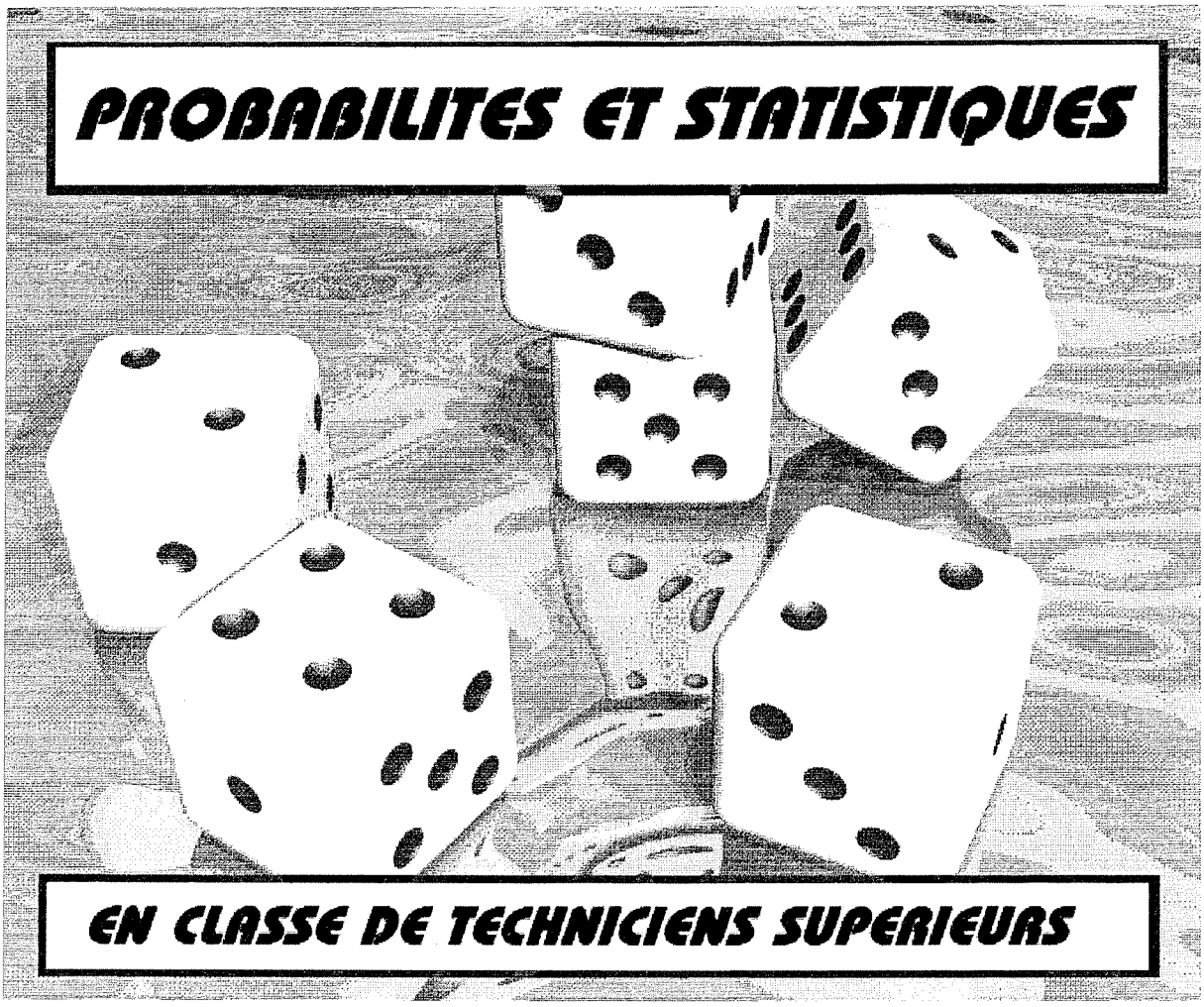
UNIVERSITE
LOUIS PASTEUR
STRASBOURG

PROBABILITES ET STATISTIQUES



EN CLASSE DE TECHNICIENS SUPERIEURS

1996

IREM
10, rue du Général Zimmer
67084 STRASBOURG CEDEX
Tél. 88 41 63 07 Secrétariat
88 41 64 40 Bibliothèque

I
Introduction

Cette brochure est un recueil de différents documents sur des parties qui nous ont semblé délicates dans l'enseignement des probabilités en classe de techniciens supérieurs.
Elle n'a pas l'ambition d'aborder tous les points figurant aux programmes des différents brevets de techniciens supérieurs.
Elle n'est pas présentée sous la forme d'un cours complet et organisé, mais nous espérons que les collègues y puiseront des exemples pertinents et y trouveront des compléments permettant un ancrage plus solide de leur enseignement.

Pour chacun des thèmes abordés, nous avons adopté la présentation suivante.
Une première page précise s'il s'agit :

- d'un document pour l'élève que l'on peut utiliser ainsi dans une classe.
- d'un document pour l'élève avec des compléments pour les enseignants.
- d'un document plus spécialement destiné aux enseignants et permettant de répondre à un certain nombre de questions au delà du programme des classes de techniciens supérieurs (questions qui nous ont été posées par les enseignants lors de différentes rencontres de professeurs de ces classes).

Les notions abordées figurent au sommaire de la page suivante.
En ce qui concerne les notions de probabilité conditionnelle, indépendance, variable aléatoire, qu'il convient souvent de reprendre en classe de techniciens supérieurs , le lecteur pourra se reporter à notre brochure:
Enseigner les probabilités en classe de Terminale, IREM de STRASBOURG, mars 1994

Nous espérons que cette nouvelle brochure sera utile à nos collègues, confrontés à la lourde tâche d'initier leurs étudiants des classes de techniciens supérieurs aux subtilités des probabilités et statistiques.

Probabilités et statistiques en classe de techniciens supérieurs

Cette brochure a été rédigée par

Claire DUPUIS
Mohamed ATLAGH
André BASTIAN
Bernard GOESEL
Christiane HEUSCH
Bernard KOCH
Dominique PERNOUX
Suzette ROUSSET-BERT

SOMMAIRE

Document pour l'élève

Ce document contient quelques observations complémentaires sur l'indépendance destinées aux professeurs.

Opérations sur les variables aléatoires

Objectif:

Définir le vocabulaire et donner quelques résultats concernant les couples de variables aléatoires sur des exemples simples.

OPERATIONS SUR LES VARIABLES ALEATOIRES

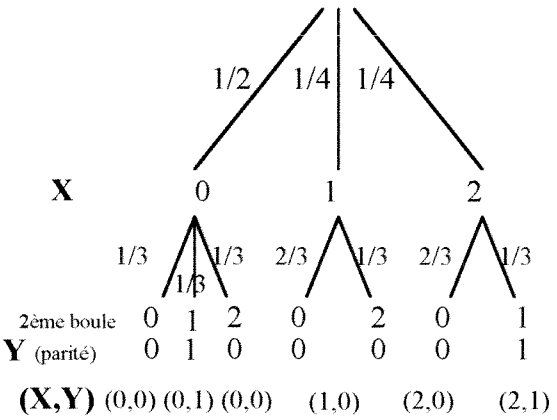
OPERATIONS SUR LES VARIABLES ALEATOIRES

Une urne contient 4 boules numérotées 0, 0, 1 et 2. On considère les deux expériences aléatoires suivantes :
Expérience A : on tire une boule au hasard dans l'urne. Sans remise on effectue un deuxième tirage.
Expérience B : on tire une boule au hasard. On remet la boule tirée dans l'urne et ont effectue un deuxième tirage.
Soit X la variable aléatoire prenant pour valeur le numéro x porté par la première boule tirée et Y celle prenant pour valeur 0 si le numéro porté par la deuxième boule tirée est pair et 1 si le numéro porté par la deuxième boule tirée est impair . Y code ainsi la parité de la deuxième boule tirée (1=impair et 0=pair).

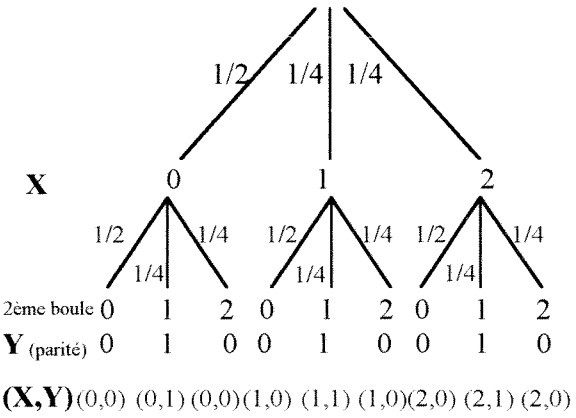
I. Loi d'un couple aléatoire.

Les lois de chacune des variables aléatoires X et Y se déduisent des arbres suivants :

Expérience A : (sans remise)



Expérience B : (avec remise)



On obtient pour ces deux expériences les mêmes lois pour X et Y.

x_i	0	1	2
$P(X=x_i)$	1/2	1/4	1/4

y_i	0	1
$P(Y=y_i)$	3/4	1/4

Dans les deux cas A et B on obtient donc la même espérance et la même variance pour les variables aléatoires X et Y :

E(X)=3/4 V(X)=11/16
E(Y)=1/4 V(Y)=3/16

On note (X,Y) le couple aléatoire prenant les valeurs (x,y). On introduit la loi de probabilité de ce couple par les probabilités élémentaires :

$$p_{ij} = P\left[(X = x_i) \text{ et } (Y = y_j)\right]$$

En regroupant ces probabilités (que l'on peut déduire des arbres ci-dessus) dans un tableau à double entrée on obtient la **loi de probabilité du couple** (X,Y) :

Expérience A (sans remise) :

Y↓ X→	0	1	2	somme
0	1/3	1/4	1/6	3/4
1	1/6	0	1/12	1/4
somme	1/2	1/4	1/4	

Expérience B (avec remise) :

Y↓ X→	0	1	2	somme
0	3/8	3/16	3/16	3/4
1	1/8	1/16	1/16	1/4
somme	1/2	1/4	1/4	

On obtient par sommation dans les marges grisées du tableau les probabilités obtenues pour les lois de probabilités des variables aléatoires X et Y . Les lois de X et Y sont dans ces circonstances appelées "**lois marginales**" de (X,Y).
On remarque que pour les deux expériences A et B, les lois de X et de Y sont les mêmes, mais les lois des couples (X,Y) sont différentes. La seule connaissance des lois de X et de Y est donc insuffisante pour donner la loi du couple (X,Y).

