sur 3 postes de travail de 6 ouvriers, 3 postes de travail de nature différente : $A_6^3 = \frac{6!}{(6-3)!} = 6 \cdot 5 \cdot 4 = 120$

3°) Combinaisons

C'est le nombre de possibilités pour n objets d'être regroupés par p sans tenir compte de l'ordre. Il y a p! permutations possibles sur les arrangements et donc

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

Exemples

- Combinaison 2 parmi trois : $C_3^2 = \frac{3!}{(3-2)!2!} = \frac{3 \cdot 2}{2} = 3$
- Répartition sur 3 postes de travail de 6 ouvriers, 3 postes de travail de identiques : $C_6^3 = \frac{6!}{(6-3)! \cdot 3!} = \frac{6 \cdot 5 \cdot 4}{3 \cdot 2} = 20$

Propriétés importantes

- $C_n^p = C_n^{n-p}$

  en effet $C_n^{n-p} = \dfrac{n!}{(n-p)![n-(n-p)]!} = \dfrac{n!}{(n-p)!\,p!} = C_n^p$
- $C_n^p = C_{n-1}^{p-1} + C_{n-1}^p$

  en effet : $C_{n-1}^{p-1} = \dfrac{(n-1)!}{(p-1)![(n-1)-(p-1)]!} = \dfrac{(n-1)!}{(p-1)!(n-p)!}$ et

  $C_{n-1}^p = \dfrac{(n-1)!}{p![(n-1)-p]!}$

  Si l'on met ces deux expressions sur le même dénominateur p!(n-p)! il vient $C_{n-1}^{p-1} \dfrac{p\cdot(n-1)!}{p!(n-p)!}$ et $C_{n-1}^p \dfrac{(n-p)(n-1)!}{p!(n-p)!}$ et l'addition de ces deux termes donne le résultat.
- Triangle de Pascal
- Formule du binôme $(x+y)^n = C_n^n \cdot x^n + C_n^{n-1} \cdot x^{n-1} \cdot y + ... + C_n^0 \cdot y^n$

4°) Remplir un « registre »

				....			

n cases dans lesquelles on peut mettre m valeurs différentes, combien y a-t-il de possibilités pour remplir ?

Réponse : $(m)^n$

Notions de calcul des probabilités

I) Tribu d'événements

1°) Introduction

Lancer d'un dé. Les événements : obtenir 1, obtenir 2, ... obtenir 6 mais aussi obtenir un nombre pair, obtenir 2 ou 3 ...

L'ensemble $\Omega = \{1,2,3,4,5,6\}$ est l'univers

Tout événement appartient aux parties de $\Omega = P(\Omega)$

2°) Définition

Ω quelconque F est une tribu si

- $\Omega \in F$
- $\forall A \in F, \bar{A} = C_\Omega A \in F$
- $\forall (A_i)_{i\in I} \in F, A_i$ famille dénombrable de F, $\cup(A_i) \in F$

Propriétés

- $\emptyset \in F$
- $\forall A, B \in F, A \cap B \in F, A \cup B \in F, A \setminus B \in F$
- $\forall (A_i)_{i\in I} \in F, A_i$ famille dénombrable de F, $\cap(A_i) \in F$

Exemples de tribu

- tribu grossière : $F = \{\Omega, \emptyset\}$

- tribu de Bernouilli $F=\{\Omega, A, \bar{A}, \emptyset\}$
- tribu la plus fine $F=P(\Omega)$

3°) Opérations logiques sur les événements

- A et B correspond à $A\cap B$
  Exemple : une machine fabrique des pièces telles que $14{,}85\ mm \leqslant d \leqslant 15{,}15\ mm$ correspond aux événements A et B : A $14{,}85\ mm \leqslant d$ et B $d \leqslant 15{,}15\ mm$
- Evénement incompatibles : ne peuvent être réalisé dans la même épreuve simultanément : $A\cap B=\emptyset$
  Exemple : C ($14{,}85\ mm \leqslant d \leqslant 15{,}15\ mm$) et D($d > 15{,}05\ mm$) sont incompatibles.
- A ou B correspond à $A\cup B$
  Exemple : jet d'un dé A(X est pair) et B($X \geqslant 4$)
  $A\cup B=\{2{,}4{,}5{,}6\}$

