Chapitre 1 LES STATISTIQUES DESCRIPTIVES

A. Statistiques a une variable

1. Vocabulaire de la statistique

Un ensemble d'objets ou de personnes d'une étude statistique est appelé **population**. Un élément de cette population est appelé **individu**.
L'étude statistique porte sur un caractère. Si le caractère est **quantitatif**, les mesures sont alors les valeurs d'une **variable statistique** (ex : un âge, une taille...). Si le caractère est **qualitatif**, on est obligé de le quantifier (ex : sexe...).
La variable est dite **discrète** si elle ne prend que des valeurs isolées (ex : entières). Elle est **continue** si elle peut prendre toutes la valeurs d'un intervalle (ex : **R**).
L'**effectif** d'une population est le nombre d'individus total de cette population. La **fréquence** d'un caractère est le nombre d'individus possédant ce caractère divisé par l'effectif total de la population.

2. Les variables discrètes

a) Représentation

On représente les variables aléatoires discrètes sous forme d'histogramme ou de camembert grâce aux différentes fréquences

b) Caractéristiques

(1) La moyenne

Soit n valeurs distinctes ou non de la variable. Si cette variable prend p valeurs distinctes (p n), $x_1,\ldots,x_p$, d'effectifs respectifs $n_1,\ldots n_p$ alors la moyenne est donnée par :

$$\overline{x}=\frac{1}{n}\sum_{i=1}^{p} n_i x_i$$

(2) Propriété

si pour tout i, on peut opérer un changement de variable affine du type : $x_i = aX_i + b$
alors $\overline{x} = a\overline{X} + b$.

(3) La variance

Elle est donnée par la formule : : $V=\frac{1}{n}\sum_{i=1}^{p} n_i (x_i - \overline{x})^2$.

(4) L'écart-type

Il est donné par la formule : $\sigma = \sqrt{V}$

(5) Propriétés

La formule suivante est plus pratique pour le calcul de V : $V = \frac{1}{n}\sum_{i=1}^{p} n_i x_i^2 - \bar{x}^2$

De plus si pour tout i, $x_i = aX_i + b$ alors $V_x = a^2 V_X$ et $\sigma_x = a\sigma_X$.

3. Les variables continues

a) *Représentation*

Pour leur représentation, on regroupe en général dans des classes adjacentes d'amplitudes pas forcément égales. Ceci est représenté dans le tableau ci-dessous :

classes	$[X_0 ; X_1[$	$[X_1 ; X_2[$	…………………	$[X_{p-1} ; X_p]$
centre des classes	x_1	x_2		x_p
effectifs	n_1	n_2		n_p
fréquences	n_1/n	n_2/n		n_p/n

La représentation s'effectue alors grâce à un histogramme dont les rectangles sont de largeur l'amplitude de la classe et dont **l'aire est proportionnelle à l'effectif.**

b) *Caractéristiques*

Pour calculer moyenne et écart-type, on prend les formules connues avec les x_i centres des classes c'est à dire $x_i = \frac{X_{i-1} + X_i}{2}$

B. Statistiques a deux variables

1. Tableau de données. Nuage de points.

On observe que dans certains cas, il semble exister un lien entre 2 caractères d'une population (ex : entre poids et taille, entre l'épaisseur d'un mur et sa résistance thermique...).
On définit alors une série statistique à 2 **variables** x et y, prenant des valeurs $x_1,\dots,x_n$ et $y_1,\dots,y_n$.

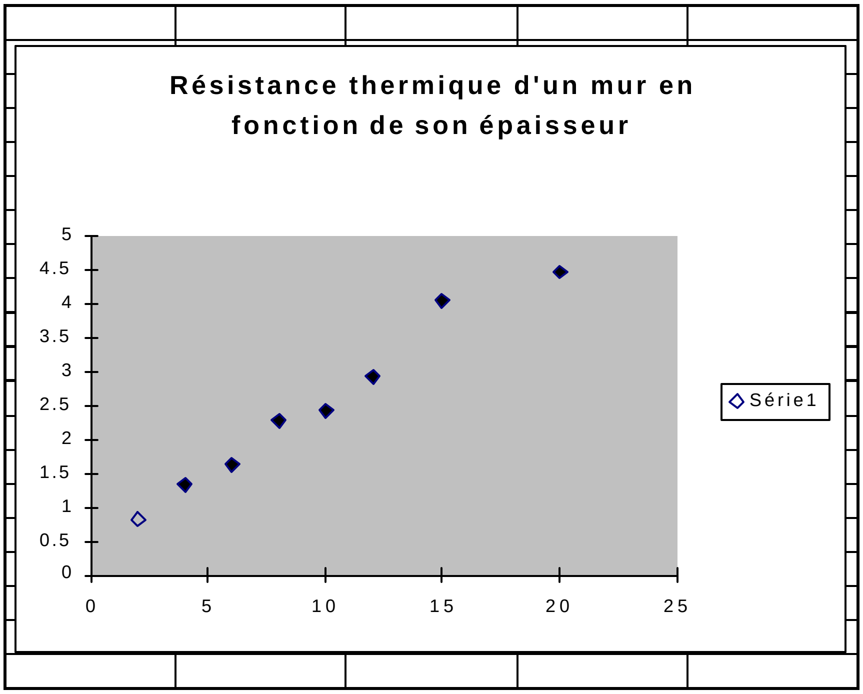
a) *Tableau de données.*

Le nom est explicite. On représente les différentes valeurs de x et y dans un tableau à deux entrées.

x en mm	2	4	6	8	10	12	15	20
y en m^2.°C	0,83	1,34	1,63	2,29	2,44	2,93	4,06	4,48

b) *Nuage de points*

Le plan P étant muni d'un repère **orthogonal**, on peut associer au couple $(x_i ; y_i)$ de la série statistique double, le point M_i de coordonnées x_i et y_i. L'ensemble des points M_i obtenus constitue le nuage de points représentant la série statistique.



c) Point moyen

On appelle point moyen d'un nuage de n points M_i de coordonnées (x_i ; y_i), le point G de coordonnées

$$x_G = \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$y_G = \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$$

C. Ajustement affine

1. Méthode graphique

a) Ajustement à la règle

On trace au jugé une droite D passant par plus près possible des points du nuage de points, en s'efforçant d'équilibrer le nombre de points situés au dessus et au dessous de la droite D. L'équation de D est alors de la forme y = ax + b. Pour retrouver cette équation, il suffit alors de connaître 2 points de D.

b) Ajustement affine par la méthode de Mayer

On partage le nuage de points en deux nuages de points de nombres équivalents. On calcule alors le point moyen de chaque nuage qu'on appelle G_1 et G_2. La droite (G_1G_2) est la droite de Mayer. Elle passe de plus par le point G. C'est une bonne approximation, si le nuage de points est allongé.

Ex : $2 \leq x_i \leq 8$ et $10 \leq x_i \leq 20$ alors on obtient y = 0,21x + 0,47.

2. Ajustement affine par la méthode des moindres carrés

a) La droite de régression

Soit D une droite d'ajustement. Soit $M_i(x_i\ ;\ y_i)$ un point du nuage. P_i est le point de même abscisse x_i que M_i situé sur la droite D d'équation y = ax + b. Q_i est le point de même ordonnée y_i que M_i situé sur la droite D' d'équation x = a'x + b'.

On appelle **droite de régression de y en x**, la droite D tell que : $\sum_{i=1}^{n} M_iP_i^2 = \sum_{i=1}^{n}\left[y_i - (ax_i + b)\right]^2$ soit minimale.

On appelle **droite de régression de x en y**, la droite D' telle que : $\sum_{i=1}^{n} M_iQ_i^2 = \sum_{i=1}^{n}\left[x_i - (a'y_i + b')\right]^2$ soit minimale.

b) Covariance d'une série statistique double

C'est le nombre $\text{cov}(x,y) = \sigma_{xy} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})$. Pratiquement on utilise plutôt la formule :

$$\sigma_{xy} = \frac{1}{n}\sum_{i=1}^{n} x_i y_i - \bar{x}\bar{y}$$

c) Equations des droites de régression

On montre que la droite de régression D de y en x a pour équation y = ax + b avec $a = \frac{\sigma_{xy}}{\sigma_x^2}$ et b vérifiant $b = \bar{y} - a\bar{x}$.

De même on montre que la droite de régression D' de x en y a pour équation x = a'y + b' avec $a' = \frac{\sigma_{xy}}{\sigma_y^2}$ et b' vérifiant $b' = \bar{x} - a'\bar{y}$.

Les droites D et D' passent toutes les deux par le point moyen G.

3. Coefficient de corrélation

a) Définition

Le coefficient de corrélation d'une série statistique double est le nombre r défini par $r = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$

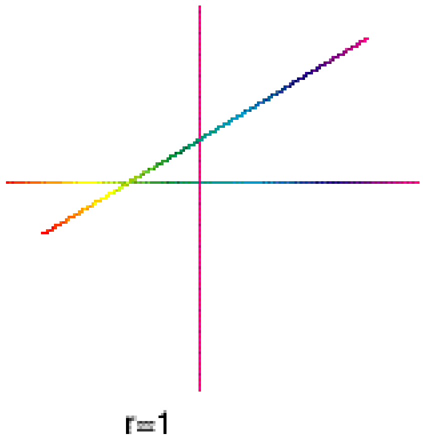
b) Propriétés

r, a, b, σ_{xy}, a' et b' sont tous de même signe.