II. Indépendance de deux variables aléatoires :

Dans chacun des 6 cas de l'expérience B on remarque que :

$$P\left[(X = x_i) \text{ et } (Y = y_j)\right] = P(X = x_i) \times P(Y = y_j)$$

Cette propriété n'est pas vérifiée pour l'expérience A. On a par exemple :

$$P\left[(X = 1) \text{ et } (Y = 1)\right] = 0 \text{ alors que } P(X = 1) \times P(Y = 1) = \frac{1}{16}$$

Dans le cas de l'expérience B on dira que les variables aléatoires X et Y sont indépendantes.

Définition :
Soit X et Y deux variables aléatoires définies dans le même univers.
X prend les valeurs $x_1, x_2, \ldots, x_n$
Y prend les valeurs $y_1, y_2, \ldots, y_p$
X et Y sont deux variables aléatoires indépendantes si et seulement si pour tout couple (x_i, y_j) :

$$P\left[(X = x_i) \text{ et } (Y = y_j)\right] = P(X = x_i) \times P(Y = y_j)$$

III. Somme de deux variables aléatoires.

Soit S la variable aléatoire somme de X et et de Y : S=X + Y.
Les lois de probabilités des couples permettent d'établir la loi de S pour chacune des expériences :

Expérience A : (sans remise)

s_i	0	1	2	3
$P(S=s_i)$	1/3	5/12	1/6	1/12

On calcule l'espérance et la variance de S=X+Y : E(X+Y)=1 et V(X+Y)=5/6.

Expérience B : (avec remise)

s_i	0	1	2	3
$P(S=s_i)$	3/8	5/16	1/4	1/16

On calcule l'espérance et la variance de S=X+Y : E(X+Y)=1 et V(X+Y)=14/16.

On remarque que dans les deux cas : E(X+Y) = E(X) + E(Y).

En revanche :

Dans le cas A, V(X) + V(Y) = 11/16+3/16 = 14/16 n'est pas égal à V(X+Y)

Dans le cas général nous admettrons que :

Pour tout couple de variables aléatoires (X,Y) on a : E(X+Y) = E(X) + E(Y). Si X et Y sont indépendantes on a de plus : V(X + Y) = V(X) + V(Y)

La réciproque de la seconde affirmation n'est pas vraie. L'exercice n°4 proposé plus loin dans ce chapitre fournit un contrexemple.

IV. Différence de deux variables aléatoires.

Soit D la variable aléatoire D = X - Y. Comme précédemment, à l'aide de la loi du couple, on construit la loi de X-Y.

Expérience A : (sans remise)

d_i	-1	0	1	2
$P(D=d_i)$	1/6	1/3	1/3	1/6

On calcule l'espérance et la variance de S=X+Y : E(X+Y)=1/2 et V(X+Y)=11/12.

Expérience B : (avec remise)

d_i	-1	0	1	2
$P(d=d_i)$	1/8	7/16	1/4	3/16

On calcule l'espérance et la variance de S=X+Y : E(X+Y)=1/2 et V(X+Y)=14/16.

On remarque, là aussi, que dans les deux cas : E(X - Y) = E(X) - E(Y).

En revanche :

Dans le cas A, V(X) + V(Y) = 11/16+3/16 = 14/16 n'est pas égal à V(X - Y)
Dans le cas B, V(X - Y) = V(X) + V(Y).

Dans le cas général nous admettrons que :

Pour tout couple de variables aléatoires (X,Y) on a : E(X - Y) = E(X) - E(Y). Si X et Y sont indépendantes on a de plus : V(X - Y) = V(X) + V(Y)

V. Somme de 2 variables aléatoires indépendantes suivant chacune 1 loi de Poisson

X_1 suit une loi de Poisson de paramètre λ_1

X_2 suit une loi de Poisson de paramètre λ_2

On veut savoir quelle loi suit la variable aléatoire $Y = X_1 + X_2$
On prend alors un exemple :

X_1 suit une loi de Poisson de paramètre $\lambda_1 = 1$
On peut donc lire dans la table les probabilités $p(x_1 = k)$ pour $0 \leq k \leq 6$

X_2 suit une loi de Poisson de paramètre $\lambda_2 = 2$

On peut donc lire dans la table les probabilités $\rho(x_2 = k)$ pour $0 \leq k \leq 8$
On se limitera donc à $0 \leq k \leq 6$.
On cherche la loi de $Y = X_1 + X_2$. On se limitera pour Y aux valeurs de 0 jusqu'à 12.
On peut construire un tableau donnant la loi du couple (X_1, X_2) puis en déduire la loi de Y.

Poisson $\lambda_1 = 1 \rightarrow$ Poisson $\lambda_2 = 2$ ↓	0 (0,368)	1 (0,368)	2 (0,184)	3 (0,061)	4 (0,015)	5 (0,003)	6 (0,0005)
0 (0,135)	0,0497						
1 (0,271)	0,0997						
2 (0,271)							
3 (0,180)							
4 (0,090)							
5 (0,036)							
6 (0,012)							

Y	0	1	2	3	4	5	6	7	8	9	10	11	12
proba	0,05	0,149	0,224	0,224	0,168								

Compléter le tableau précédent. Vérifier que l'on obtient les valeurs des probabilités correspondant à la loi de Poisson de paramètre $\lambda_3 = 3 = 1 + 2$

On admet que ce résultat est général

Si la variable X_1 suit une loi de Poisson de paramètre λ_1 ;

la variable X_2 suit une loi de Poisson de paramètre λ_2 ;

les variables X_1 et X_2 sont indépendantes ;

Alors la variable aléatoire $Y = X_1 + X_2$ suit une loi de Poisson de paramètre $\lambda_1 + \lambda_2$.

VI. EXERCICES

Exercice n° 1 : dans cet exercice, l'indépendance est postulée a priori

Un concessionnaire livre des appareils. X est la variable aléatoire représentant le nombre d'appareils vendus en une semaine. Les probabilités de vente sont données dans le tableau

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

Ces probabilités sont invariables d'une semaine à l'autre

Soit Z la variable aléatoire prenant pour valeur le nombre d'appareils vendus au bout de deux semaines. Donner les valeurs possibles pour Z et calculer les probabilités correspondantes. Calculer E(X) ; E(Z) ; Var (X) et Var(Z).

Exercice n° 2

Une entreprise vend deux articles A et B dont les ventes suivent des variables aléatoires indépendantes X et Y caractérisées par :

$E(X) = 10\,000$ et $\sigma_X = 25$
$E(Y) = 9\,000$ et $\sigma_Y = 20$

Les prix de vente unitaires et les coûts de fabrication sont respectivement pour A et B :

$P_A = 180$ et $C_A = 110$
$P_B = 200$ et $C_B = 100$

L'entreprise supporte en outre des coûts de fabrication fixes égaux à 1 000 000 F

1) Donner une expression du chiffre d'affaires total : R, du coût total C et de la marge totale B.
   Caractériser ces trois variables aléatoires.
2) L'entreprise souhaiterait augmenter sa marge de 5%. De combien doit on augmenter le prix de vente de l'article A, les autres données restant constantes ?

Exercice n° 3 : destiné à montrer que la propriété Var (X + Y) = Var X + Var Y n'entraîne pas nécessairement que X et Y sont indépendantes.

On lance deux dés cubiques normaux. On note U et V les points amenés par chacun de ces dés.
On pose X = U + V et Y = U - V

1) Donner les lois de probabilités de X et Y.
   Vérifier que X et Y ne sont pas indépendantes en calculant par exemple la probabilité d'avoir simultanément X = 5 et Y = 0.

2) Calculer E(X) ; Var (X) ; E (Y) ; Var (Y)
   Vérifier que X + Y = 2 U ; en déduire (Var (X + Y)
   Vérifier que Var (X) + Var (Y) = Var (X + Y)

Eléments de corrigés

Exercice n° 1

X_2 ↓ X_1 →	0	1	2	3	4
0	0,01	0,02	0,03	0,03	0,01
1	0,02	0,04	0,06	0,06	0,02
2	0,03	0,06	0,09	0,09	0,03
3	0,03	0,06	0,09	0,09	0,03
4	0,01	0,02	0,03	0,03	0,01

$Z = X_1 + X_2$	0	1	2	3	4	5	6	7	8
proba	0,01	0,04	0,10	0,18	0,23	0,22	0,15	0,06	0,01

X_1 est la variable aléatoire prenant pour valeurs le nombre d'appareils vendus la première semaine

X_2 est la variable aléatoire prenant pour valeurs le ombre d'appareils vendus la deuxième semaine

$Z = X_1 + X_2$

$E(X_1) = E(X_2) = 2,1$ $\quad$ $Var(X_1) = Var(X_2) = 1,29$

$E(Z) = 4,2$ $\quad$ $Var(Z) = Var(X_1) + Var(X_2) = 2,58$

Attention à l'erreur classique $Z = 2X$
En effet, la variable 2X ne prend que les valeurs 0, 2, 4, 6,8, et ne correspond pas au nombre d'appareils vendus au bout de deux semaines.

$E(2X) = 2\,E(X) = 4,2$ mais $Var(2X) = 4\,Var(X) = 5,16$.

Exercice n°2

1°) $R = 180X + 200Y$ $\quad E(R) = 3\,600\,000$
$Var(R) = 180^2$ Var X + 200^2 Var Y = 3 6250 000
$\sigma(R) \approx 6\,020{,}8$

$C = 110X + 100Y + 1\,000\,000$ $\quad E(C) = 3\,000\,000$
$Var(C) = 110^2$ Var X + 100^2 Var Y = 11 562 500
$\sigma(C) \approx 3\,400{,}37$

$M = R - C = 70X + 100Y - 1\,000\,000$
$E(M) = 6\,00\,000$ $\quad$ Var (M) = 7 062 500
$\sigma(M) \approx 2\,657{,}5$

Attention de ne pas écrire V(M) = V(R - C) = V(R) + V(C) !On ne peut rien dire quant à l'indépendance de R et C !