II) Probabilité d'un événement

1°) Introduction

Jet de dé : $prob(\{1\})=\frac{1}{6}$ et $prob(\{2{,}3\})=\frac{1}{3}$

On peut construire une application notée prob de $P(\Omega)\rightarrow \mathbb{R}^{+¿}$

2°) Définitions

On appelle probabilité sur un espace probabilisable (Ω, F) toute application notée prop ou P ou p de F dans R⁺ vérifiant :

- $prob(\Omega)=1$
- $\forall (A_i)_{i\in I}\in F$, famille d'événements 2 à 2 disjoints $prob(\cup A_i)=\sum prob(A_i)$

L'espace $(\Omega, F, prob)$ est appelé espace probabilisé

3°) Propriétés

- $\forall A\in F, 0 \leqslant prob(A)\leqslant 1$
- $\forall A\in F, prob(\bar{A})=1-prob(A)$
- $prob(\emptyset)=0$
- $\forall A, B\in F, A\subset B \Rightarrow prob(A)\leqslant prob(B)$
- $\forall A, B\in F, A\cap B=\emptyset \Rightarrow prob(A\cup B)=prob(A)+prob(B)$
- $\forall A, B\in F, prob(A\cup B)=prob(A)+prob(B)-prob(A\cap B)$

4°) Probabilité

$$p=\frac{nb\ de\ cas\ favorable}{nd\ de\ cas\ possibles}$$

Exemple : probabilité d'obtenir 1 boule noire d'une urne contenant 3 noires et 2 blanches : p=3/5

Remarque : lancer d'un dé : la probabilité d'obtenir un 2 est 1/6 (p(2)=1/6)

On peut le lancer 10 fois sans obtenir de 2.

Fréquence : lors d'une répétition de n fois une certaine épreuve X est apparu nX fois, la fréquence relative sera alors $f_X = \frac{n_X}{n}$

Probabilité : c'est une limite théorique quand le nombre d'expériences augmente pour tendre vers l'infini :

$$p(X) = \lim_{n\to\infty} \frac{n_X}{n} = \lim_{n\to\infty} f_X$$

III) Probabilité conditionnelle

1°) Définition

B étant un événement de probabilité non nulle, on appelle probabilité conditionnelle de A sachant B le rapport :

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

On peut vérifier que cette fonction qui associe à tout événement A la proabilité de A si B est une probabilité.

2°) Indépendance

Deux événements A et B sont indépendants si $prob(A\cap B) = prob(A)\cdot prob(B)$
L'indépendance de deux événements revient à poser :
$prob(A/B) = p(A)$ ou $prob(B/A) = p(B)$

n événements sont dits mutuellement indépendants deux à deux si pour toute partie
$I \subset \{1, ..n\}$
$P(\cap A_i)_{i\in I} = \prod A_i$
Cette propriété entraîne l'indépendance 2 à 2 de tous les couples mais la réciproque est fausse.

Théorème de Bayes

- Pour tout couple A, B d'événements $prob(B/A) = \frac{prob(A/B)\cdot prob(B)}{prob(A)}$
- Si $B_1, \dots B_n$ forment un ensemble complet d'événements alors
  $prob(A) = \sum_j prob(A/B_j)\cdot prob(B_j)$ et donc :
  $$prob(B_i/A) = \frac{prob(A/B_i)\cdot prob(B_i)}{\sum_j prob(A/B_j)\cdot prob(B_j)}$$

Variable aléatoire (V.A.)

I) Variable aléatoire, loi de probabilité

Une grandeur est appelée variable aléatoire (V.A.) lorsqu' elle varie de façon imprévisible.
On définit :
D : Domaine de définition de la VA (numérique)
X : peut être continue ($D \subset \mathbb{R}$) ou discrète {pile,face} points d'un dé (lancé) {1,2,3,4,5,6}
L'ensemble des couples [x_i,P(x_i)] constitue la loi de probabilité.
En pratique, la donnée des couples [x_i,P(x_i)] ne se conçoit que pour une variable discrète.