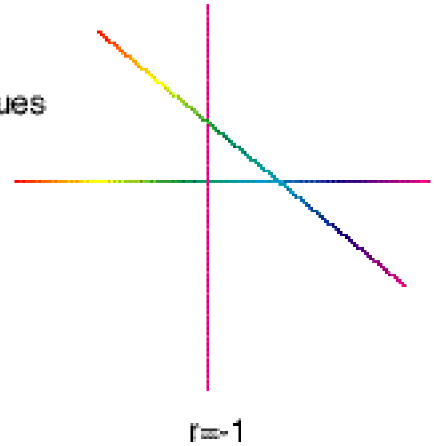
-1 r 1

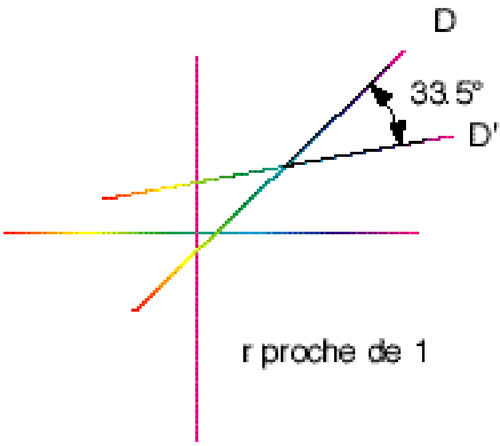
c) *Interprétation graphique*

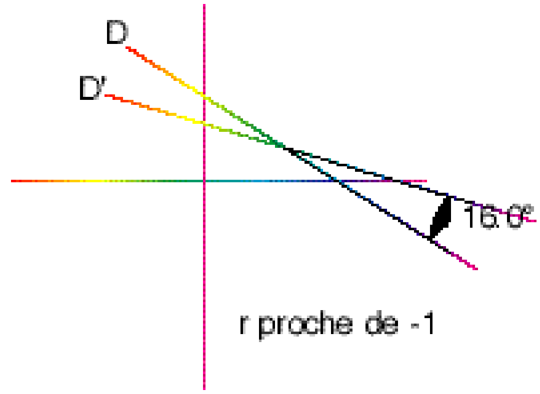
si $r^2 = 1$ alors $aa' = 1$. Les droites D et D' sont alors confondues ; on dit que l'ajustement affine est parfait.
si 0,7 | r | <1 alors les deux droites D et D' sont proches l'une de l'autre (en fait l'angle entre les deux est inférieur à 45°) ; on dit que l'ajustement affine est justifié.
si | r | <0,7 alors l'angle entre les deux droites est supérieur à 45°. L'ajustement affine ne se justifie pas.

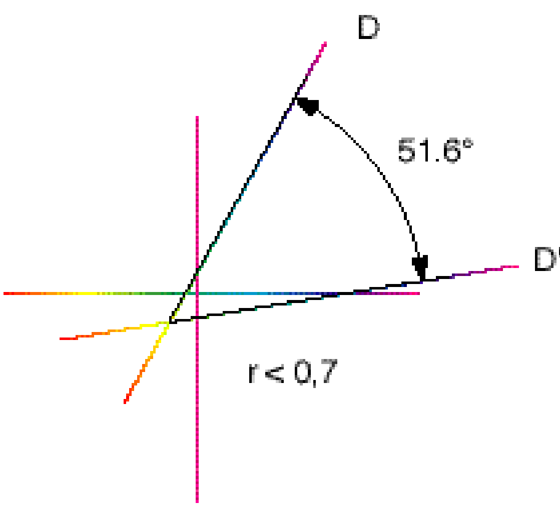


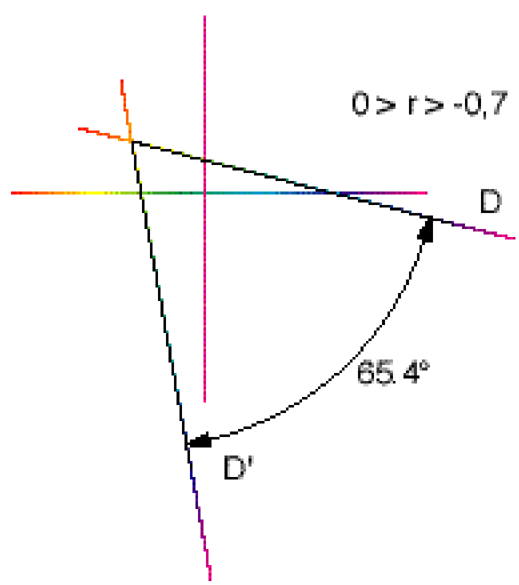
D et D' confondues











Chapitre 2 PROBABILITES ET ANALYSE COMBINATOIRE

A. Notion d'expérience aléatoire

1. Définition

Une expérience ayant un nombre fini d'issues possibles est appelé **expérience aléatoire** s'il est impossible de savoir à l'avance quelle en sera l'issue. L'ensemble de toutes les issues possibles est appelé l'**univers** des possibles associé à cette expérience; Il est généralement noté Ω. Chaque sous ensemble de Ω contenant un seul élément, c'est à dire chaque issue possible est appelé événement élémentaire.

2. Remarque

Si on note chaque issues possibles $e_1,\ldots,e_n$ alors $\Omega=\{e_1,\ldots,e_n\}$ et chaque $\{e_i\}$ est alors un événement élémentaire.
Une expérience aléatoire est déterminée par l'expérience que l'on effectue et donc l'univers aussi, c'est à dire que si on change d'expérience aléatoire, on change aussi d'univers !

B. Vocabulaire des événements

1. Définition

Soit E une expérience aléatoire et Ω l'univers des possibles associé à cette expérience. L'ensemble de toutes les parties de Ω, $P(\Omega)$ est l'ensemble des événements lié à Ω.
Ω est l'événement certain. $\varnothing$ est l'événement impossible.

2. Composition d'événements

a) Evénement A È B

La loi $\cup$ dans $P(\Omega)$ correspond à l'emploi du « ou inclusif » entre deux événements.

b) Evénement A Ç B

La loi $\cap$ dans $P(\Omega)$ correspond à l'emploi du « et » entre deux événements.
Dans le cas où $A\cap B=\varnothing$, on dit que les deux événements A et B sont disjoints ou incompatibles.

c) Evénement contraire

Soit A un événement lié à une expérience aléatoire E d'univers associé Ω. A est donc une partie de Ω. Ainsi $C_A=\{e_i\in\Omega / e_i\notin A\}$ est associé à l'événement qui n'est réalisé que si A ne l'est pas. On l'appelle complémentaire de A ou « non A ». Il est noté $\overline{A}$.
On a alors : $\overline{A\cup B}=\overline{A}\cap\overline{B}$ et $\overline{A\cap B}=\overline{A}\cup\overline{B}$ et $\overline{(n>p)}=(n\le p)$

C. Axiomatique du calcul des probabilités

1. Axiomes du calcul des probabilités

Soit E une expérience aléatoire et Ω son univers associé. On appelle probabilité, notée p, toute application de l'ensemble des événements $P(\Omega)$ dans $\mathbb{R}$ vérifiant les trois axiomes suivants :

A1 : $\forall A \in P(\Omega) \quad 0 \leq p(A) \leq 1$

A2 : $p(\Omega)=1$

A3 : si deux événements A et B sont incompatibles alors $p\left(A \cup B\right) = p(A) + p(B)$

2. Conséquences

La probabilité d'un événement $A=\{e_1,\ldots,e_n\}$ est telle que $p(A)=p(e_1)+\ldots+p(e_n)$.

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

$$p(\emptyset) = 0$$

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

3. Cas particulier important : l'équiprobabilité

L'équiprobabilité correspond au cas où tous les événements élémentaires ont la même probabilité. S'il y a n événements élémentaires chacun possède une probabilité de 1/n d'apparaître. Dans ce cas, on peut écrire la formule suivante :

$$p(A) = \frac{\text{nbre d'éléments de } A}{\text{nbre d'éléments de } \Omega} = \frac{\text{nbre de cas favorables}}{\text{nbre de cas possibles}}$$

D. Probabilité conditionnelle

1. Définition

Soit p une probabilité sur un univers Ω et soit A un événement de probabilité non nulle. La probabilité que l'événement B soit réalisé sachant que A l'est déjà est défini par $p(B/A) = \dfrac{p(A \cap B)}{p(A)}$. On l'appelle probabilité conditionnelle.

2. Propriétés

$$p(\Omega / A) = 1$$

$$p(B \cup C / A) = p(B / A) + p(C / A)$$

$$p(A \cap B) = p(A / B)p(B) = p(B / A)p(A)$$

Cette formule est appelée formule des probabilités composées.

3. Exemple

deux machines M1 et M2 fabriquent des tiges. Elles produisent respectivement 1/3 et 2/3 de la production. La machine M1 sort 5% de tiges défectueuses et M2 en sort 6%.

Soit les événements A : « la tige est fabriquée par M1 » B : « la tige est fabriquée par M2 »
D : « la tige est défectueuse ».
1) Quelle est la probabilité que la tige soit fabriquée par M1 ? c'est p(A)=1/3.
2) On tire une tige de la production de M1. Quelles est la probabilité qu'elle soit défectueuse?
C'est p(D/A)=5/100.
3) On tire une tige de la production. Quelle est la probabilité pour qu'elle provienne de M1 et qu'elle soit défectueuse ? C'est $P(A \cap D) = p(D/A)p(A) = \frac{1}{60}$.
4) On tire une tige de la production. Quelle est la probabilité pour qu'elle soit défectueuse ?
C'est $p(D) = p((A \cap D) \cup (B \cap D)) = p(D/A)p(A) + p(D/B)p(B) = \frac{17}{300}$.
5) Quelle est la probabilité qu'une pièce défectueuse ait été fabriquée par M1 ?
C'est $P(A/D) = \frac{p(A \cap D)}{p(D)} = \frac{1}{60} \times \frac{300}{17} = \frac{5}{17}$.

E. Evénements indépendants

1. Définition

Soit Ω l'univers des possibles d'une expérience aléatoire E et p une probabilité associée. Deux événements sont dits indépendants relativement à p si : $p(A \cap B)=p(A)p(B)$

2. Remarque

Il ne faut pas confondre indépendant et incompatible
Si deux événements sont indépendants alors $p(A \cap B)=p(A)p(B)$ donc $p(B/A)=p(B)$ et $p(A/B)=p(A)$; cela signifie que deux événements sont indépendants si et seulement si la réalisation de l'un n'influe pas sur la réalisation de l'autre.