Exercice n° 3

X = U + V	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
p en $\frac{1}{36}$	1	2	3	4	5	6	5	4	3	2	1

Y= **U - V**	**- 5**	**- 4**	**- 3**	**- 2**	**- 1**	**0**	**1**	**2**	**3**	**4**	**5**
p en $\frac{1}{36}$	1	2	3	4	5	6	5	4	3	2	1

p (X = 5 et Y = 0) = 0
Or p (x = 5) = $\frac{4}{36}$ et p(Y = 0) = $\frac{6}{36}$. Donc X et Y ne sont pas indépendants

E(U) = 3,5 $\quad$ Var (U) = 2,916
E(U+V) = 7 $\quad$ E(U-V) = 0
Var (U+V) = 5,833 $\quad$ Var (U-V) = 5,833
X+Y = 2U d'où $\quad$ Var (X+Y) = 4 Var U = 11,666 = Var X + Var Y

DOCUMENT POUR LE PROFESSEUR

COMPLÉMENTS SUR L'INDÉPENDANCE DE VARIABLES ALÉATOIRES

Pour les élèves de T.S. : seule l'indépendance de deux variables aléatoires est explicitement au programme. Pour aborder le théorème de la limite centrée, ou pour démontrer "proprement" la loi suivie par la moyenne $\frac{X_1 + X_2 + \ldots + X_n}{n}$, il est nécessaire de parler de l'indépendance de n variables aléatoires.

Pour les élèves, on "glisse" rapidement de l'indépendance de 2 variables aléatoires à l'indépendance de n variables aléatoires sans soulever de problème.

Pour aller plus loin

• Soient X_1 .....X_k, k variables aléatoires discrètes définies sur le même univers Ω. On dit que ces variables sont mutuellement indépendantes, si et seulement si, quels que soient $\alpha_1 \in X_1(\Omega)$.... $\alpha_k \in X_k(\Omega)$...

$p(X_1 = \alpha_1 \text{ et } X_2 = \alpha_2 \ldots\ldots\ldots \text{ et } X_k = \alpha_k) = p(X_1 = \alpha_1).\, p(X_2 = \alpha_2) \ldots\ldots\ldots p(X_k = \alpha_k)$

- Soit une famille de variables aléatoires mutuellement indépendantes , alors toute sous-famille est formée de variables aléatoires mutuellement indépendantes. C'est la notion de **variables aléatoires mutuellement indépendantes qui sera employée dans la suite du cours**. Ne pas confondre avec la notion de variables aléatoires 2 à 2 indépendantes.

Remarque :

- si n variables aléatoires sont mutuellement indépendantes alors elles sont indépendantes 2 à 2, 3 à 3, ...
- Par contre, si n variables aléatoires sont indépendantes 2 à 2, elles peuvent ne pas être mutuellement indépendantes.

Ex : On lance 2 dés équilibrés successivement.

X_1 est la variable aléatoire qui prend la valeur 1 si le 1er dé amène un nombre pair

X_2 est la variable aléatoire qui prend la valeur 1 si le $2^{ème}$ dé amène un nombre pair

Y est la variable qui prend la valeur 1 si la somme des résultats obtenus est paire.

	résultat 1e lancer	**1**	**2**	**3**	**4**	**5**	**6**
résultat 2e lancer	$X_1 \rightarrow$ $X_2 \downarrow$	0	1	0	1	0	1
1	0	Y = 1		Y = 1		Y = 1	
2	1		Y = 1		Y = 1		Y = 1
3	0	Y = 1		Y = 1		Y = 1	
4	1		Y = 1		Y = 1		Y = 1
5	0	Y = 1		Y = 1		Y = 1	
6	1		Y = 1		Y = 1		Y = 1

X_1 et X_2 sont indépendantes

En effet :

$$p(X_1 = 1 \text{ et } X_2 = 1) = \frac{9}{36} = \frac{1}{4} = \frac{1}{2} x \frac{1}{2} = p(X_1 = 1) \times p(X_2 = 1)$$

$$p(X_1 = 0 \text{ et } X_2 = 0) = \frac{9}{36} = \frac{1}{4} = \frac{1}{2} x \frac{1}{2} = p(X_1 = 0) \times p(X_2 = 0)$$

$$p(X_1 = 0 \text{ et } X_2 = 1) = \frac{9}{36} = \frac{1}{4} = \frac{1}{2} x \frac{1}{2} = p(X_1 = 0) \times p(X_2 = 1)$$

$$p(X_1 = 1 \text{ et } X_2 = 0) = \frac{9}{36} = \frac{1}{4} = \frac{1}{2} x \frac{1}{2} = p(X_1 = 1) \times p(X_2 = 0)$$

On démontre de même que X_1 et Y sont indépendantes

X_2 et Y sont indépendantes

X_1, X_2, Y sont elles mutuellement indépendantes ?

$$p(X_1 = 1 \text{ et } X_2 = 1 \text{ et } Y = 1) = \frac{9}{36} = \frac{1}{4}$$

$$p(X_1 = 1) \times p(X_2 = 1) \times p(Y = 1) = \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8}$$

Donc, X_1, X_2, Y ne sont pas mutuellement indépendantes.

Document pour l'élève

Variable aléatoire continue

Objectif:

Donner un exemple permettant d'introduire les notions de
variable aléatoire continue
fonction densité
fonction de répartition

VARIABLE ALEATOIRE CONTINUE

On vous présente un bout de ficelle de L cm de long, tendu entre deux points A et B. On vous demande de couper ce bout de ficelle à l'endroit de votre choix. On mesure alors la longueur du morceau de ficelle d'extrémité A. On se propose de construire un modèle mathématique permettant d'attribuer une valeur à la probabilité d'événements tels que :

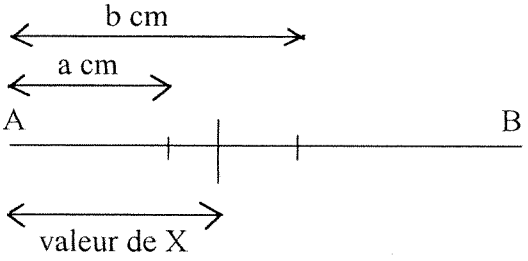
le morceau de ficelle a une longueur inférieure ou égale à...

le morceau de ficelle a une longueur exactement égale à...

Par analogie avec ce qu'on appelle une variable aléatoire discrète, on est amené à introduire une variable aléatoire continue X qui prend pour valeur la longueur du morceau de ficelle en cm.

1°) Premier modèle

Soit a et b deux nombres quelconques fixés vérifiant $0 \leq a \leq b \leq L$. L'événement $a \leq X \leq b$ est réalisé si l'on coupe la ficelle de cette façon :

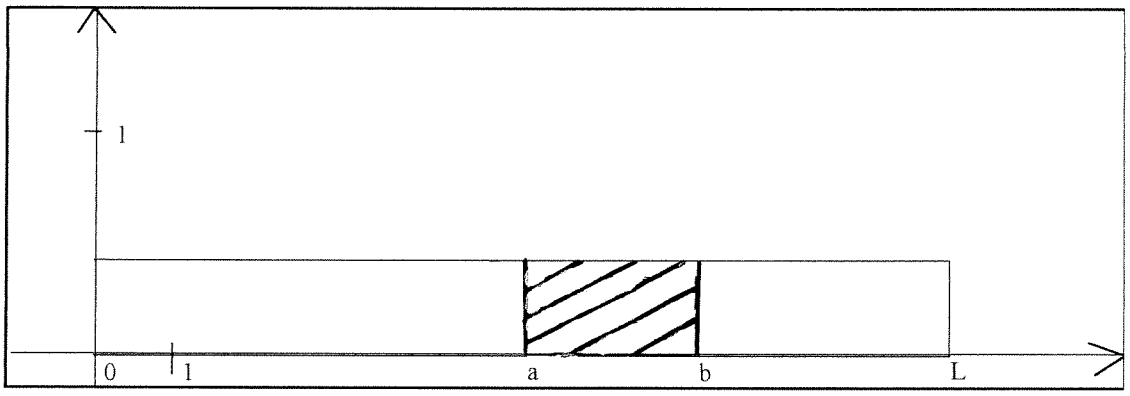


On peut décider, ce qui dans un premier temps est assez « naturel », d'attribuer à l'événement $a \leq X \leq b$ une probabilité proportionnelle à la longueur de l'intervalle [a,b].

♦a) Trouver $p(a \leq X \leq b)$ sachant que $p(a \leq X \leq b)$ est proportionnelle à la longueur de l'intervalle [a,b] et que bien sûr $p(0 \leq X \leq L) = 1$.

♦b) Il peut être intéressant de visualiser $p(a \leq X \leq b)$.

Quelle doit être l'équation $y=f(x)$ du bord supérieur du rectangle ci-dessous pour que $p(a \leq X \leq b)$ soit égale à la mesure de l'aire hachurée en unité d'aire ?



♦ c) Faire une figure pour L=10 (unités : 1 cm pour 1 sur l'axe des x et 10 cm pour 1 sur l'axe des y).

Quelle est la probabilité que

le morceau de ficelle mesure entre 3 et 4 cm ?

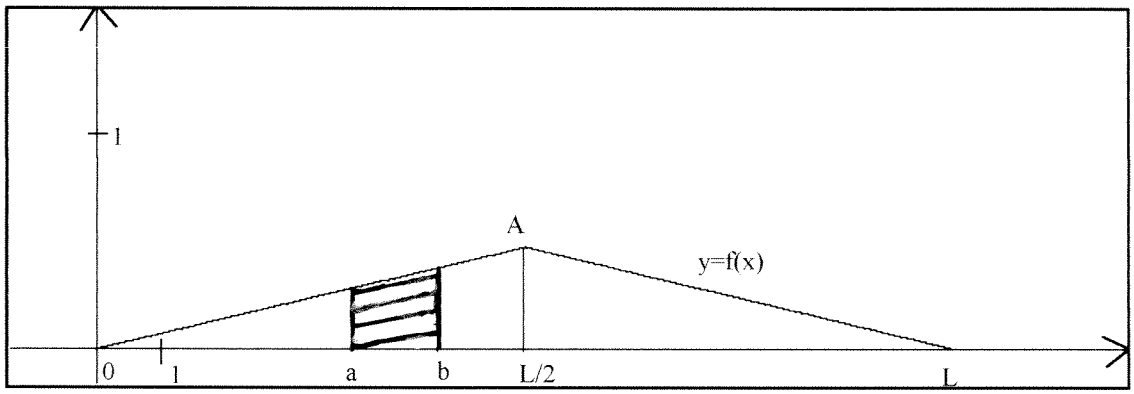
le morceau de ficelle mesure au plus 4 cm ?

le morceau de ficelle mesure exactement 4 cm ?

♦ d) Soit F la fonction qui à x associe p(X≤x) qui, ici, est égal à p(0≤X≤x). Représenter cette fonction dans un repère bien choisi dans le cas où L=10.

2°) Deuxième modèle

On peut supposer qu'on aura plutôt tendance à couper la ficelle dans la région du centre. Pour traduire ce fait on peut par exemple remplacer la figure du 1°)b par une figure telle que p(a≤X≤b) dépende non seulement de la longueur de l'intervalle [a,b] mais aussi de la position de cet intervalle. On peut par exemple envisager une figure de ce type :



avec p(a≤X≤b) égale à la mesure en unité d'aire de l'aire du trapèze hachuré.