1°) V.A. discrète

données sous forme de tableau ou de graphique :

X	-1	0	1
P	0,5	0,2	0,3

On définit la fonction de répartition :

$$F(x_i)=p(X \leqslant x_i)=\sum_{n=1}^{i} p_n$$

- elle est strictement croissante
- elle est en marches d'escaliers
- p(a<X<b) = F(b) – F(a)
- $F(+\infty)=1$

2°) V.A. continue

Elle est caractérisée par sa densité de probabilité f(x) :

f(x)dx = P(x < X < x+dx)

Cette fonction possède un certain nombre de propriétés :
- $f(x) \geqslant 0$
- $P(a<X<b)=\int_a^b f(x)\,dx$
- $\int_D f(x)\,dx=1$

On définit aussi une fonction de répartition :

$$F(x)=\int_{infD}^{x} f(t)\,dt=P(X \leqslant x)$$

Propriétés
- F(x) n'est jamais décroissante
- $F(x) \in [0,1]$
- F(inf(D)) = 0
- F(sup(D)) = 1

- P(a<x<b) = F(b) – F(a)
- si F(x) est dérivable $\frac{dF(x)}{dx}=f(x)$

3°) Changement de variable

Cas discret : commençons par un exemple

X	1	2	4
P	0,5	0,2	0,3

Tableau de répartition de Y= 1/(3-X)

Y	-1	0,5	1
P	0,3	0,5	0,2

Cas continu :

X possède une densité de probabilité f(x)
Y est une fonction de X : Y = η(X)
X est donc une fonction de Y : X = ψ(Y)
Trouver g(y), la densité de probabilité de Y ?
Invariance de probabilité s'écrit P(x<X<x+dx) = P(y<Y<y+dy)
ou encore f(x)dx = g(y)dy
Il vient $g(y)=f(x)\frac{dx}{dy}=f(\psi(y))\cdot\psi'(y)$
Or $dy=|\eta'(x)|\cdot dx$
Donc

$$g(y)=\frac{f(x)}{|\eta'(x)|}=\frac{f(\psi(y))}{|\eta'(\psi(y))|}$$

Exemples :
1) Y= η(X) = c+X avec c cste η'(X) = 1
X = ψ(Y) = Y – c
g(y) = f(y-c)/1
Il y a conservation de la densité de probabilité, celle-ci étant simplement translatée de c.
2) Y= η(X) = cX avec c cste η'(X) = c
X = ψ(Y) = Y/c
g(y) = f(y/c)/c
Il y a conservation de la densité de probabilité, celle-ci étant simplement divisée par c ainsi que l'abscisse.

II) Paramètres associés à une variable aléatoire

1°) Espérance mathématique

$E(X)=\mu=\int_D x\cdot f(x)\cdot dx$ dans le cas d'une variable aléatoire continue.

$E(X)=\mu=\sum_{i=1}^{n} x_i\, p(x_i)$ Dans le cas d'un variable aléatoire discrète.

Propriétés
- si X = c alors E(X) = c si c est une constante.
- E(cX) = c.E(X) si c est une constante.
- Si X_i V.A. quelconque : $E(X_1 \pm X_2 \pm X_3 \pm \ldots \pm X_n)=E(X_1)\pm E(X_2)\pm E(X_3)\pm \ldots \pm E(X_n)$
- si X_1 et X_2 sont indépendantes : $E(X_1 \cdot X_2)=E(X_1)\cdot E(X_2)$
- si Y=g(X) alors $E(Y)=\int_D g(x) f(x)\, dx$

2°) Variance

C'est un paramètre qui mesure la dispersion. Il est noté $V(X)=\sigma^2$
V(X)= E[(X-μ)²]

$V(X)=\sum_{i=1}^{n} (x_i-\mu)^2 \cdot p(x_i)$ dans le cas d'une variable aléatoire discrète.

$V(X)=\int_D (x_i-\mu)^2 \cdot f(x)\cdot dx$ dans le cas d'une variable aléatoire continue.

Propriétés
- si X = c alors V(X) = 0 si c est une constante.
- V(cX) = c^2.V(X) si c est une constante.
- V(c+X) = V(X) si c est une constante.
- V(X) = E(X^2)-μ^2. Propriété importante pour le calcul. On a évidemment :

  $E(X^2)=\int_D x^2 \cdot f(x)\cdot dx$ pour une variable aléatoire continue.

  $E(X^2)=\sum_{i=1}^{n} x_i^2 \cdot p(x_i)$ pour une variable aléatoire discrète.