F. Eléments d'analyse combinatoire

1. Les p-listes

Elles correspondent à un tirage successif et avec remise c' est à dire que les répétitions sont possibles et que l'ordre est important.
Le nombre de p-listes d'un ensemble E à n éléments est n^p.

2. Les suites de p éléments distincts

Elles correspondent à un tirage successif sans remise, c'est à dire que les répétitions sont impossibles et l'ordre est important.

a) Les permutations

Toute suite de n éléments distincts choisis parmi les n éléments d'un ensemble E est appelé permutation de n éléments. Le nombre total de permutations d'un ensemble de n éléments est : n!.

b) <u>Les arrangements</u>

Toute suites de p éléments distincts choisis parmi n éléments distincts (n p) est appelé arrangement de p éléments parmi n. Le nombre total d'arrangements de p éléments parmi n est :

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

3. Nombre de parties à p éléments d'un ensemble à n éléments (p n)

Il correspond à un tirage simultané c'est à dire que les répétitions sont non possibles et que l'ordre n'est pas important.

Toute partie de p éléments distincts choisis parmi n éléments (p n) est appelée combinaison de p éléments parmi n. Le nombre total de combinaisons d'un ensemble à n éléments est : $C_n^p = \dfrac{A_n^p}{p!}$.

4. Propriétés des arrangements et combinaisons

$A_n^p = \dfrac{n!}{(n-p)!}$ et $C_n^p = \dfrac{n!}{p!(n-p)!}$ formules évidentes à montrer à partir de la définition....

$C_n^p = C_n^{n-p}$ en effet il y a autant de parties de E à p éléments que de parties de E à n-p éléments..

$C_{n-1}^{p-1} + C_{n-1}^p = C_n^p$ cette formule peut se montrer à partir des formules de définition mais aussi par des considérations ensembliste : soit $a \in E$ et $E'=E\setminus\{a\}$ alors lors du choix d'un sous ensemble F à n éléments de E deux cas peuvent se produire à savoir $a \in F$ cela revient à choisir une partie à p-1 éléments de E' (C_{n-1}^{p-1} façon de la faire) ou $a \notin F$ cela revient à choisir une partie à p éléments de E' (C_{n-1}^p façon de le faire) on a donc :
$C_{n-1}^{p-1} + C_{n-1}^p = C_n^p$

Valeurs particulières :

$A_n^0 = 1$ $\quad C_n^0 = 1$

$A_n^1 = n$ $\quad C_n^1 = n$

$A_n^n = n!$ $\quad C_n^n = 1$

5. Triangle de Pascal - binôme de Newton :

<u>Triangle de pascal</u> : la construction est basée sur la propriété $C_{n-1}^{p-1} + C_{n-1}^p = C_n^p$.

n \ p	0	1	2	3	4	5
0	1					
1	1	1				
2	1	2	1			
3	1	3	3	1		
4	1	4	6	4	1	
5	1	5	10	10	5	1

<u>Binôme de Newton :</u> $\boxed{(a+b)^n = \sum_{p=0}^{p=n} C_n^p a^p b^{n-p}}$

Par exemple : $(a+b)^5 = a^5 + 5a^4b + 10a^3b^2 + 10a^2b^3 + 5ab^4 + b^5$

Remarque : dans les cas particulier où a=b=1 on obtient $\sum_{p=0}^{p=n} C_n^p = 2^n$

Chapitre 3 LES VARIABLES ALÉATOIRES

A. *Variables aléatoires discrètes sur un univers fini*

1. Convention d'écriture

Soit $\Omega = \{\omega_1,...,\omega_N\}$ un univers fini probabilisé. On appelle variable aléatoire, notée X, définie sur Ω toute application de Ω dans $\mathbb{R}$.
On note $X(\Omega) = \{x_1,\ldots,x_n\}$ ($n \in \mathbb{N}$) l'ensemble image de Ω par X.
$X = x_i$ est la partie de Ω formées de toutes les éventualités ω_k ayant pour image x_i. Il y en a n, formant une partition de Ω.
$X > x$ est la partie de Ω formée de toutes les éventualités dont le nombre image est supérieur strictement à x.
$x \leq X \leq y$ est la partie de Ω formée de toutes les éventualités dont le nombre image est compris entre x et y.

2. Loi de probabilité

Soit Ω un univers fini probabilisé st X une variable aléatoire sur Ω. On appelle loi de probabilité de X, l'application qui à chaque valeur image x_i fait correspondre la probabilité p_i de la partie $(X = x_i)$ de Ω.
On la représente alors sous forme d'un tableau :

$X = x_i$	x_1	x_2	…	x_i	…	x_n
$p(X=x_i)$	p_1	p_2	…	p_i	…	p_n

On a donc $p(\Omega) = \sum_{i=1}^{n} p_i = 1$.

3. Fonction de répartition

On appelle fonction de répartition de la variable aléatoire X, l'application F de $\mathbb{R}$ dans [0 ; 1] qui associe à tout réel x la probabilité $p(X \leq x)$ c'est à dire $F(x) = p(X \leq x)$.
Elle est croissante, continue par morceau et en escalier.
De plus on a : $p(X > x) = 1 - F(x)$ et $p(x \leq X \leq y) = F(y) - F(x)$.

4. Valeurs caractéristiques d'une variable aléatoire à valeurs discrètes

Soit X une variable aléatoire prenant les valeurs $\{x_1\ldots x_n\}$ avec les probabilités respectives $\{p_1\ldots p_n\}$.

a) Espérance

On appelle espérance mathématique de X le nombre $E(X) = \sum_{i=1}^{n} p_i x_i$. Ce nombre s'interprète comme la moyenne m des valeurs x_i pondérées par leur probabilité p_i.

b) Propriétés

Soit k une constante alors : $\begin{array}{l} E(k) = k \\ E(X+k) = E(X) + k \\ E(kX) = kE(X) \end{array}$

c) Variance et écart-type

On appelle variance de X le nombre $V(X) = \sum_{i=1}^{n} p_i x_i^2 - \left(E(X)\right)^2$. L'écart-type est : $\sigma(X) = \sqrt{V(X)}$

d) Propriétés

Soit k une constante alors : $\begin{array}{l} V(k) \\ V(X+k) = V(X) \\ V(kX) = k^2 V(X) \end{array}$

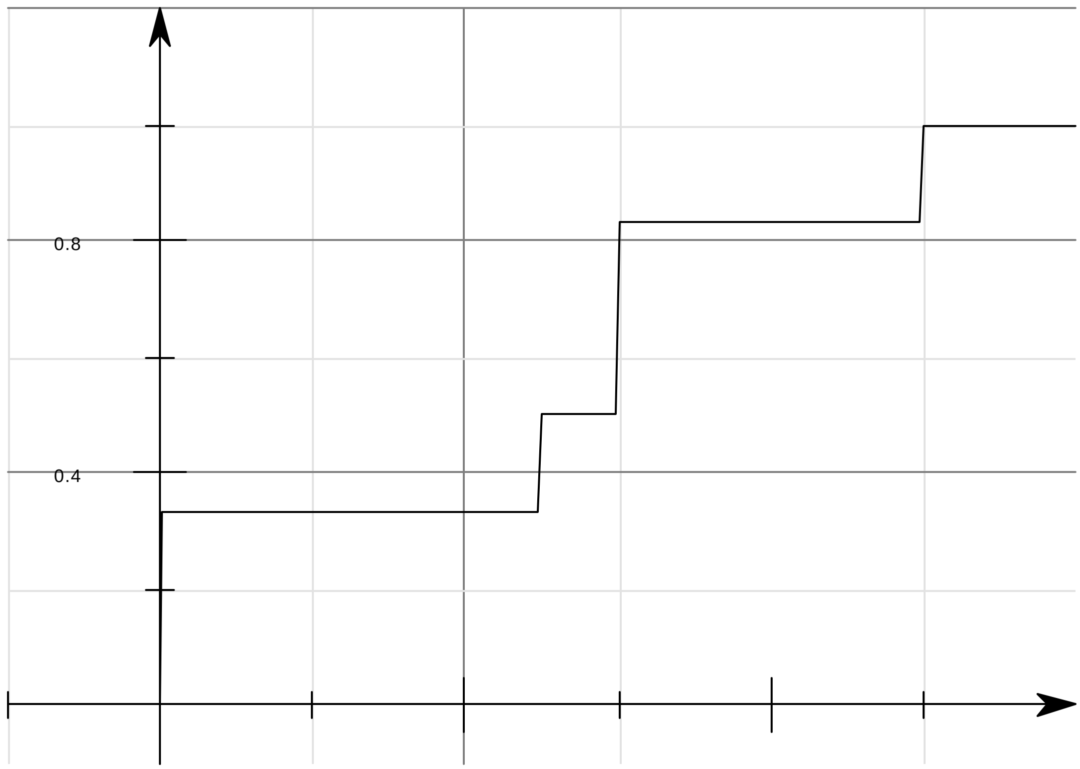
5. Exemple

Contre une mise convenable, on lance un dé marqué as, roi, dame, valet, dix et neuf.
L'as rapporte 10 F   Le roi et la dame 6 F   Le valet 5 F   le 10 et le 9 rien

Loi de probabilité

$X = x_i$	0	5	6	10
p_i	1/3	1/6	1/3	1/6

Fonction de répartition F



E(X)=4,5 V(X)=12,58 et σ(X)=3,55

6. <u>Loi binomiale</u>

Une variable aléatoire X à valeurs entières : 0,1 , 2 , …n suit une loi binomiale de paramètres n et p si et seulement si pour tout k appartenant à {1,…,n} on a : $p(X = k) = C_n^k p^k q^{n-k}$.
On l'utilise chaque fois qu'une même expérience a 2 éventualités. Elle est notée B(n,p).