♦ a) Trouver quelle doit être l'ordonnée du point A pour que « l'aire sous la courbe » soit égale à l'unité d'aire puis faire un dessin exact dans le cas où L=10 (unités : 1 cm pour 1 sur l'axe des x et 10 cm pour 1 sur l'axe des y).

♦ b) Quelle est la probabilité que

le morceau de ficelle mesure entre 3 et 4 cm ?

le morceau de ficelle mesure au plus 4 cm ?

le morceau de ficelle mesure exactement 4 cm ?

♦ c) L est à nouveau quelconque.

Vérifier que :

$$f(x) = \frac{4x}{L^2} \quad \text{si} \quad 0 \le x \le \frac{L}{2}$$

$$f(x) = -\frac{4x}{L^2} + \frac{4}{L} \quad \text{si} \quad \frac{L}{2} \le x \le L$$

♦ d) On pose $F(x) = p(X \leq x) = p(0 \leq X \leq x)$.

Vérifier que:

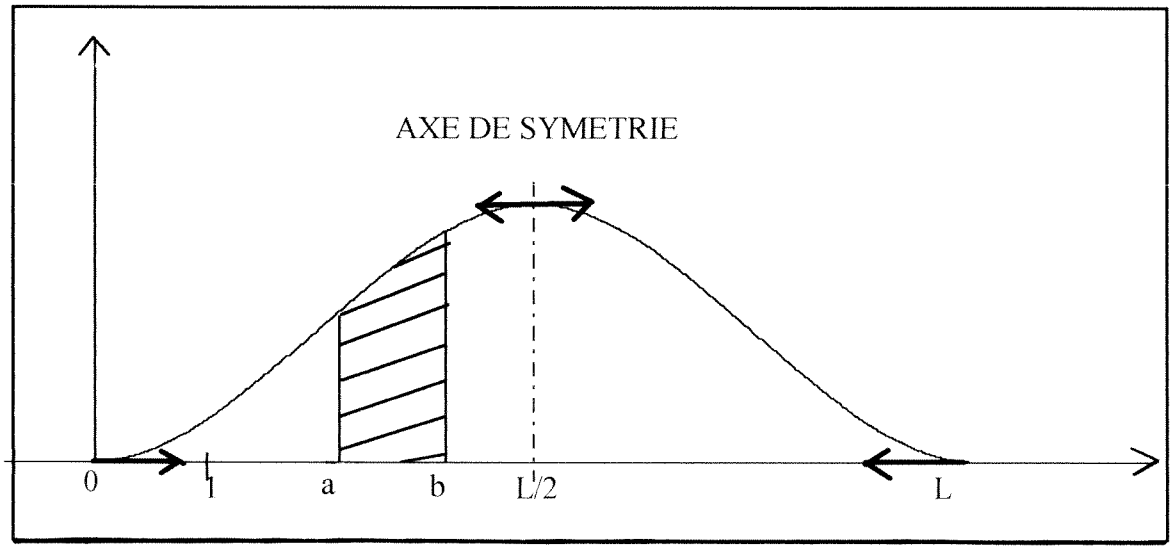
$$F(x) = \frac{2x^2}{L^2} \quad \text{si} \quad 0 \leq x \leq \frac{L}{2}$$

$$F(x) = -\frac{2x^2}{L^2} + \frac{4}{L}x - 1 \quad \text{si} \quad \frac{L}{2} \leq x \leq L$$

♦ e) Représenter la fonction F dans un repère bien choisi dans le cas où L=10.

3°) Troisième modèle

On peut choisir pour f une fonction qui ne soit pas affine par morceaux. On peut par exemple chercher une fonction dont la courbe représentative soit de ce type :



$p(a \leq X \leq b)$ étant toujours égale à la mesure, en unité d'aire, de l'aire de la région hachurée c'est-à-dire à $\int_a^b f(x)dx$.

Pour toute la suite de l'exercice on supposera que L=10.

On peut décider de chercher une fonction f définie sur [0,5] par une formule du type $f(x)=ax^3+bx^2+cx+d$.

♦ a) En tenant compte du fait que :

la courbe C représentant f passe par l'origine

la courbe C admet au point d'abscisse 0 une tangente horizontale

la courbe C admet au point d'abscisse 5 une tangente horizontale

le domaine limité par la courbe C, l'axe des x et les droites d'équations x= 0 et x=5 a une aire égale à 0,5 unité d'aire (car la surface totale « sous la courbe » doit avoir une aire égale à une unité d'aire)

démontrer que $f(x) = -0{,}0032\,x^3 + 0{,}024\,x^2$.

♦ b) Tracer la courbe C représentant f (pour $0 \leq x \leq 10$)

(unités : 1 cm pour 1 sur l'axe des x et 10 cm pour 1 sur l'axe des y).

♦ c) Quelle est la probabilité que

le morceau de ficelle mesure entre 3 et 4 cm ?

le morceau de ficelle mesure au plus 4 cm ?

le morceau de ficelle mesure exactement 4 cm ?

♦ d) On pose $F(x) = p(X \leq x) = p(0 \leq X \leq x)$.

Vérifier que : si $0 \leq x \leq 5$ alors $F(x) = -0{,}0008\, x^4 + 0{,}008\, x^3$.

Expliquer pourquoi si $5 \leq x \leq 10$ alors $F(x) = 1 - F(10-x)$.

♦ e) Représenter la fonction F dans un repère bien choisi.

CONCLUSION :

1°) X est une variable aléatoire qui peut prendre pour valeur n'importe quel nombre de l'intervalle [0,L]. On dit que c'est une variable aléatoire continue.

Ce qui est à noter c'est que, pour tout a de l'intervalle [0,L], on a $p(X=a)=0$ (la probabilité que le morceau de ficelle mesure **exactement** a cm est nulle).

Pour définir la loi de la variable X, on a donc été amené à s'intéresser aux intervalles de type [a,b], avec $0 \leq a \leq b \leq L$, et à définir $p(a \leq X \leq b)$.

Pour cela on a introduit une fonction f définie sur [0,L] ayant les propriétés suivantes :

pour tout x de [0,L] , $f(x) \geq 0$

l'aire du domaine limité par la courbe C représentant f, l'axe des x et les droites d'équations x=0 et x=L est égale à une unité d'aire.

On a alors défini $p(a \leq X \leq b)$ par :

$p(a \leq X \leq b) = \int_a^b f(x)dx$ (mesure, en unité d'aire, de l'aire du domaine limité par la courbe C, l'axe des x et les droites d'équations x=a et x=b).

Remarque :

$p(x \leq X \leq x+\Delta x)$ est égale à la mesure de l'aire d'un petit domaine qui peut être confondue, si Δx est « près de 0 », avec l'aire d'un petit rectangle de largeur Δx et de hauteur f(x). Donc, si Δx est « près de 0 », $p(x \leq X \leq x + \Delta x) \approx f(x)\Delta x$ soit

$$f(x) \approx \frac{p(x \leq X \leq x + \Delta x)}{\Delta x}$$

(qu'on peut considérer comme une « densité moyenne de probabilité » sur l'intervalle $[x,x+\Delta x]$ de longueur Δx).

En fait :

$$f(x) = \lim_{\Delta x \to 0}\left(\frac{p(x \le X \le x + \Delta x)}{\Delta x}\right)$$

ce qui explique que f(x) soit appelée densité de probabilité en x.

Il faut bien noter que f(x) n'est pas une probabilité mais que la fonction f sert à définir des probabilités (la probabilité que X appartienne à un petit intervalle de longueur Δx étant sensiblement égale à $f(x).\ \Delta x$).

2°) Pour tout x de [0,L], on a aussi défini $F(X) = p(X \le x)$.
La fonction F est appelée fonction de répartition de la variable aléatoire X.

Comme $p(a \le X \le b) = p(X \le b) - p(X < a) = p(X \le b) - p(X \le a)$ car $p(X = a) = 0$ on a donc :
$p(a \le X \le b) = F(b) - F(a)$.

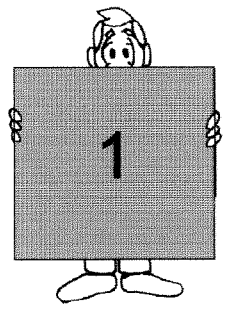
F est une primitive de f (pour notre exemple $F(x) = \int_0^x f(t)dt$ et F est donc ici la primitive de f qui s'annule pour x=0). On peut d'ailleurs retrouver cette propriété en remarquant que

$$f(x) = \lim_{\Delta x \to 0}\left(\frac{F(x + \Delta x) - F(x)}{\Delta x}\right) = F'(x)$$

.

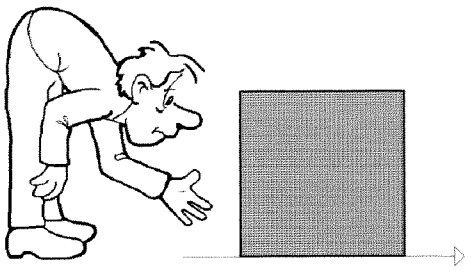
En définitive, une variable aléatoire continue X est définie soit par la donnée de sa fonction densité f soit par la donnée de sa fonction de répartition F et : $p(a \le X \le b) = \int_a^b f(x)dx = F(b) - F(a)$.

Remarque :
pour avoir une idée de la signification de f et F, on peut retenir que, grossièrement, les intervalles auxquels X a le plus de chance d'appartenir sont ceux sur lesquels f(x) est élevé ou, ce qui revient au même, ceux sur lesquels F(x) augmente rapidement.

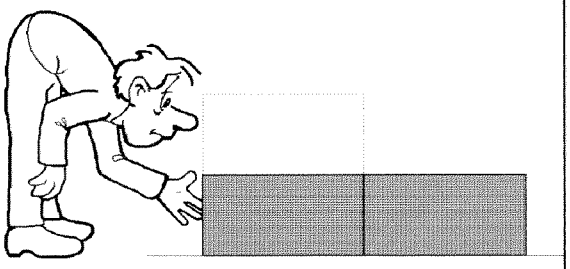


Voilà, j'ai ma probabilité à répartir sur un intervalle !