$\sigma=\sqrt{V(X)}$ est l'écart-type
σ/μ est l'écart-type relatif
σ^2/μ^2 est la variance relative.

Lois de distribution des Variables Aléatoires (V.A.)

I) Loi binomiale (ou de Bernouilli)

On dit que la V.A. discrète est répartie suivant la loi binomiale si ses valeurs possibles sont 0, 1, ... m, ... n et les probabilités respectives valent

$P_m = P[X = m] = C_n^m \cdot p^m \cdot q^{n-m}$ avec q=1-p

On la notera B(n,p)

Propriétés

- μ = np
- σ^2 = npq
- $\sigma = \sqrt{npq}$
- quand n est grand et que ni p ni q sont proches de 0, la loi binomiale tend vers la loi normale. Ainsi $z = \frac{X - np}{\sqrt{npq}} \in N(0,1)$ Cette approximation est très bonne dès que np et nq sont tous deux plus grands que 5.

Exemple

V.A. X binomiale B(50, 0.2) on cherche $p(X \leqslant 6)$
Une table (non donnée ici) donnerait 0,1034
μ = 10 et $\sigma = \sqrt{npq}$ = 2,828
On transforme $p(X \leqslant 6)$ en $p(X < 6.5)$ (approximation du discret par le continu) donc
$z = \frac{6{,}5 - 10}{2.828} = -1{,}24$ donne p=0,107

II) Distribution de Poisson

$$P(X = m) = \frac{\lambda^m \cdot \mathrm{e}^{-\lambda}}{m!}$$

Le produit n.p donne le λ de la loi de poisson.

III)Répartition exponentielle

On dit qu'une variable aléatoire X obéit à la loi de répartition si $f(x) = \lambda \cdot e^{-\lambda \cdot x}$ où λ est le paramètre de la loi exponentielle. La fonction de répartition de la variable X à répartition exponentielle et de paramètre λ est :

$$F(x) = \left\{ \begin{array}{c} 0 \ \ pour\ x \leqslant 0 \\ 1 - e^{-\lambda \cdot x}\ pour\ x > 0 \end{array} \right\}$$

L'espérance mathématique est $E(X) = \frac{1}{\lambda}$ et la variance $V(X) = \frac{1}{\lambda^2}$ ainsi l'écart -type est :

$$\sigma(X)=\frac{1}{\lambda}$$

IV) Loi Normale (ou loi de Gauss)

1°) Définition

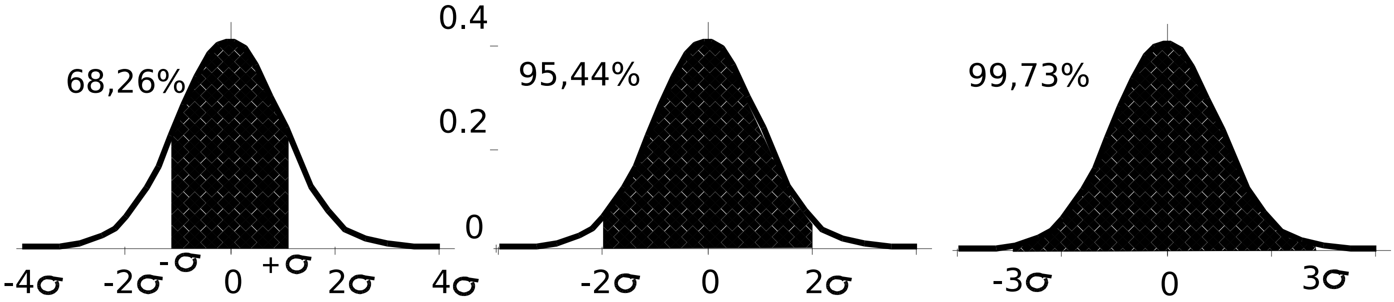
Une variable aléatoire est distribuée normalement si sa densité de probabilité s'écrit :

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

Sa fonction de répartition est donc :

$$F(x)=\frac{1}{\sigma\sqrt{2\cdot\pi}}\int_{-\infty}^{x} e^{\frac{-(t-\mu)^2}{2\sigma^2}}\,dt$$

Cette fonction est connue sous le nom de fonction d'erreur et notée souvent erf. Cette fonction est tabulée mais nous donnons quelques unes de ses propriétés :



99% se trouve dans $\mu \pm 2{,}58\ \sigma$

Dans le cas général nous noterons une distribution gaussienne par :

N(μ,σ ²) ou même N(μ,σ). N(0,1) est une loi de répartition centrée normale.