Son espérance est alors E(X) = np, sa variance V(X) = npq.

B. <u>Variables aléatoires dénombrables sur un univers infini</u>

1. <u>Loi de Poisson</u>

Une variable aléatoire dénombrable(c'est à dire établissant une bijection avec $\mathbb{N}$) X suit une loi de Poisson de paramètre m (m > 0) sis et seulement si : $p(X = k) = e^{-m} \dfrac{m^k}{k!}$.
Son espérance est E(X)=m, sa variance V(X) = k.

C. <u>Variables aléatoires continues</u>

1. <u>Définition</u>

Une variable aléatoire continue est une variable aléatoire X dont l'ensemble des valeurs est $\mathbb{R}$ ou un intervalle I de $\mathbb{R}$.
Une telle variable est généralement définie par sa fonction de répartition $F: x \mapsto F(x) = p(X \leq x)$.

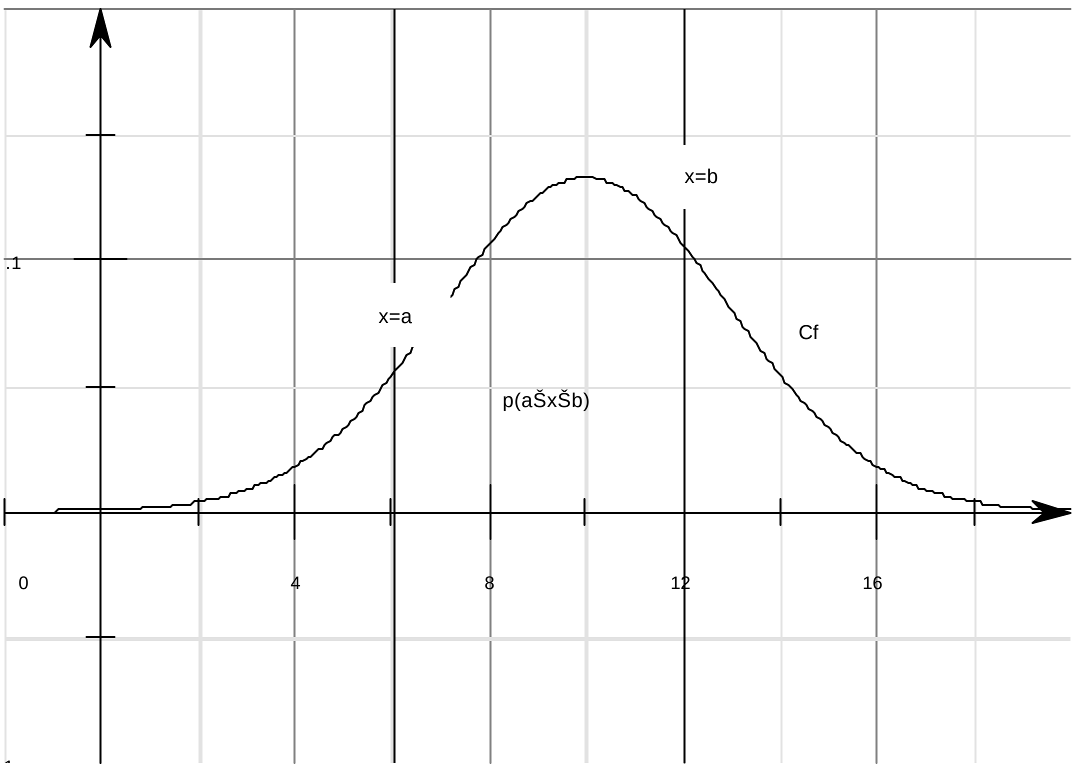
2. Fonction densité de probabilité

On désigne par fonction densité de probabilité, la fonction dérivée f de la fonction de répartition F. On alors : $\int f(x)dx = F(x)$ et $\int_a^b f(x)dx = F(b) - F(a) = p(a \leq X \leq b)$. c'est à dire que l'aire mesurée entre les droites d'équation x = a et x = b, la courbe représentative de f et l'axe des abscisses correspond à p(a x b).

$$\int_{-\infty}^{a} f(x)dx = F(a) = p(X \leq a)$$

On a $\int_{a}^{+\infty} f(x)dx = 1 - F(a) = p(X \geq a) = 1 - p(X < a)$

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



3. Valeurs caractéristiques

Soit X une variable aléatoire continue alors on a :

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

$$V(X) = \int_{-\infty}^{+\infty} f(x)\left[x - E(X)\right]^2 dx = E(X^2) - \left[E(X)\right]^2$$

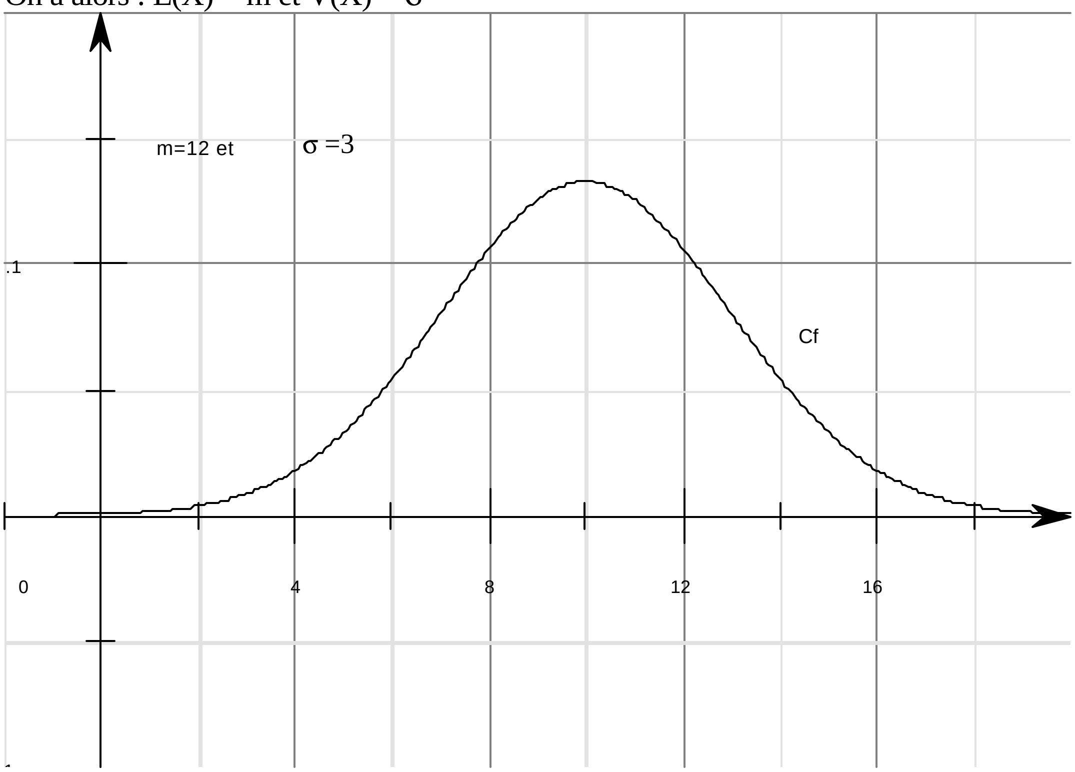
D. La loi normale ou loi de Laplace-Gauss

1. Définition

Une variable aléatoire X suit une loi normale $\mathcal{N}$ (m;σ) de paramètres m et σ lorsque sa densité de probabilité est la fonction f définie sur $\mathbb{R}$ par : $f(x) = \dfrac{1}{\sigma\cdot\sqrt{2\pi}}\, e^{-\frac{1}{2}\left[\frac{x-m}{\sigma}\right]^2}$.

Cette loi est souvent qualifiée de loi du hasard ; elle est très fréquente dans des mesures répétées d'une même grandeur.

On a alors : E(X) = m et V(X) = σ^2



2. La loi normale centrée réduite

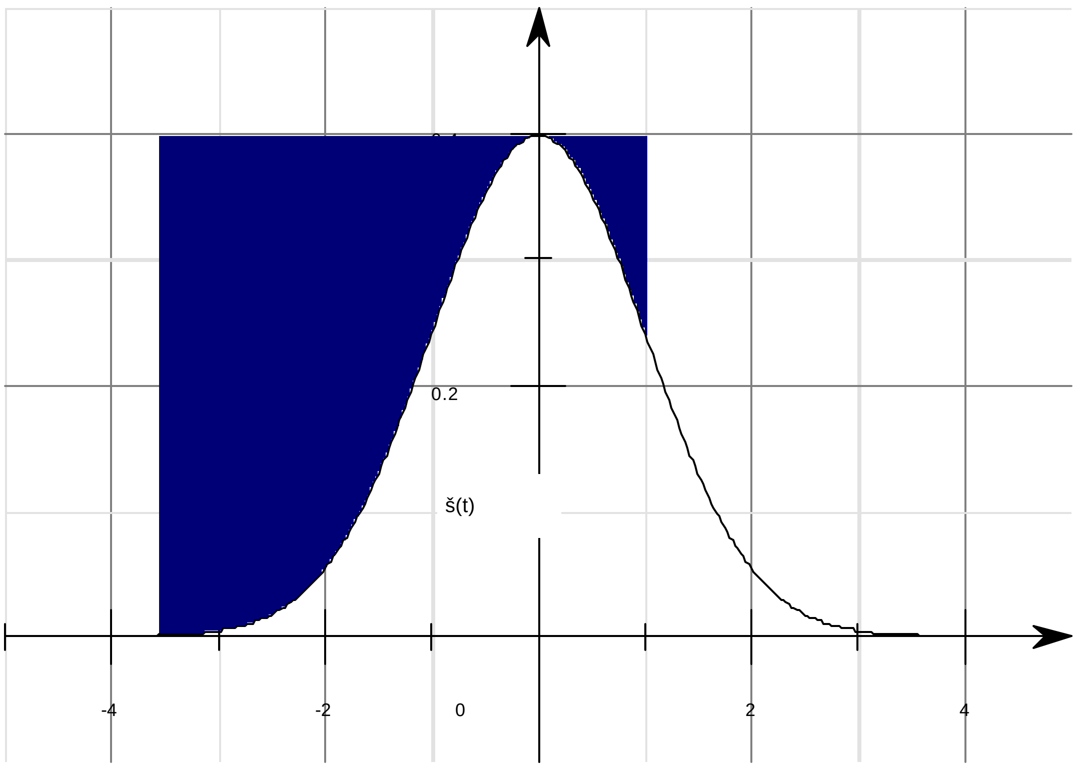
Si une variable aléatoire suit la loi normale $\mathcal{N}$ (m;σ) il est difficile de calculer F(x) pour n'importe quel x; Il existe alors une loi qui est tabulée qui nous permet grâce aux théorème suivant de calculer facilement F(x) pour $\mathcal{N}$ (m;σ) .

a) Théorème :

si une variable aléatoire X suit une loi normale $\mathcal{N}$ (m;σ) alors la variable aléatoire $T = \dfrac{X-m}{\sigma}$ suit la loi normale centrée réduite $\mathcal{N}$ (0;1).