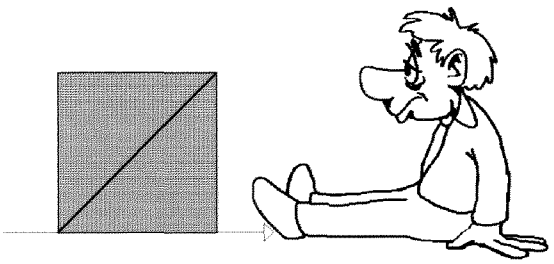
Le plus simple serait un intervalle de longueur 1



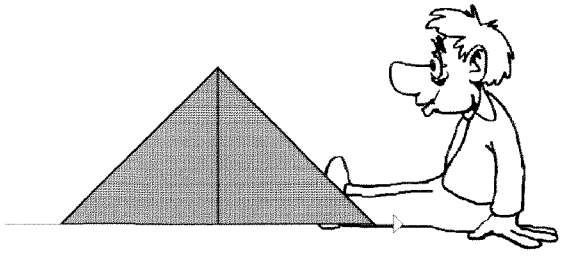
Ou à la rigueur de longueur 2



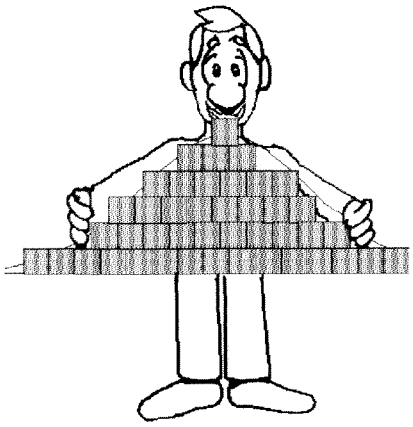
Moins uniforme c'est plus compliqué



Mais tellement plus beau !



Et là c'est presque normal !



Document pour l'élève

Ce document est précédé de quelques commentaires sur la convergence destinés aux enseignants.

Diverses approximations de la loi binomiale

Objectif:

Aborder l'approximation de la loi binomiale par les lois de Poisson et Gauss sous le double aspect numérique et graphique.

888888888888888888

Loi des grands nombres

Approximation de la loi binomiale par les lois de Poisson et de Gauss

Commentaires concernant les difficultés de la notion mathématique à exposer :

Les démonstrations de la convergence d' une loi binomiale vers une loi de Poisson ou vers une loi de Gauss font intervenir des méthodes avancées de l' analyse : formule de Stirling, développements limités , et la caractérisation de la convergence en loi utilisant les fonctions caractéristiques.

Ces démonstrations apparaissent donc hors des ambitions d' un enseignement en classe de B.T.S..

La mise en oeuvre d' un support visuel (ici des comparaisons de courbes que l' étudiant peut lui même réaliser à l' aide d' un logiciel) semble par contre pouvoir emporter la conviction des élèves quant au fait que telle loi binomiale pourra commodément être remplacée par une loi normale ou une loi de Poisson, dans des conditions qu' ils peuvent être à même de découvrir.

Une telle démarche devrait bien convenir à des élèves qui sont surtout de futurs " utilisateurs ", et non de futurs " théoriciens ".

Commentaires théoriques :

Rappelons à ce sujet les critères de convergence en loi :

Une suite (X_n) de variables aléatoires converge en loi vers une variable X si une des conditions suivantes est réalisée :

- la suite (F_n) des fonctions de répartition correspondantes converge vers la fonction de répartition F de X en tout point de continuité de F,
  ou si quels que soient les points a et b de continuité de F :
  $P(a < X_n < b) \underset{n}{\rightarrow} P(a < X < b)$,
- si les variables aléatoires X_n et si la variable X prennent leurs valeurs dans un même ensemble discret $\mathcal{D}$:
  $\forall d \in \mathcal{D}$: $P(X_n = d) \underset{n}{\rightarrow} P(X = d)$.
- si les variables aléatoires X_n et si la variable aléatoire X ont pour densités les fonctions f_n et la fonction f :
  la suite $(f_n(x))$ converge vers $f(x)$ pour toute valeur de x, à l' exception éventuelle d' un ensemble dénombrable de telles valeurs.

Les deux premières conditions sont en fait des conditions nécessaires et suffisantes.

Le support visuel que nous comptons mettre en oeuvre trouve sa justification dans la troisième condition.

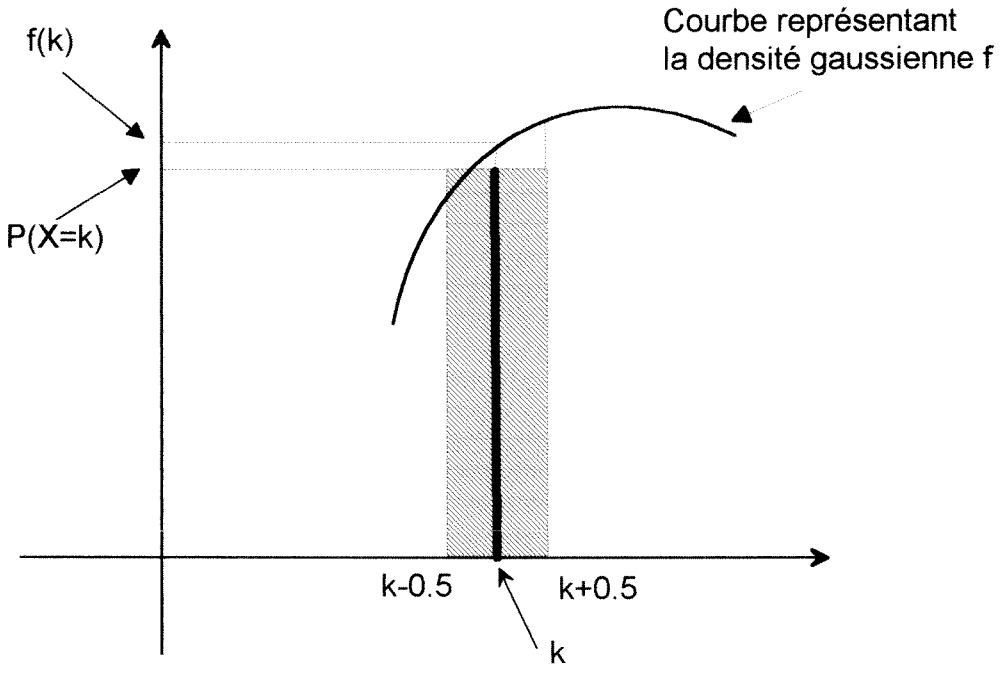
Objectifs pédagogiques :

Dans une première activité, nous rendrons les élèves attentifs au fait que l' on peut, sous certaines conditions, approcher une v.a. X de loi binomiale de paramètres n, p par une v.a. Y de loi normale de paramètres $m = n.p$, $\sigma = \sqrt{n.p.(1-p)}$, et de densité :

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

Nous nous bornerons à faire remarquer aux élèves, pour des entiers k, a, b :
$P(X = k) = P(k - 0.5 < X < k + 0.5) \approx P(k - 0.5 < Y < k + 0.5)$,
et :
$P(a \leq X \leq b) = P(a - 0.5 \leq X \leq b + 0.5) \approx P(a - 0.5 \leq Y \leq b + 0.5)$

La première comparaison pourra être effectuée en comparant l' aire du rectangle hachuré et l' aire de la partie de plan située entre la courbe représentant f et l' axe des abscisses, ce entre les abscisses $k - 0.5$ et $k + 0.5$.



Nous n' aurons pas ainsi mis en évidence de notion de convergence, mais fait sentir pourquoi les variables X et Y sont équivalentes, au sens où leur probabilité d' appartenance à certains intervalles sont sensiblement les mêmes.

Dans une seconde activité, nous suggérons que l' approximation de la loi binomiale par une loi normale n' est pas toujours la meilleure. Il est en effet possible de démontrer que des distributions binomiales peuvent dans certaines conditions converger en loi vers une distribution de Poisson.

Pour mesurer les écarts entre les distributions envisagées :

binomiale ⇔ Gauss et *binomiale ⇔ Poisson*,

l' élève devra préciser un outil . La somme des valeurs absolues des écarts est ainsi évoquée pour mesurer une approximation globale.

Une troisième activité propose une série de cas où un élève s' aidant de tableaux de nombres et de graphiques doit choisir une approximation. Certains cas sont de décision difficile, l' approximation de la binomiale par l' une des lois n' apparaissant pas plus justifiée que son approche par l' autre. Il y a lieu d' autre part d' être critique quant à la validité des conclusions apportées : La mesure brute des différents écarts : $\left| B_{n,p}(k) - P_{\lambda}(k) \right|$ ou $\left| B_{n,p}(k) - f(k) \right|$ masque le fait que ces écarts n' ont pas tous la même portée suivant les valeurs de k : ces écarts pourraient être pondérés par les probabilités des ensembles sur lesquels ils portent.

Préliminaires

Dans cette fiche, les notations suivantes seront utilisées :

$B_{n,p}$ désigne la distribution binomiale de paramètres n,p, où p est la probabilité d' un succès, et n le nombre de répétitions de l' épreuve.
$B_{n,p}(k)$ désigne donc la probabilité d' obtenir k succès lors d' une série de n répétitions d' une expérience, X désignant le nombre de ces succès.
$B_{n,p}(k) = P(X = k) = C_n^k \times p^k \times (1-p)^{n-k}$

Cette distribution binomiale sera ici représentée par une fonction en escalier valant $B_{n,p}(k)$ sur l' intervalle$[k-0.5, k+0.5[$.

La fonction f désigne la densité de la loi gaussienne de paramètres $m = n.p$ et $\sigma = \sqrt{n.p.(1-p)}$. Nous avons donc :
$f(x) = \frac{1}{\sigma.\sqrt{2\pi}}.e^{-\frac{(x-m)^2}{2\sigma^2}}$.

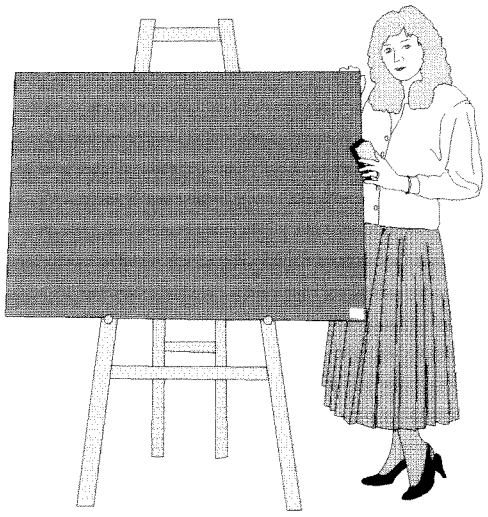
Enfin, P_λ désigne la distribution de Poisson de paramètre $\lambda = n.p$: $P_\lambda(k) = \frac{\lambda^k}{k!}.e^{-\lambda}$.