2°) Propriétés

Soit N(μ,σ) une loi de répartition sa moyenne est alors μ et son écart type σ .

A partir d'une distribution (répartition), pour trouver une distribution centrée normale, on pose $y'=\frac{y-\mu}{\sigma}$ -μ enlève la valeur moyenne (centre) et /σ normalise la distribution.

Soit $X\in N(\mu,\sigma)$ alors si Y=aX+b $Y\in N(a\cdot\mu+b,\sigma\cdot\sqrt{a})$

3°) utilisation de la table

On trouve en général des tables donnant non pas la fonction de répartition, mais en fait

pour x>0 $F(x)=\frac{1}{\sigma\sqrt{2\cdot\pi}}\int_0^x e^{\frac{-(t-\mu)^2}{2\sigma^2}}\,dt$ La fonction de partition est facilement trouvée à l'aide de symétries et en remarquent que l'aire complète à gauche de l'axe des y est 0,5.
(Réalisée avec le tableur OooCalc et formule : =LOI.NORMALE(A41+K1;0;1;1)–0,5)

	0	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2257	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2517	0,2549
0,7	0,2580	0,2611	0,2642	0,2673	0,2704	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2995	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1,2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4306	0,4319
1,5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441
1,6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1,7	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633
1,8	0,4641	0,4649	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4699	0,4706
1,9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4761	0,4767
2	0,4772	0,4778	0,4783	0,4788	0,4793	0,4798	0,4803	0,4808	0,4812	0,4817
2,1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
2,2	0,4861	0,4864	0,4868	0,4871	0,4875	0,4878	0,4881	0,4884	0,4887	0,4890
2,3	0,4893	0,4896	0,4898	0,4901	0,4904	0,4906	0,4909	0,4911	0,4913	0,4916
2,4	0,4918	0,4920	0,4922	0,4925	0,4927	0,4929	0,4975	0,4932	0,4934	0,4936
2,5	0,4938	0,4940	0,4941	0,4943	0,4945	0,4946	0,4948	0,4949	0,4951	0,4952
2,6	0,4953	0,4955	0,4956	0,4957	0,4959	0,4960	0,4961	0,4962	0,4963	0,4964
2,7	0,4965	0,4966	0,4967	0,4968	0,4969	0,4970	0,4971	0,4972	0,4973	0,4974
2,8	0,4974	0,4975	0,4976	0,4977	0,4977	0,4978	0,4979	0,4979	0,4980	0,4981
2,9	0,4981	0,4982	0,4982	0,4983	0,4984	0,4984	0,4985	0,4985	0,4986	0,4986
3	0,4987	0,4987	0,4987	0,4988	0,4988	0,4989	0,4989	0,4989	0,4990	0,4990
3,1	0,4990	0,4991	0,4991	0,4991	0,4992	0,4992	0,4992	0,4992	0,4993	0,4993
3,2	0,4993	0,4993	0,4994	0,4994	0,4994	0,4994	0,4994	0,4995	0,4995	0,4995
3,3	0,4995	0,4995	0,4995	0,4996	0,4996	0,4996	0,4996	0,4996	0,4996	0,4997
3,4	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4998
3,5	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998
3,6	0,4998	0,4998	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3,7	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3,8	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3,9	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000

4°) Théorème central limite

Il exprime le fait que quand le nombre d'expérience augmente les lois tendent vers une distribution gaussienne.

Estimation

I) Le problème de l'échantillonnage

La théorie de l'échantillonnage consiste à déterminer des propriétés sur des échantillons prélevés dans une population dont on connaît déjà des propriétés. On ne considère ici que des échantillons aléatoires, c'est à dire constitués d'éléments pris au hasard dans une population. Le tirage des éléments d'un échantillon peut être fait sans remise; On dit qu'il est exhaustif. Sinon si le tirage est fait avec remise, on dit qu'il est non exhaustif ; dans ce cas les tirages sont indépendants. Dans la plupart des cas, la population ayant un grand effectif, dans laquelle on tire une faible proportion d'éléments, on assimile un tirage sans remise à un tirage avec remise.