Sa fonction de répartition est notée $\pi(x) = \int_{-\infty}^{x} f(t)dt$ avec $f(x) = \dfrac{1}{\sqrt{2\pi}}\, e^{\frac{-x^2}{2}}$

Sa lecture se fait grâce à une table (cf annexe).



b) <u>Exemples de calculs</u>

p(T 1,67)=π(1,67)=0,9525
p(T 1,25)=1-p(T < 1,25)=1-0,8944=0,1056
p(T -1,67)=p(T 1,67)=1-p(T 1,67)
p(-t T t)=2π(t) - 1

3. <u>Carte de contrôle</u>

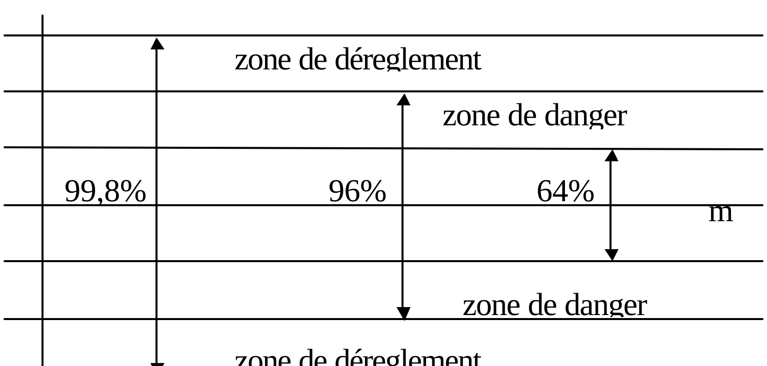
Soit X une variable aléatoire suivant une loi normale N (m;σ) alors $T=\frac{X-m}{\sigma}$ suit une loi normale centrée réduite N (0 ; 1) d'où

$p(m-\sigma<X<m+\sigma)=p(-1<T<1)=2\pi(1)-1=0,64=64\%$

$p(m-2\sigma<X<m+2\sigma)=p(-2<T<2)=2\pi(2)-1=0,96=96\%$

$p(m-3\sigma<X<m+3\sigma)=p(-3<T<3)=2\pi(3)-1=0,998=99,8\%$

ce que l'on représente sous forme de carte de contrôle :



4. Approximation des lois

Dans certaines conditions on peut par commodité approximer certaines lois par une loi normale:

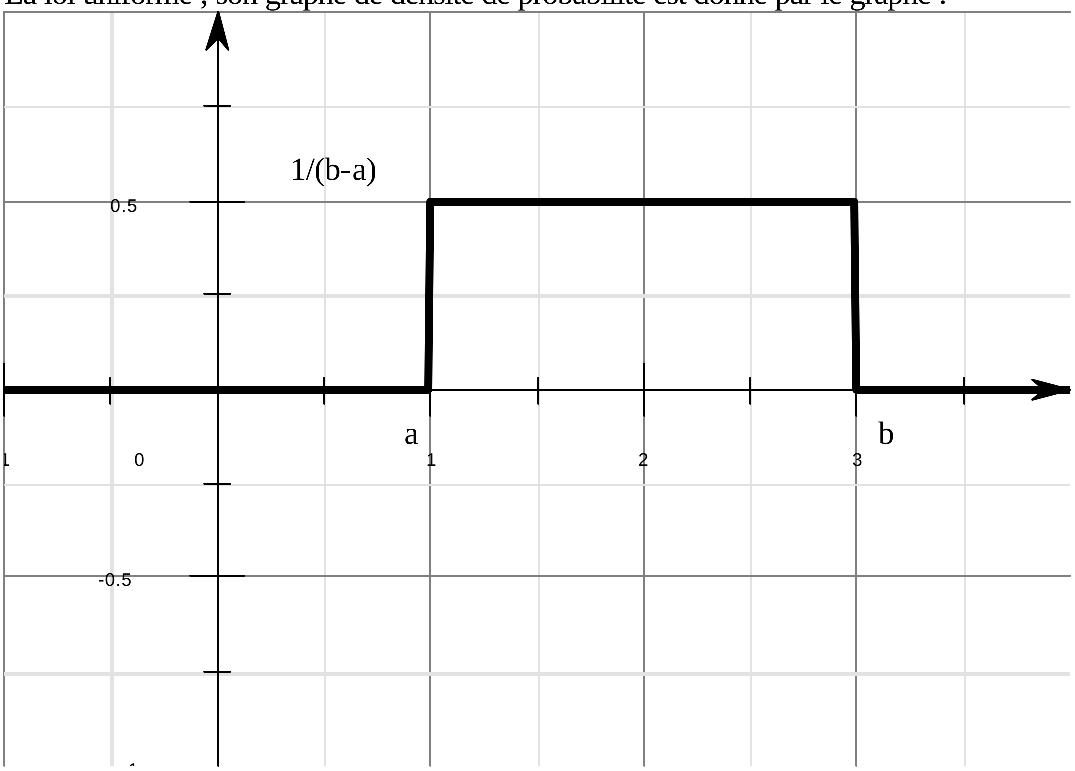
$B(n,p) \approx P(np)$ si $p < 0{,}1$ $npq \leq 10$ et $n > 30$

on a $B(n,p) \approx \aleph(np,\sqrt{npq})$ si $npq > 10$ et $n \geq 50$

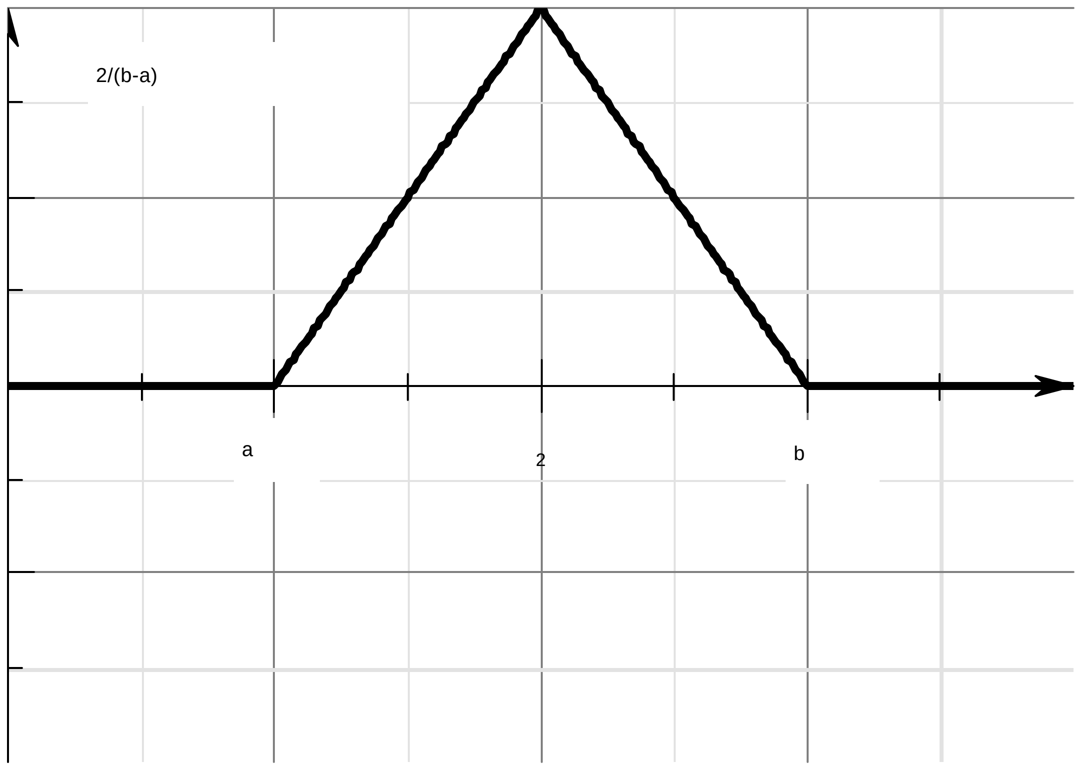
$P(m) \approx \aleph(m,\sqrt{m})$ si $m > 20$

E. D'autres exemples de lois continues

La loi uniforme ; son graphe de densité de probabilité est donné par le graphe :



La loi triangulaire : elle est donnée par son graphe fonction densité de probabilité :


2/(b-a)
a
2
b

Chapitre 4 ECHANTILLONNAGE

A. 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.

B. Distribution d'échantillonnage des moyennes

Considérons une population ayant une certaine propriété avec une moyenne m et un écart-type σ. Soit $\overline{X}$ 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, $\overline{X}$ suit approximativement la loi normale

$\mathcal{N}(m; \frac{\sigma}{\sqrt{n}})$.