Cette distribution sera représentée par une fonction étagée prenant la valeur $P_\lambda(k)$ sur l' intervalle $[k-0.5, k+0.5[$.

Attention !

Revoir :

- *les notions de v.a. discrète et continue*
- *les lois binomiale, de Gauss, et de Poisson*
- *l' utilisation d' une table de loi normale*



Première activité

Une urne contient cent boules, quarante sont noires, et soixante sont blanches. On tire au hasard et avec remise trente fois une de ces boules. Soit X la variable aléatoire représentant le nombre de boules noires obtenues au cours d' une telle expérence.

Préciser la loi suivie par X, et établir un tableau comme ci-dessous pour les valeurs de k : 0; 1; 2.....17,

k	P(X=k)

déterminer l' espérance et l' écart-type de cette distribution,

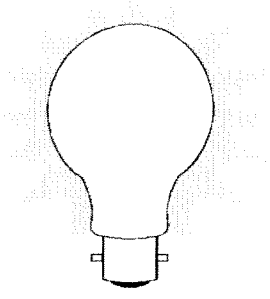
calculer : $P(7 \leq X \leq 15)$.

Sur le graphique de la page suivante la distribution de X a été représentée par une fonction en escalier.
Représenter aussi cette distribution par un diagramme en bâtons.
La courbe continue représente la densité f de la loi normale de paramètres $m = n.p$ et $\sigma = \sqrt{n.p.(1-p)}$ correspondant à $n = 30$ et $p = 0.4$.

Constater que $P(X = k)$ peut être approché par $\int_{k-0.5}^{k+0.5} f(t)dt$, et pourquoi $P(7 \leq X \leq 15)$ peut être approché par $\int_{7-0.5}^{15+0.5} f(t)dt$.

En résumé :

Nous dirons que la v.a. X qui suit une loi binomiale a été approchée par une v.a. Y gaussienne, les v.a. X et Y ayant des probabilités d' appartenance à un même intervalle sensiblement égales.

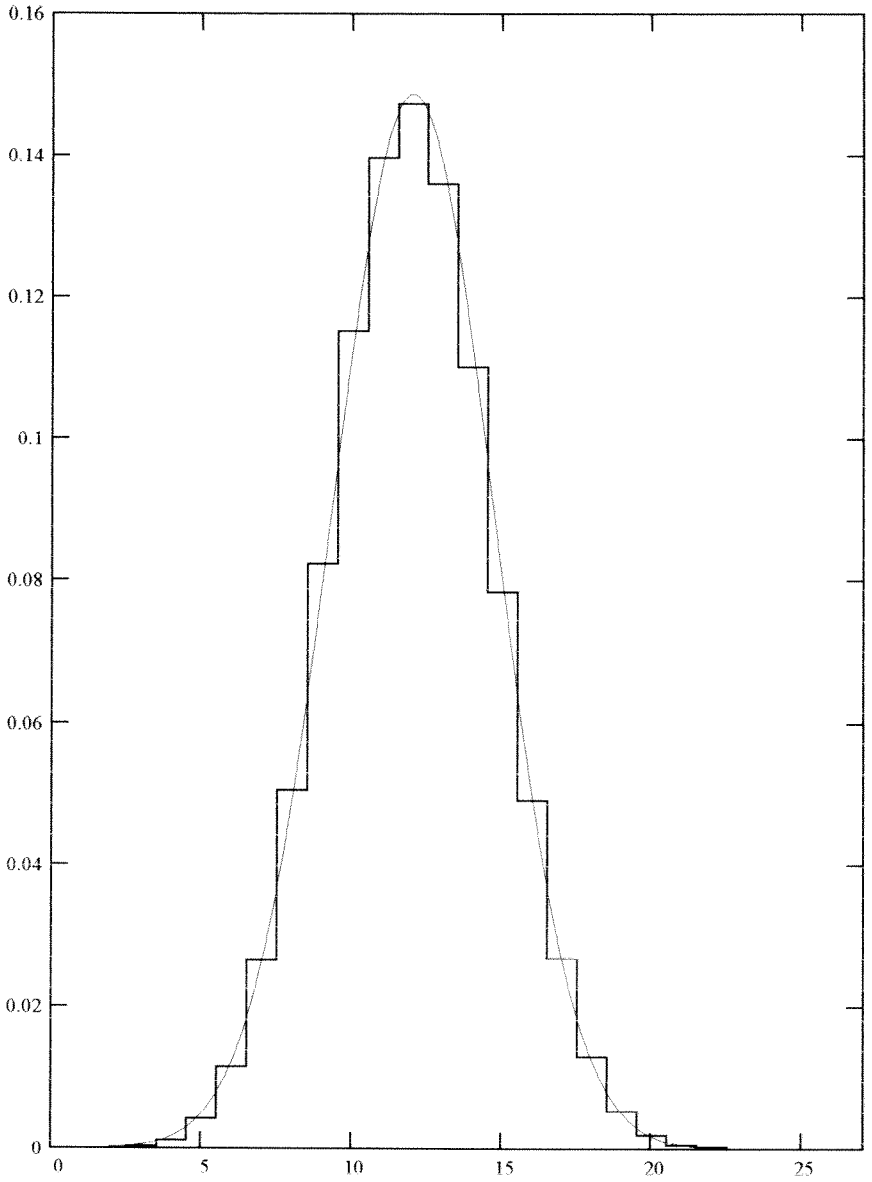


Voici une représentation de l' approximation de la loi binomiale obtenue pour :

$$n = 30 \qquad \text{et} \qquad p = 0.4$$

par une loi gaussienne $N(m, \sigma)$ avec :

$$m = 12 \qquad \text{et} \qquad \sigma = 2.683$$



Deuxième activité

Une urne contient cent boules, cinq sont noires, et quatre-vingt quinze sont blanches. On en tire successivement vingt avec remise. On désigne par X la v.a. représentant le nombre de boules noires obtenue lors d' une série de vingt tirages.

Préciser la loi, l' espérance et l' écart-type de X.

On donne le tableau suivant précisant les valeurs de $B(k) = B_{n,p}(k), f(k)$, et $P(k) = P_{\lambda}(k)$. Compléter ce tableau.

k	B(k)	f(k)	P(k)	|B(k) − f(k)|	|B(k) − P(k)|
0	0.3584859	0.2418089	0.3678794	0.1166771	0.0093935
1	0.3773536	0.4093061	0.3678794	0.0319525	0.0094742
2	0.1886768	0.2418089	0.1839397	0.0531321	0.0047371
3	0.0595821	0.0498591	0.0613132	0.0097231	0.0017311
4	0.0133276	0.0035881	0.0153283	0.0097395	0.0020007
5	0.0022446	9.012215610^{-5}	0.0030657	0.0021545	8.2101610^{-4}
6	2.953481610^{-4}	7.900365510^{-7}	5.109436710^{-4}	2.945581210^{-4}	2.155955110^{-4}
		Totaux :			

Préciser quelle loi, loi de Poisson ou loi normale vous semble à l' examen de ce tableau la plus apte à approcher la loi binomiale pour la détermination de $P(0 < X < 6)$.

Commentaires :

Voici le tableau complet des valeurs des différentes lois pour k = 0 ; 1 ;...20

k	B(k)	f(k)	P(k)	|B(k) - f(k)|	|B(k) - P(k)|
0	0.358486	0.241809	0.367879	0.116677	0.009394
1	0.377354	0.409306	0.367879	0.031953	0.009474
2	0.188677	0.241809	0.18394	0.053132	0.004737
3	0.059582	0.049859	0.061313	0.009723	0.001731
4	0.013328	0.003588	0.015328	0.009739	0.002001
5	0.002245	$9.012216 \cdot 10^{-5}$	0.003066	0.002155	$8.21016 \cdot 10^{-4}$
6	$2.953482 \cdot 10^{-4}$	$7.900366 \cdot 10^{-7}$	$5.109437 \cdot 10^{-4}$	$2.945581 \cdot 10^{-4}$	$2.155955 \cdot 10^{-4}$
7	$3.108928 \cdot 10^{-5}$	$2.41719 \cdot 10^{-9}$	$7.299195 \cdot 10^{-5}$	$3.108686 \cdot 10^{-5}$	$4.190267 \cdot 10^{-5}$
8	$2.658952 \cdot 10^{-6}$	$2.581203 \cdot 10^{-12}$	$9.123994 \cdot 10^{-6}$	$2.658949 \cdot 10^{-6}$	$6.465043 \cdot 10^{-6}$
9	$1.865931 \cdot 10^{-7}$	0	$1.013777 \cdot 10^{-6}$	$1.865931 \cdot 10^{-7}$	$8.27184 \cdot 10^{-7}$
10	$1.080276 \cdot 10^{-8}$	0	$1.013777 \cdot 10^{-7}$	$1.080276 \cdot 10^{-8}$	$9.057495 \cdot 10^{-8}$
11	$5.168784 \cdot 10^{-10}$	0	$9.216156 \cdot 10^{-9}$	$5.168784 \cdot 10^{-10}$	$8.699277 \cdot 10^{-9}$
12	$2.040309 \cdot 10^{-11}$	0	$7.68013 \cdot 10^{-10}$	$2.040309 \cdot 10^{-11}$	$7.476099 \cdot 10^{-10}$
13	$6.608289 \cdot 10^{-13}$	0	$5.907792 \cdot 10^{-11}$	$6.608289 \cdot 10^{-13}$	$5.841709 \cdot 10^{-11}$
14	$1.739024 \cdot 10^{-14}$	0	$4.219851 \cdot 10^{-12}$	$1.739024 \cdot 10^{-14}$	$4.202461 \cdot 10^{-12}$
15	0	0	$2.813234 \cdot 10^{-13}$	0	$2.809573 \cdot 10^{-13}$
16	0	0	$1.758271 \cdot 10^{-14}$	0	$1.757669 \cdot 10^{-14}$
17	0	0	$1.034277 \cdot 10^{-15}$	0	$1.034203 \cdot 10^{-15}$
18	0	0	0	0	0
19	0	0	0	0	0
20	0	0	0	0	0

Nous avons :

$$\sum |B(k) - f(k)| = 0.22370 \quad \text{et} \quad \sum |B(k) - P(k)| = 0.02842.$$

La loi la plus apte à approcher la loi binomiale est ici la loi de Poisson.