1°) Distribution d'échantillonnage des moyennes

Considérons une population ayant une certaine propriété avec une moyenne μ et un écart-type σ. Soit X la variable aléatoire qui à tout échantillon aléatoire prélevé avec remise et d'effectif N fixé, associe la moyenne de cet échantillon. Pour N suffisamment grand, X suit approximativement la loi normale N(μ, $\sigma/\sqrt{N}$).

$\hat{X}_N=\overline{X_N}=\frac{\sum_{i=1}^{N} X_i}{N}$ où les X_i sont des échantillons est appelé estimateur. Il s'agit bien d'une variable aléatoire car si vous faites plusieurs fois l'expérience d'échantillonnage vous trouverez un résultat différent. Comme toute variable aléatoire elle a une espérance mathématique et une variance et le point important est que sa variance est la variance de la population notée σ^2 divisée par N.

2°) Propriété des estimateurs

<u>Convergence</u> : un estimateur est dit correct (ou convergent) s'il converge en probabilité vers le paramètre lorsque n tend vers l'infini.

$$\forall \epsilon \forall \eta, \exists N \in \mathbb{N} : \forall n > N \qquad Prob\{|\hat{\theta}_n-\theta|<\epsilon\}>1-\eta$$

<u>Biais</u> : un estimateur $\hat{\theta}_N$ est sans biais lorsque son espérance mathématique est égale à la valeur θ du paramètre quel que soit N :

$$\forall N \in \mathbb{N} \qquad E[\hat{\theta}_N]=\theta$$

<u>Efficacité</u> : un estimateur sans biais $\hat{\theta}_N$ est efficace si la variance est la plus faible parmi les variances des autres estimateurs sans biais.

Si l'on reprend le seul estimateur que l'on a défini jusqu'à présent $\hat{X}_N=\overline{X_N}=\frac{\sum_{i=1}^{N} X_i}{N}$ il converge en probabilité vers E[X]. Cette propriété justifie la coutume intuitive d'identifier en pratique l'espérance mathématique E[X] d'une variable aléatoire à la moyenne d'un échantillon

$$\overline{X_N}=\frac{\sum_{i=1}^{N} X_i}{N}$$

3°) Estimateur de la variance

On démontre que la quantité que l'on doit prendre est :

$$\hat{\sigma}^2=\frac{\sum_{i=1}^{N}(x_i-\bar{X}_N)^2}{N-1}$$ pour avoir une estimation sans biais.

II) Estimation par intervalle de confiance

Les estimations ponctuelles sont hélas liées au choix de l'échantillon ; il faut donc rechercher un nouveau type d'estimation de la moyenne d'une population ou d'un pourcentage. On cherche des intervalles qui, généralement, à 95% ou 99% des cas, contiennent la moyenne m inconnue ou le pourcentage p d'une certaine propriété que possède la population.

1°) Théorème limite central

Soit $\{X_i\}$ i=1,2, ... une suite de variables aléatoires indépendantes de même loi, de valeur moyenne μ et d'écart-type σ. La suite $\{Y_n\}$ i=1,2, ... de variables aléatoires définie par :

$$Y_N=\frac{\sum_{i=1}^{N} X_i-N\cdot m}{\sigma\cdot\sqrt{N}}$$ converge en loi lorsque n augmente indéfiniment, vers une variable aléatoire qui suit la loi normale centrée réduite N(0,1).

Remarque : On peut écrire Y_N sous la forme :

$$Y_N=\frac{\sum_{i=1}^{N} X_i-N\cdot m}{\sigma\cdot\sqrt{N}}=\frac{\frac{1}{N}\sum_{i=1}^{n} X_i-m}{\frac{\sigma}{\sqrt{N}}}=\frac{\overline{X_N}-m}{\frac{\sigma}{\sqrt{N}}}$$

2°) De la moyenne

a) 1er cas

Soit P la population : μ la moyenne est inconnue σ l'écart-type est connu. Soit un échantillon : la moyenne $\overline{X_N}$ est connue, N l'effectif est connu. On se place dans le cas où l'on peut considérer que la variable aléatoire X, qui à tout échantillon de taille N fixée, associe la moyenne de cet échantillon, suit une loi normale N(μ, $\sigma/\sqrt{N}$).