Rem :

- n suffisamment grand quand n   30.
- si la population est-elle même normale, on peut utiliser ce résultat même si n est petit.
- lorsque les échantillons de taille n sont prélevés sans remise dans une population d'effectif N, on peut utiliser le résultat précédent en prenant $\frac{\sigma}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}}$ au lieu de $\frac{\sigma}{\sqrt{n}}$.
- il ne faut pas confondre l'écart-type $\frac{\sigma}{\sqrt{n}}$ de la variable aléatoire qui prend pour valeurs les moyennes d 'échantillons de taille n, et l'écart-type d'un échantillon.

C. Distribution d'échantillonnage des pourcentages

Considérons une population dont un pourcentage p d'éléments possède une certaine propriété. Soit F la variable aléatoire, qui à tout échantillon aléatoire prélevé avec remise d'effectif n fixé, associe le pourcentage d'éléments de cet échantillon possédant cette propriété. Pour n suffisamment grand, F suit approximativement la loi normale $\mathcal{N}\left(p, \sqrt{\frac{pq}{n}}\right)$ avec q = 1 - p.

Chapitre 5 ESTIMATION

A. Introduction

C'est le problème inverse de l'échantillonnage ; c'est à dire connaissant des renseignements sur un on plusieurs échantillons, on cherche à en déduire des informations sur la population totale.

B. Estimation ponctuelle

1. Moyenne

De manière générale, on choisit la moyenne $\bar{x}_e$ d'un échantillon prélevé au hasard dans une population comme meilleure estimation ponctuelle de la moyenne inconnue m de cette population.

2. Proportion

De même, on choisit la proportion f_e des éléments possédant une certaine propriété dans un échantillon prélevé aléatoirement dans une population comme meilleure estimation ponctuelle de la proportion inconnue p des éléments de cette population ayant cette propriété.

3. Variance. Ecart-type

On choisit le nombre $\frac{n}{n-1}\sigma_e^2$ où n est l'effectif et σ_e^2 la variance d'un échantillon prélevé au hasard dans une population, comme meilleure estimation ponctuelle de la variance inconnue σ^2 de cette population et on prend $\sqrt{\frac{n}{n-1}}\sigma_e$ comme meilleure estimation ponctuelle de l'écart-type σ inconnue de cette population.

C. 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. De la moyenne

a) 1er cas

Soit P la population : **m** la moyenne est **inconnue**
s l'écart-type est **connu**
Soit un échantillon **:** $\bar{x}_e$ la moyenne est **connue**
n l'effectif est **connu**

On se place dans le cas ou l'on peut considérer que la variable aléatoire $\overline{X}$, qui à tout échantillon de taille n fixée, associe la moyenne de cet échantillon, suit une loi normale $N(m; \frac{\sigma}{\sqrt{n}})$.

Alors l’intervalle de confiance de la moyenne m de la population, avec le coefficient de confiance $2\pi(t)$ - 1, lu dans la table de la loi normale centrée réduite $\mathcal{N}(0\,;1)$ est :

$$\left[\bar{x}_e - t\frac{\sigma}{\sqrt{n}}\,;\,\bar{x}_e + t\frac{\sigma}{\sqrt{n}}\right]$$

Cette méthode conduit dans $100(2\pi(t) - 1)$ cas sur 100, pourcentage choisi à l’avance, à un intervalle de confiance contenant m.
Les cas usuels les plus fréquent sont :

- coefficient de confiance 95% alors t = 1,96
- coefficient de confiance 99% alors t=2,58.

b) <u>2ème cas</u>

Soit P la population : **m** la moyenne est **inconnue**
s l’écart-type est **inconnu**
Soit un échantillon **:** $\bar{x}_e$ la moyenne est **connue**
$\mathbf{s_e}$ l’écart-type est **connu**
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 $\bar{X}$, qui à tout échantillon de taille n fixée, <u>**n < 30**</u>**,** 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 m de la population, avec le coefficient de confiance $2\delta(t)$ - 1, lu dans la table de la loi de Student à n - 1 degrés de liberté est :

$$\left[\bar{x}_e - t\frac{\sigma_e}{\sqrt{n-1}}\,;\,\bar{x}_e + t\frac{\sigma_e}{\sqrt{n-1}}\right]$$

c) <u>3ème cas</u>

Soit P la population : **m** la moyenne est **inconnue**
s l’écart-type est **inconnu**
Soit un échantillon **:** $\bar{x}_e$ la moyenne est **connue**
$\mathbf{s_e}$ l’écart-type est **connu**
n l’effectif est **connu et il est supérieur à 30.**

On se place dans le cas ou l’on peut considérer que la variable aléatoire $\bar{X}$, qui à tout échantillon de taille n fixée, <u>**n   30**</u>**,** associe la moyenne de cet échantillon, suit une loi normale $\mathcal{N}(m; \frac{\sigma}{\sqrt{n}})$.
Alors l’intervalle de confiance de la moyenne m de la population, avec le coefficient de confiance $2\pi(t)$ - 1, lu dans la table de la loi de la loi normale centrée réduite $\mathcal{N}(0\,;1)$:

$$\left[\bar{x}_e - t\frac{\sigma_e}{\sqrt{n-1}}\,;\,\bar{x}_e + t\frac{\sigma_e}{\sqrt{n-1}}\right]$$

2. De la proportion

A l'aide d'un échantillon, on définit de même un intervalle de confiance de la proportion p **inconnue** d'une caractéristique de la population.

a) 1er cas

Soit P la population : **p** la proportion est **inconnue**
Soit un échantillon : $\mathbf{f_e}$ la proportion est **connue**
n l'effectif est **connu et inférieur strictement à 30.**
Alors l'intervalle de confiance de la proportion p de la population avec le coefficient de confiance 2δ(t) - 1, lu dans la table de la loi de Student à n - 1 degrés de liberté est :

$$\left[f_e - t\sqrt{\frac{f_e(1-f_e)}{n-1}} ; f_e + t\sqrt{\frac{f_e(1-f_e)}{n-1}}\right].$$

b) 2ème cas

Soit P la population : **p** la proportion est **inconnue**
Soit un échantillon : $\mathbf{f_e}$ la proportion est **connue**
n l'effectif est **connu et supérieur à 30.**
Alors l'intervalle de confiance de la proportion p de la population avec le coefficient de confiance 2π(t) - 1, lu dans la table de la loi normale centrée réduite $N(0 ; 1)$ est :

$$\left[f_e - t\sqrt{\frac{f_e(1-f_e)}{n-1}} ; f_e + t\sqrt{\frac{f_e(1-f_e)}{n-1}}\right].$$

3. Exemples

a) 1er exemple

Dans une population P de grand effectif, on prélève de manière non exhaustive, un échantillon de 100 personnes dont on note la masse en kg:

masse	62	64	68	10	74
effectif	5	18	42	27	8

- Calculer la moyenne et l'écart-type de cet échantillon: $\bar{x}_e = 68 \text{ kg} \quad \sigma_e = 3 \text{ kg}$

2. Donner un intervalle de confiance de la moyenne m des masses des personnes de P au coefficient de confiance 95% : nous sommes dans le 3ème cas

$$m \in \left[68 - 1{,}96\frac{3}{\sqrt{100-1}} ; 68 + 1{,}96\frac{3}{\sqrt{100-1}}\right] = [67{,}4 ; 68{,}6]$$

b) 2ème exemple

Lors d'un contrôle de qualité sur une population d'appareils ménagers, au cours d'un mois de fabrication, on prélève de manière non exhaustive un échantillon de 1000 appareils. Après un test de conformité, on constate que 60 appareils ont un défaut. Donner un intervalle de confiance du pourcentage p d'appareils défectueux au risque de 5%.

$$p \in \left[\frac{60}{1000} - 1,96\sqrt{\frac{\frac{60}{1000}\left(1-\frac{60}{1000}\right)}{1000-1}} ; \frac{60}{1000} + 1,96\sqrt{\frac{\frac{60}{1000}\left(1-\frac{60}{1000}\right)}{1000-1}} \right] = [0,045;0,075] = [4,5\%;7,5\%]$$

D. ***Analyse de sensibilité des coefficients d'une droite d'ajustement***

Il est intéressant dans les applications d'obtenir un intervalle de confiance à ε près des coefficients réels, inconnus évidemment $\hat{a}$ et $\hat{b}$, à partir des données mesurées pour cela on admet que :

Si on appelle $Z = \frac{\sqrt{V(X)V(Y) - Cov(X,Y)^2}}{V(X)\sqrt{n-2}}$ alors les variables aléatoires :

$T = \frac{a-\hat{a}}{Z}$ et $U = \frac{(b-\hat{b})}{Z}\frac{\sqrt{n}}{\sqrt{\sum_{i=1}^{n} x_i^2}}$ obéissent toutes les deux à une loi de Student à n-2 degrés de liberté.

Exemple : Déterminer des intervalles de confiance à 95% (0.05) des coefficients a et b

Etape 1 : *Détermination de t*

P($|T|$>t)= ε avec ε=0.05 donne dans la table de Student à 12-2=10 degrés de liberté t=2.228 c'est à dire
P($|T|$<t)=95% ou encore
P(-2.28<T<2.28)=95%

Etape 2 : *Calcul de Z*

$Z^2 = \frac{2206 x 7857 - 4394}{(12-2)2206^2} = 0.243$ et Z=0.243

Etape 3 : *Encadrement de a*

$T = \frac{a-\hat{a}}{Z}$ et a=1.68 par conséquent : a-tZ<$\hat{a}$<a+tZ finalement :

1.68-2.28x0.243<$\hat{a}$<1.68+2.28x0.243

1.12<$\hat{a}$<2.23

Etape 4 : *Encadrement de b*

$U = \frac{(b-\hat{b})}{Z} \frac{\sqrt{n}}{\sqrt{\sum_{i=1}^{n} x^2_i}}$ et b=-1.3 par conséquent : $b - tZ\frac{\sqrt{\sum_{i=1}^{n} x^2_i}}{\sqrt{n}} < \hat{b} < b + tZ\frac{\sqrt{\sum_{i=1}^{n} x^2_i}}{\sqrt{n}}$ et donc :

$$-1.3 - 2.28x0.24x\sqrt{\frac{224822}{12}} < \hat{b} < -1.3 + 2.28x0.24x\sqrt{\frac{224822}{12}}$$

$$-76.2 < \hat{b} < 73.6$$

Chapitre 6 TESTS STATISTIQUES

A. Principe des tests

Partons d'un exemple...Une machine fabrique des tiges d'acier. Si la machine est réglée correctement, l'utilisateur obtient une population de tiges de longueurs moyenne m et d'écart-type σ. On désire savoir si cette machine se dérègle. Ainsi, on prélèvera, à intervalles réguliers, des échantillons pour mesurer la longueur effective des tiges.
Nous faisons alors l'hypothèse H_0 dite hypothèse nulle que la machine est bien réglée. On teste alors cette hypothèse: 2 cas se présentent :

- la machine est bien réglée, on accepte H_0.
- la machine est mal réglée, on rejette H_0 et donc on accepte H_1 dite hypothèse alternative.