Notons que la somme des distances observées entre les termes $B(k)$ et $f(k)$ d' une part, et les termes $B(k)$ et $P(k)$ d' autre part résulte essentiellement des termes correspondant aux grandes valeurs de $B(k)$, comme le montrent les valeurs :

$$\sum_{k=0}^{k=6} |B(k) - f(k)| = 0.22367 \quad \text{et} \quad \sum_{k=0}^{k=6} |B(k) - P(k)| = 0.02837 .$$

Troisième activité

Nous avons vu précédemment que la loi binomiale peut dans certains cas utilement être remplacée par une loi de Poisson ou une loi de Gauss.

Les résultats de la théorie sont les suivants :

Théorème 1 :

Soit (X_n) une suite de v.a. définies sur un même espace probabilisé, chacune suivant une loi binomiale de paramètres n et p_n.
Si $(n \times p_n)$ converge vers le réel λ, alors la suite (X_n) converge en loi vers une v.a. de loi de Poisson de paramètre λ.

En pratique : une v.a. X suivant une loi binomiale de paramètres n assez grand et p petit pourrait être approchée par une v.a. Y de loi de Poisson de paramètre $\lambda = n \times p$.

Théorème 2 : *(Th. central limite)*

Soit (X_n) une suite de v.a. définies sur un même espace probabilisé, chacune suivant une loi binomiale de paramètres n et p (*p ici ne varie pas*), la suite $\left[\frac{X_n - n \times p}{\sqrt{n \times p \times (1-p)}} \right]$ converge en loi vers une v.a. Y normale, centrée et réduite.

En pratique : une v.a. X suivant une loi binomiale de paramètres n assez grand et p non voisin de 0 ou 1 pourrait être approchée par une v.a. Y suivant une loi normale de paramètres $m = n \times p$ et $\sigma = \sqrt{n \times p \times (1-p)}$.

Le problème auquel se trouve donc confronté l' utilisateur est le suivant :

Donner des critères de décision permettant de choisir telle ou telle approximation.

Nous allons essayer de faire progresser ce problème sur les exemples qui suivent, nous préciserons ensuite ces critères.

Première question :

A l' examen des tableaux qui suivent et illustrant les cas :

$n = 30$; $p = 0.4$
$n = 30$; $p = 0.2$
$n = 10$; $p = 0.2$

préciser quelle est l' approximation qui vous semble la mieux adaptée.

Pour $n = 30$ et $p = 0.4$ nous obtenons :

k	B(k)	f(k)	P(k)	|B(k) − f(k)|	|B(k) − P(k)|
0	$2.210739 \cdot 10^{-7}$	$6.749926 \cdot 10^{-6}$	$6.144212 \cdot 10^{-6}$	$6.528852 \cdot 10^{-6}$	$5.923138 \cdot 10^{-6}$
1	$4.421478 \cdot 10^{-6}$	$3.333986 \cdot 10^{-5}$	$7.373055 \cdot 10^{-5}$	$2.891838 \cdot 10^{-5}$	$6.930907 \cdot 10^{-5}$
2	$4.274096 \cdot 10^{-5}$	$1.43321 \cdot 10^{-4}$	$4.423833 \cdot 10^{-4}$	$1.005801 \cdot 10^{-4}$	$3.996423 \cdot 10^{-4}$
3	$2.659437 \cdot 10^{-4}$	$5.36213 \cdot 10^{-4}$	0.00177	$2.702693 \cdot 10^{-4}$	0.001504
4	0.001197	0.001746	0.005309	$5.492607 \cdot 10^{-4}$	0.004112
5	0.004149	0.004948	0.012741	$7.993519 \cdot 10^{-4}$	0.008592
6	0.011524	0.012204	0.025481	$6.799235 \cdot 10^{-4}$	0.013957
7	0.026341	0.026198	0.043682	$1.435643 \cdot 10^{-4}$	0.017341
8	0.050487	0.048943	0.065523	0.001544	0.015036
9	0.082275	0.079581	0.087364	0.002694	0.005089
10	0.115185	0.112618	0.104837	0.002568	0.010348
11	0.139619	0.138703	0.114368	$9.160776 \cdot 10^{-4}$	0.025251
12	0.147375	0.148677	0.114368	0.001302	0.033007
13	0.136039	0.138703	0.10557	0.002664	0.030468
14	0.110127	0.112618	0.090489	0.002491	0.019638
15	0.078312	0.079581	0.072391	0.001269	0.005921
16	0.048945	0.048943	0.054293	$1.702019 \cdot 10^{-6}$	0.005348
17	0.026872	0.026198	0.038325	$6.743069 \cdot 10^{-4}$	0.011453
18	0.012938	0.012204	0.02555	$7.341396 \cdot 10^{-4}$	0.012612
19	0.005448	0.004948	0.016137	$4.996276 \cdot 10^{-4}$	0.010689
20	0.001997	0.001746	0.009682	$2.514831 \cdot 10^{-4}$	0.007685
21	$6.34124 \cdot 10^{-4}$	$5.36213 \cdot 10^{-4}$	0.005533	$9.791099 \cdot 10^{-5}$	0.004898
22	$1.729429 \cdot 10^{-4}$	$1.43321 \cdot 10^{-4}$	0.003018	$2.962189 \cdot 10^{-5}$	0.002845
23	$4.01027 \cdot 10^{-5}$	$3.333986 \cdot 10^{-5}$	0.001574	$6.762843 \cdot 10^{-6}$	0.001534
24	$7.797748 \cdot 10^{-6}$	$6.749926 \cdot 10^{-6}$	$7.87246 \cdot 10^{-4}$	$1.047822 \cdot 10^{-6}$	$7.794482 \cdot 10^{-4}$
25	$1.24764 \cdot 10^{-6}$	$1.189366 \cdot 10^{-6}$	$3.778781 \cdot 10^{-4}$	$5.827369 \cdot 10^{-8}$	$3.766304 \cdot 10^{-4}$
26	$1.599538 \cdot 10^{-7}$	$1.823952 \cdot 10^{-7}$	$1.744053 \cdot 10^{-4}$	$2.244138 \cdot 10^{-8}$	$1.742453 \cdot 10^{-4}$
27	$1.579791 \cdot 10^{-8}$	$2.434403 \cdot 10^{-8}$	$7.751345 \cdot 10^{-5}$	$8.546127 \cdot 10^{-9}$	$7.749765 \cdot 10^{-5}$
28	$1.128422 \cdot 10^{-9}$	$2.827828 \cdot 10^{-9}$	$3.322005 \cdot 10^{-5}$	$1.699406 \cdot 10^{-9}$	$3.321892 \cdot 10^{-5}$
29	$5.188147 \cdot 10^{-11}$	$2.858873 \cdot 10^{-10}$	$1.374623 \cdot 10^{-5}$	$2.340058 \cdot 10^{-10}$	$1.374618 \cdot 10^{-5}$
30	$1.152922 \cdot 10^{-12}$	$2.515463 \cdot 10^{-11}$	$5.498491 \cdot 10^{-6}$	$2.400171 \cdot 10^{-11}$	$5.49849 \cdot 10^{-6}$

Pour $n = 30$ et $p = 0.2$ nous obtenons :

k	B(k)	f(k)	P(k)	|B(k) − f(k)|	|B(k) − P(k)|
0	0.001238	0.004282	0.002479	0.003044	0.001241
1	0.009285	0.013468	0.014873	0.004184	0.005588
2	0.033656	0.034393	0.044618	$7.361291 \cdot 10^{-4}$	0.010961
3	0.078532	0.071308	0.089235	0.007224	0.010703
4	0.132522	0.120042	0.133853	0.01248	0.00133
5	0.172279	0.164078	0.160623	0.008201	0.011656
6	0.179457	0.182091	0.160623	0.002634	0.018834
7	0.153821	0.164078	0.137677	0.010257	0.016144
8	0.110559	0.120042	0.103258	0.009483	0.007301
9	0.067564	0.071308	0.068838	0.003744	0.001275
10	0.035471	0.034393	0.041303	0.001078	0.005832
11	0.016123	0.013468	0.022529	0.002655	0.006406
12	0.006382	0.004282	0.011264	0.0021	0.004882
13	0.002209	0.001106	0.005199	0.001104	0.00299
14	$6.706437 \cdot 10^{-4}$	$2.317357 \cdot 10^{-4}$	0.002228	$4.38908 \cdot 10^{-4}$	0.001557
15	$1.788383 \cdot 10^{-4}$	$3.943935 \cdot 10^{-5}$	$8.912556 \cdot 10^{-4}$	$1.39399 \cdot 10^{-4}$	$7.124172 \cdot 10^{-4}$
16	$4.191523 \cdot 10^{-5}$	$5.449901 \cdot 10^{-6}$	$3.342208 \cdot 10^{-4}$	$3.646533 \cdot 10^{-5}$	$2.923056 \cdot 10^{-4}$
17	$8.629607 \cdot 10^{-6}$	$6.114619 \cdot 10^{-7}$	$1.179603 \cdot 10^{-4}$	$8.018145 \cdot 10^{-6}$	$1.093307 \cdot 10^{-4}$
18	$1.558123 \cdot 10^{-6}$	$5.570218 \cdot 10^{-8}$	$3.93201 \cdot 10^{-5}$	$1.502421 \cdot 10^{-6}$	$3.776197 \cdot 10^{-5}$
19	$2.460195 \cdot 10^{-7}$	$4.119998 \cdot 10^{-9}$	$1.241687 \cdot 10^{-5}$	$2.418995 \cdot 10^{-7}$	$1.217085 \cdot 10^{-5}$
20	$3.382768 \cdot 10^{-8}$	$2.474251 \cdot 10^{-10}$	$3.725062 \cdot 10^{-6}$	$3.358025 \cdot 10^{-8}$	$3.691234 \cdot 10^{-6}$
21	$4.027105 \cdot 10^{-9}$	$1.206459 \cdot 10^{-11}$	$1.064303 \cdot 10^{-6}$	$4.01504 \cdot 10^{-9}$	$1.060276 \cdot 10^{-6}$
22	$4.11863 \cdot 10^{-10}$	$4.776428 \cdot 10^{-13}$	$2.902646 \cdot 10^{-7}$	$4.113853 \cdot 10^{-10}$	$2.898527 \cdot 10^{-7}$
23	$3.581417 \cdot 10^{-11}$	$1.53538 \cdot 10^{-14}$	$7.572119 \cdot 10^{-8}$	$3.579882 \cdot 10^{-11}$	$7.568538 \cdot 10^{-8}$
24	$2.61145 \cdot 10^{-12}$	0	$1.89303 \cdot 10^{-8}$	$2.611049 \cdot 10^{-12}$	$1.892769 \cdot 10^{-8}$
25	$1.56687 \cdot 10^{-13}$	0	$4.543271 \cdot 10^{-9}$	$1.566785 \cdot 10^{-13}$	$4.543115 \cdot 10^{-9}$
26	$7.533029 \cdot 10^{-15}$	0	$1.048447 \cdot 10^{-9}$	$7.532883 \cdot 10^{-15}$	$1.04844 \cdot 10^{-9}$
27	0	0	$2.329883 \cdot 10^{-10}$	0	$2.32988 \cdot 10^{-10}$
28	0	0	$4.992606 \cdot 10^{-11}$	0	$4.992605 \cdot 10^{-11}$
29	0	0	$1.032953 \cdot 10^{-11}$	0	$1.032953 \cdot 10^{-11}$
30	0	0	$2.065906 \cdot 10^{-12}$	0	$2.065906 \cdot 10^{-12}$