Alors l'intervalle de confiance de la moyenne μ de la population, avec le coefficient de confiance 1-α, lu dans la table de la loi normale centrée réduite

N(0 ; 1) est : $I_\alpha=\left\{\overline{X_n}-u_\alpha\frac{\sigma}{\sqrt{n}};\overline{X_n}+u_\alpha\frac{\sigma}{\sqrt{n}}\right\}$ en notant u la loi normale centrée réduite.

Cette méthode conduit dans 100(1 - α) cas sur 100, pourcentage choisi à l'avance, à un intervalle de confiance contenant μ :

$$Prob(\overline{X_n} - u_\alpha \frac{\sigma}{\sqrt{n}} < \mu < \overline{X_n} + u_\alpha \frac{\sigma}{\sqrt{n}}) = 1 - \alpha$$

Les cas usuels les plus fréquent sont :
· coefficient de confiance 95% alors $u_{5\%}$ = 1,96
· coefficient de confiance 99% alors $u_{1\%}$ = 2,58.

b) <u>2ème cas</u>
Soit P la population : μ la moyenne est inconnue σ l'écart-type est inconnu. Soit un échantillon : la moyenne est estimée par $\overline{X_N}$ l'écart-type est estimé lui aussi par

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{N} (x_i - \bar{x}_N)^2}{N-1}$$

et N l'effectif est connu et il est inférieur strictement à 30. On se place dans le cas ou l'on peut considérer que la variable aléatoire X , qui à tout échantillon de taille N fixée, N < 30, associe la moyenne de cet échantillon, suit une loi de Student à N-1 degrés de liberté . Alors l'intervalle de confiance de la moyenne μ de la population, avec le coefficient de confiance 1 - α, lu dans la table de la loi de Student à N - 1 degrés de liberté est :

$$I_\alpha = \left\{ \overline{X_N} - t_{1-\alpha/2}^{(N-1)} \frac{\hat{\sigma}}{\sqrt{N}} ; \overline{X_N} + t_{1-\alpha/2}^{(N-1)} \frac{\hat{\sigma}}{\sqrt{N}} \right\}$$

Une autre manière de présenter les choses est de dire que la probabilité de trouver la moyenne vraie dans l'intervalle de confiance que l'on vient de définir est 1 - α.

$$Prob(\overline{X_N} - t_{1-\alpha/2}^{(N-1)} \frac{\hat{\sigma}}{\sqrt{N}} < \mu < \overline{X_N} + t_{1-\alpha/2}^{(N-1)} \frac{\hat{\sigma}}{\sqrt{N}}) = 1 - \alpha$$

<u>3°) Intervalle de confiance d'une variance</u>

L'estimation de la variance par $\hat{\sigma}^2 = \frac{\sum_{i=1}^{N} (x_i - \bar{x}_N)^2}{N-1}$ donne aussi une variable aléatoire qui possède elle-même une variance et une moyenne. Pour caractérisée cette estimation par rapport à la valeur vraie σ^2 inconnue, on s'intéresse au rapport $\frac{\hat{\sigma}^2}{\sigma^2}$ Ce rapport peut être exprimé comme : $\frac{\hat{\sigma}^2}{\sigma^2} = \frac{\sum_{i=1}^{N} (x_i - \bar{x}_N)^2}{(N-1) \cdot \sigma^2}$ On montre que le rapport

$\frac{(N-1) \cdot \hat{\sigma}^2}{\sigma^2} = \chi^2(N-1)$ suit une loi du qui2. Il vient alors :

$\hat{\sigma}^2 = \frac{\sigma^2}{(N-1)} \cdot \chi^2(N-1)$ ou

$$\sigma^2 = \frac{\hat{\sigma}^2 \cdot (N-1)}{\chi^2(N-1)}$$

Ainsi l'intervalle de confiance peut être estimé par avec une confiance 1 - α par :

$$I_{\alpha}=\left\{\frac{\hat{\sigma}^{2}\cdot(N-1)}{\chi^{2}_{1-\alpha/2}(N-1)};\frac{\hat{\sigma}^{2}\cdot(N-1)}{\chi^{2}_{\alpha/2}(N-1)}\right\}$$