Définition : un test statistique est une méthode permettant de prendre une **décision** à partir d'informations fournies par un **échantillon**.
Cette décision dépend donc de l'échantillon. Ainsi qu'elle que soit la décision prise, on court deux sortes de risques :

- le risque dit de 1ère espèce noté α, est la probabilité de rejeter l'hypothèse H_0 alors qu'elle est vraie en réalité : $\alpha = p(\text{rejeter } H_0 / H_0 \text{ vraie})$
- le risque dit de 2nde espèce noté β, est la probabilité d'accepter l'hypothèse H_0 alors qu'elle est fausse en réalité : $\beta = p(\text{accepter } H_0 / H_0 \text{ fausse})$.

Un test est bon si on arrive à minimiser α et β.

B. Test de comparaison à une valeur standard

1. Position de problème

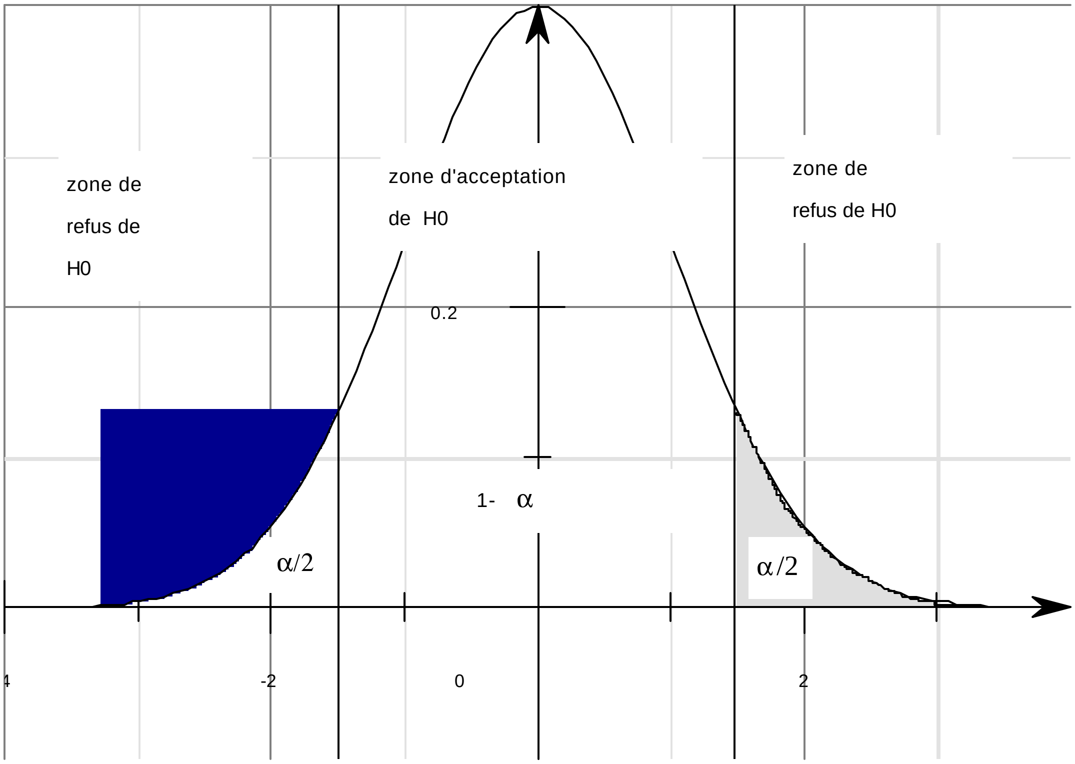
On considère une population P sur laquelle on veut étudier un paramètre γ inconnue associé à un paramètre c. Sur un échantillon de taille n, on obtient γ_e connu. Sur la base de cette valeur observée γ_e, on se propose de comparer la vraie valeur γ à une valeur γ_0 fixée à priori, constituant la valeur standard.

2. Tests relatifs à une moyenne

Soit une population P de grand effectif sur laquelle on étudie un caractère c. La moyenne m de c est inconnue. Sur un échantillon, on a trouvé une moyenne $\bar{x}_e$. On doit tester la moyenne m par rapport à une valeur notée m_0 qui est la valeur standard.

a) Test bilatéral

Soit H_0 : " $m=m_0$ " l'hypothèse nulle et H_1 : " m m_0 " l'hypothèse alternative
Soit X la variable aléatoire prenant pour valeurs les moyennes des différents échantillons de taille n   30, alors on sait que X suit une $N(m_0 ; \frac{\sigma}{\sqrt{n}})$, σ étant l'écart-type de la population P.



Il faut donc que X soit telle que $p(m_0 - t_\alpha \frac{\sigma}{\sqrt{n}} < X < m_0 + t_\alpha \frac{\sigma}{\sqrt{n}}) = 1 - \alpha$ d'où en faisant le changement de variable : $T = \dfrac{X - m_0}{\frac{\sigma}{\sqrt{n}}}$ alors T suit une loi normale centrée réduite $N(0,1)$ d'où

$p(-t_\alpha < T < t_\alpha) = 1 - \alpha$ c'est à dire $\pi(t_\alpha) = 1 - \frac{\alpha}{2}$ d'où la règle du test bilatéral :

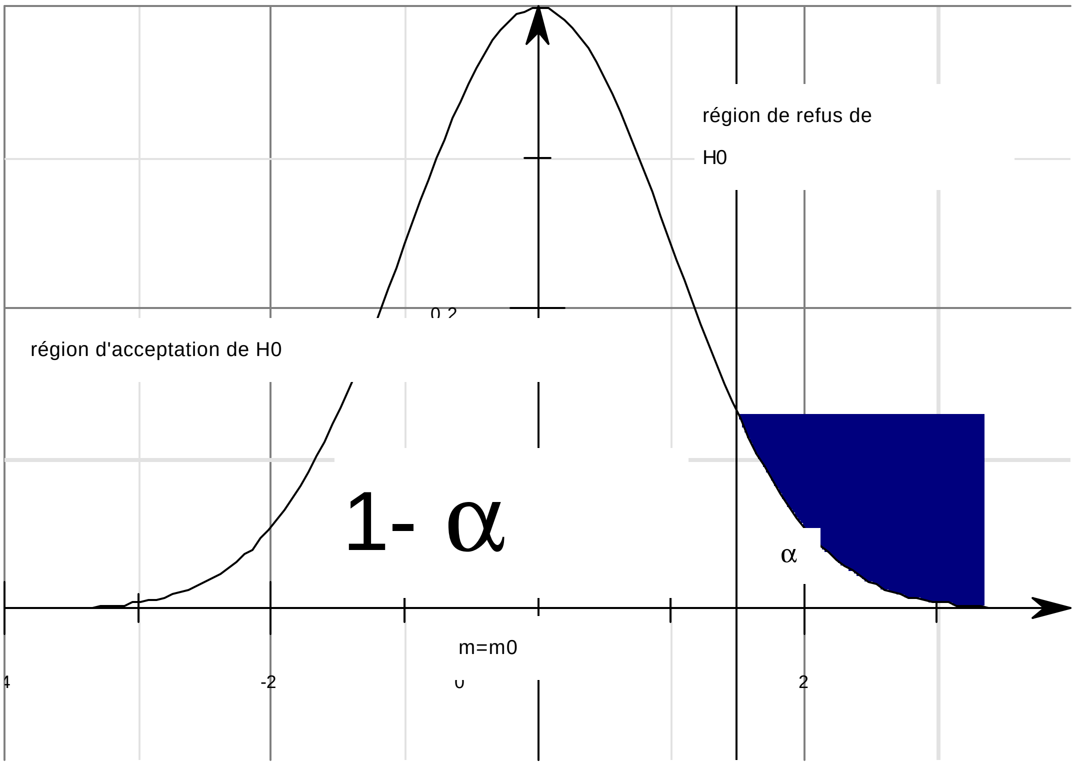
- On choisit un risque α
- on cherche dans la table de la loi normale centrée réduite $N(0,1)$, t_α tel que $\pi(t_\alpha) = 1 - \frac{\alpha}{2}$
- soit $\bar{x}_e$ la moyenne de l'échantillon de taille n alors

si $\bar{x}_e \in \left[m_0 - t_\alpha \frac{\sigma}{\sqrt{n}} ; m_0 + t_\alpha \frac{\sigma}{\sqrt{n}} \right]$, on accepte H_0 avec le risque α
sinon on rejette H_0 et donc on accepte H_1 avec un risque α.
Remarque : dans le cas usuel où α = 5% alors t_α = 1,96 et si α = 1% alors t_α = 2,58.

Si σ est inconnu (ce qui est souvent le cas) alors on prend son estimateur ponctuel $\sqrt{\frac{n}{n-1}}\sigma_e$ où σ_e est l'écart-type de l'échantillon.

b) <u>Tests unilatéraux</u>

Règle du test unilatéral à gauche

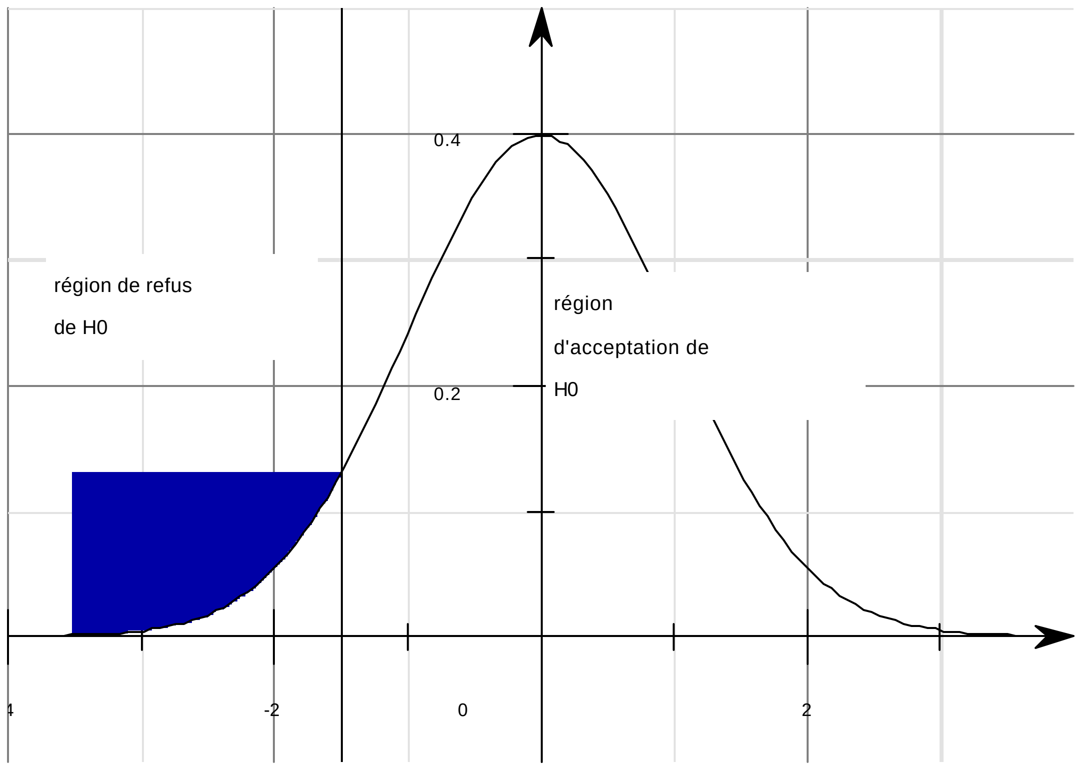


L'hypothèse nulle est H_0 : "$m = m_0$" et l'hypothèse alternative est H_1 : "$m > m_0$"
On peut la retrouver par exemple dans le cas d'un test de dépassement d'une norme.

- On choisit un risque α
- On cherche dans la table de la loi normale centrée réduite $N(0;1)$ t_α tel que $\pi(t_\alpha) = 1 - \alpha$
- Si $\bar{x}_e \leq m_0 + t_\alpha \frac{\sigma}{\sqrt{n}}$ alors on accepte H_0 sinon on refuse H_0 et donc on accepte H_1

Rem : dans les cas usuels : si $\alpha = 5\%$ alors $t_\alpha = 1{,}645$; si $\alpha = 1\%$ alors $t_\alpha = 2{,}33$

Règle du test unilatéral à droite



L'hypothèse nulle est : H_0 : "$m = m_0$" et l'hypothèse alternative est : H_1 : "$m < m_0$"
On la retrouve dans les cas de tests de non égalité d'une norme.

- On choisit un risque α
- On cherche dans la table de la loi normale centrée réduite $N(0;1)$ t_α tel que $\pi(t_\alpha) = 1 - \alpha$
- Si $\bar{x}_e > m_0 - t_\alpha \dfrac{\sigma}{\sqrt{n}}$ on accepte H_0, sinon on rejette H_0 et donc on accepte H_1

Dans ces deux cas, il est très fréquent qu'on ne connaisse pas σ ; on a alors $\dfrac{\sigma}{\sqrt{n}} = \dfrac{\sigma_e}{\sqrt{n-1}}$

3. Tests relatifs à un fréquence ou un pourcentage

Tous les tests que l'on vient de voir restent valables ; il suffit de remplacer m par p (proportion inconnue dans la population P), $\bar{x}_e$ par f_e (proportion effective sur l'échantillon) et $\dfrac{\sigma}{\sqrt{n}}$ par $\sqrt{\dfrac{f_e(1-f_e)}{n-1}}$

C. Test de comparaison de 2 populations

1. Test de comparaison de 2 moyennes

	Population P1		Population P2	
Caractères Etudiés	C		C	
Moyenne Ecart-type	m_1 σ_1	inconnus	m_2 σ_2	inconnus
	Echantillon e_1		Echantillon e_2	
Taille Moyenne Ecart-type	n_1 30 $\bar{x}_{e_1}$ σ_{e_1}	connus	n_2 30 $\bar{x}_{e_2}$ σ_{e_2}	connus

Règle du test de comparaison de 2 moyennes

L'hypothèse nulle est : H_0 : "$m_1 = m_2$" et l'hypothèse alternative est H_1 : "$m_1 \quad m_2$"

- On choisit un risque α
- On cherche dans la table de la loi normale centrée réduite $N(0;1)$ t_α tel que $\pi(t_\alpha) = 1 - \alpha/2$
- Si $\dfrac{\bar{x}_{e_1} - \bar{x}_{e_2}}{\sqrt{\dfrac{\sigma_{e_1}^2}{n_1 - 1} + \dfrac{\sigma_{e_2}^2}{n_2 - 1}}} \in \left[-t_\alpha ; t_\alpha\right]$ on accepte H_0 sinon on rejette H_0 et donc on accepte H_1
- Si H_0 est acceptée, on dit que la différence m_1-m_2 n'est pas significative au risque α

2. <u>Règle de comparaison de deux pourcentages</u>

	Population P_1	Population P_2
Caractère étudié Pourcentage	C p_1 inconnu	C p_2 inconnu
	Echantillon e_1	Echantillon e_2
Taille Pourcentage	n_1 30 f_1 connu	n_2 30 f_2 connu

L'hypothèse nulle est H_0 : "$p_1 = p_2$" et l'hypothèse alternative H_1: "$p_1 \quad p_2$"

- On choisit un risque α
- On cherche dans la table de la loi normale centrée réduite $N(0;1)$ t_α tel que $\pi(t_\alpha) = 1 - \alpha/2$
- Soit $f = \dfrac{n_1 f_1 + n_2 f_2}{n_1 + n_2}$ alors si $\dfrac{f_1 - f_2}{\sqrt{f(1 - f)\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}} \in \left[-t_\alpha ; t_\alpha\right]$ on accepte H_0 sinon on rejette H_0 et

  on accepte donc H_1.
- Si H_0 est acceptée, on dit que p_1 - p_2 n'est pas significative au risque α.

D. __Test du Khi-Deux__

On considère une distribution à laquelle on associe un ensemble fini de probabilités qui sont, soit les probabilités d'un ensemble fini d'événements élémentaires, soit les probabilités de regroupements de classes statistiques. Les classes peuvent correspondre à des modalités qualitatives prises par une variable non numériques ou à des modalités quantitatives classiques. Les cas les plus classiques sont :

- 2 classes de probabilités données p et 1-p
- k classes équiprobables
- k classes dont les probabilités sont données à priori
- n+1 classes associées aux valeurs d'une variable binomiale de paramètres n et p
- n+1 classes associées aux valeurs d'une variable aléatoire à valeurs entières positives, la n+1 ème regroupant les valeurs supérieures ou égales à n

Le but du test est de comparer globalement la distribution observée à la distribution théorique. Un risque d'erreur α est fixé dans les conditions habituelles.

On note A1, A2,…,Ak les k classes. On dispose, d'une part les k probabilités théoriques p_j , d'autre part d'un échantillon de taille n dont on connaît les effectifs observés des classes. L'hypothèse H_0 est la conformité de la distribution réelle à la distribution théorique, et s'exprimera en affirmant, pour chaque classe, l'égalité de la probabilité réelle avec la probabilité théorique, soit H_0 = « $p(A1) = p_1,\ldots,p(Ak) = p_k$ ».

L'écart entre la distribution réelle et la distribution théorique est calculé en faisant intervenir pour chaque classe deux effectifs : l'effectif observé o_j et l'effectif théorique np_j .

Une représentation graphique de la sorte est alors plus pratique :

o_1 np_1		o_j np_j		

L'exécution du test du $\aleph^2$ fait intervenir un paramètre qui est le nombre de degrés de liberté ddl et qui prend une valeur liée au nombre de cases :

- dans le cas général , ddl = k – 1 (nbre de cases moins un)
- dans le cas d'une loi binomiale, d'une loi de Poisson ou d'une loi normale ddl = k – 2

Hypothèse à tester : H_0 = « $p(A1) = p_1,\ldots,p(Ak) = p_k$ »

Ecart à tester : $\aleph^2 = \sum_{j=1}^{k} \frac{(o_j - np_j)^2}{np_j}$

Ecart critique : $\aleph_a^2$ est lu dans la table de la loi du $\aleph^2$, avec le nombre de degrés de liberté indiqué plus haut et le risque souhaité.

Règle : Si $\aleph^2 \le \aleph_a^2$ alors H_0 est accepté, sinon elle est refusée.

Une restriction d'emploi est cependant nécessaire : les **effectifs théoriques np_j** ne doivent pas être plus petits que 5. En cas de difficultés, la solution consiste à regrouper des classes contiguës pour que ces effectifs théoriques dépassent 5.