Pour $n = 10$ et $p = 0.2$ nous obtenons :

k	B(k)	f(k)	P(k)	$\lvert B(k) - f(k) \rvert$	$\lvert B(k) - P(k) \rvert$
0	0.107374	0.090361	0.135335	0.017013	0.027961
1	0.268435	0.230745	0.270671	0.03769	0.002235
2	0.30199	0.315392	0.270671	0.013402	0.031319
3	0.201327	0.230745	0.180447	0.029419	0.02088
4	0.08808	0.090361	0.090224	0.002281	0.002143
5	0.026424	0.018941	0.036089	0.007483	0.009665
6	0.005505	0.002125	0.01203	0.00338	0.006525
7	$7.86432 \cdot 10^{-4}$	$1.276217 \cdot 10^{-4}$	0.003437	$6.588103 \cdot 10^{-4}$	0.002651
8	$7.3728 \cdot 10^{-5}$	$4.102392 \cdot 10^{-6}$	$8.592716 \cdot 10^{-4}$	$6.962561 \cdot 10^{-5}$	$7.855436 \cdot 10^{-4}$
9	$4.096 \cdot 10^{-6}$	$7.058555 \cdot 10^{-8}$	$1.909493 \cdot 10^{-4}$	$4.025414 \cdot 10^{-6}$	$1.868533 \cdot 10^{-4}$
10	$1.024 \cdot 10^{-7}$	$6.500705 \cdot 10^{-10}$	$3.818985 \cdot 10^{-5}$	$1.017499 \cdot 10^{-7}$	$3.808745 \cdot 10^{-5}$

Deuxième question :
Essayez de répondre à la question précédente par simple lecture graphique.
Quels sont les cas de décision difficile ?

Parmi les critères souvent retenus, on note :

Pour $n \geq 30$ et p voisin de $\frac{1}{2}$ on accepte l' approximation par la loi normale,

pour $n \geq 30$ et $p \leq \frac{1}{10}$ on accepte l' approximation par la loi de Poisson,

pour $n \geq 50$ et $n \cdot p \geq 10$ et $n \cdot q \geq 10$ on accepte l' approximation par la loi normale.

Voici une représentation de l' approximation de la loi binomiale obtenue pour :

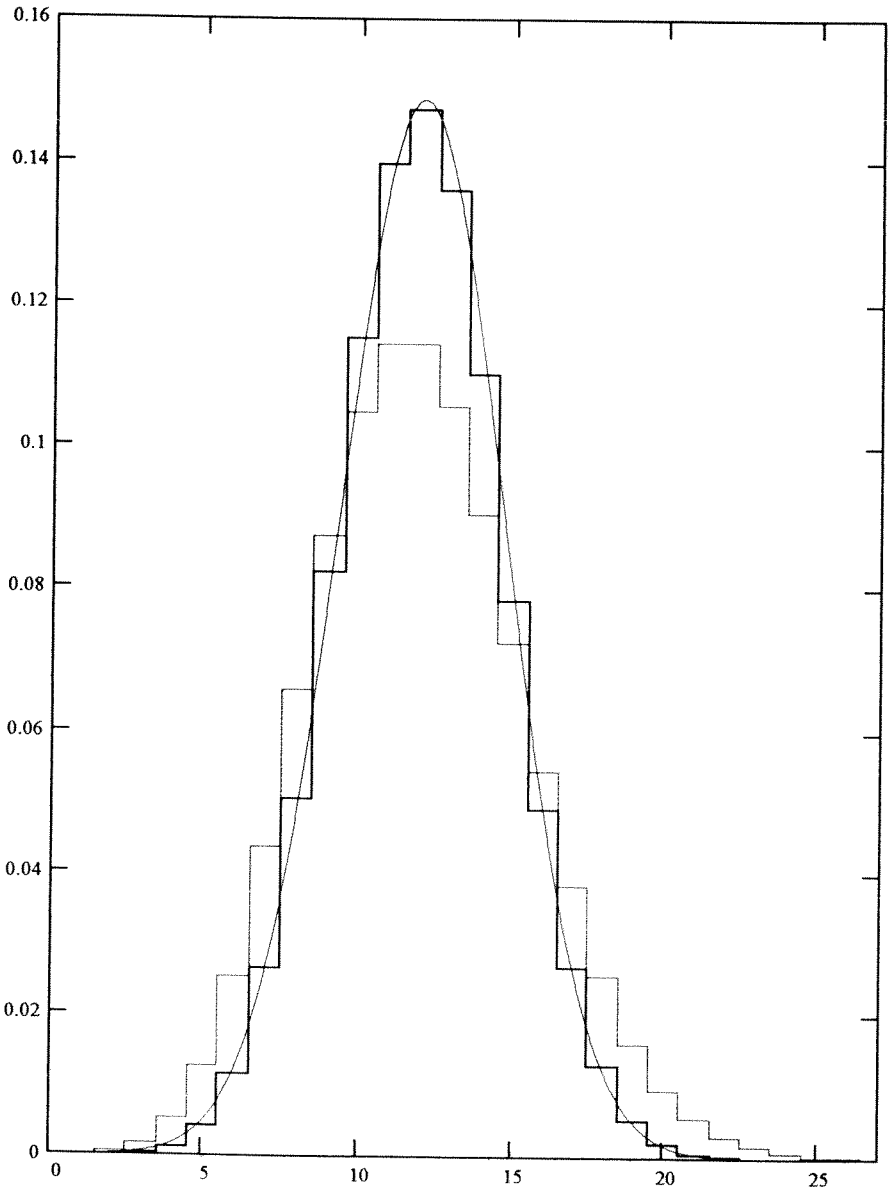
$n = 30$ et $p = 0.4$ (traits pleins)

par une loi gaussienne $N(m, \sigma)$ avec :

$m = 12$ et $\sigma = 2.683$ (courbe)

et une représentation de l' approximation de cette même loi binomiale par une loi de Poisson de paramètre

$\lambda = 12$ (traits fins)



Voici une représentation de l' approximation de la loi binomiale obtenue pour :

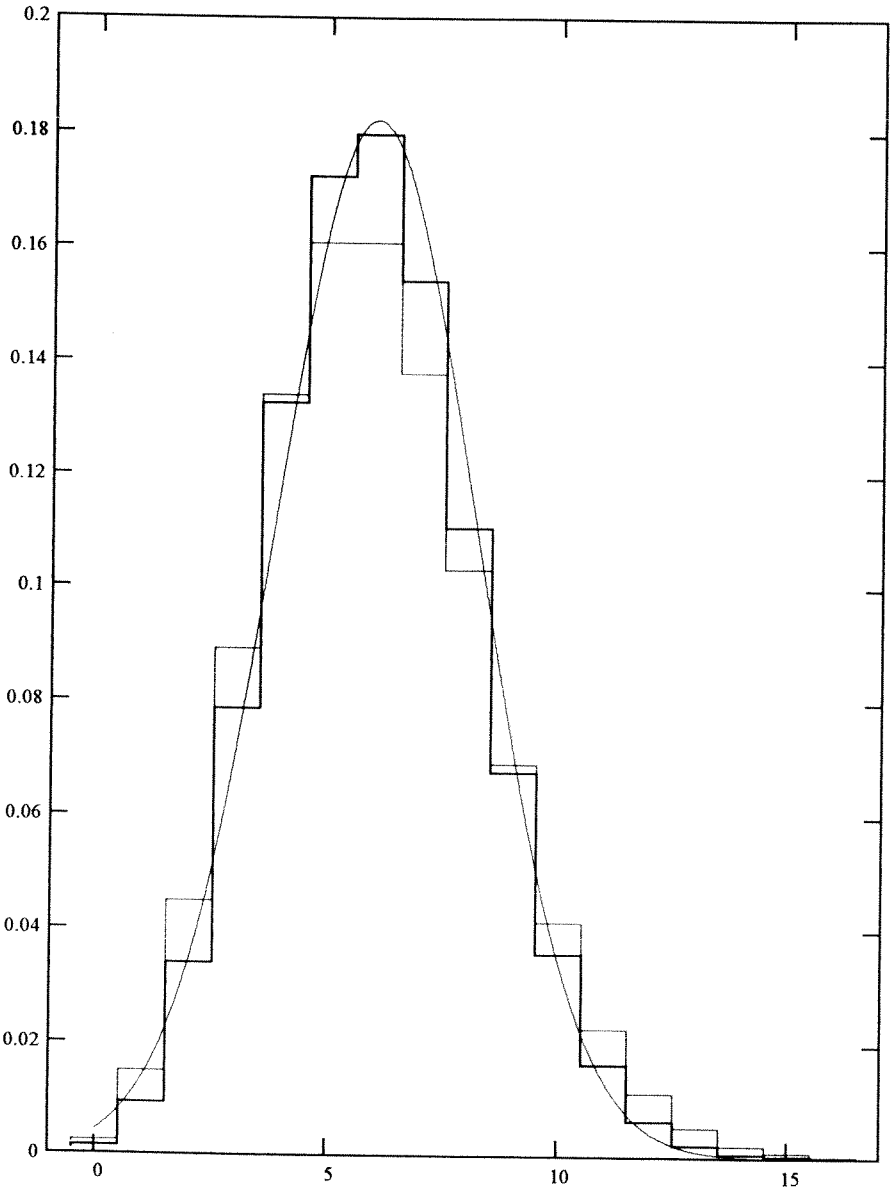
$n = 30$ et $p = 0.2$ (traits pleins)

par une loi gaussienne $N(m, \sigma)$ avec :

$m = 6$ et $\sigma = 2.191$ (courbe)

et une représentation de l' approximation de cette même loi binomiale par une loi de Poisson de paramètre

$\lambda = 6$ (traits fins)



Voici une représentation de l' approximation de la loi binomiale obtenue pour :

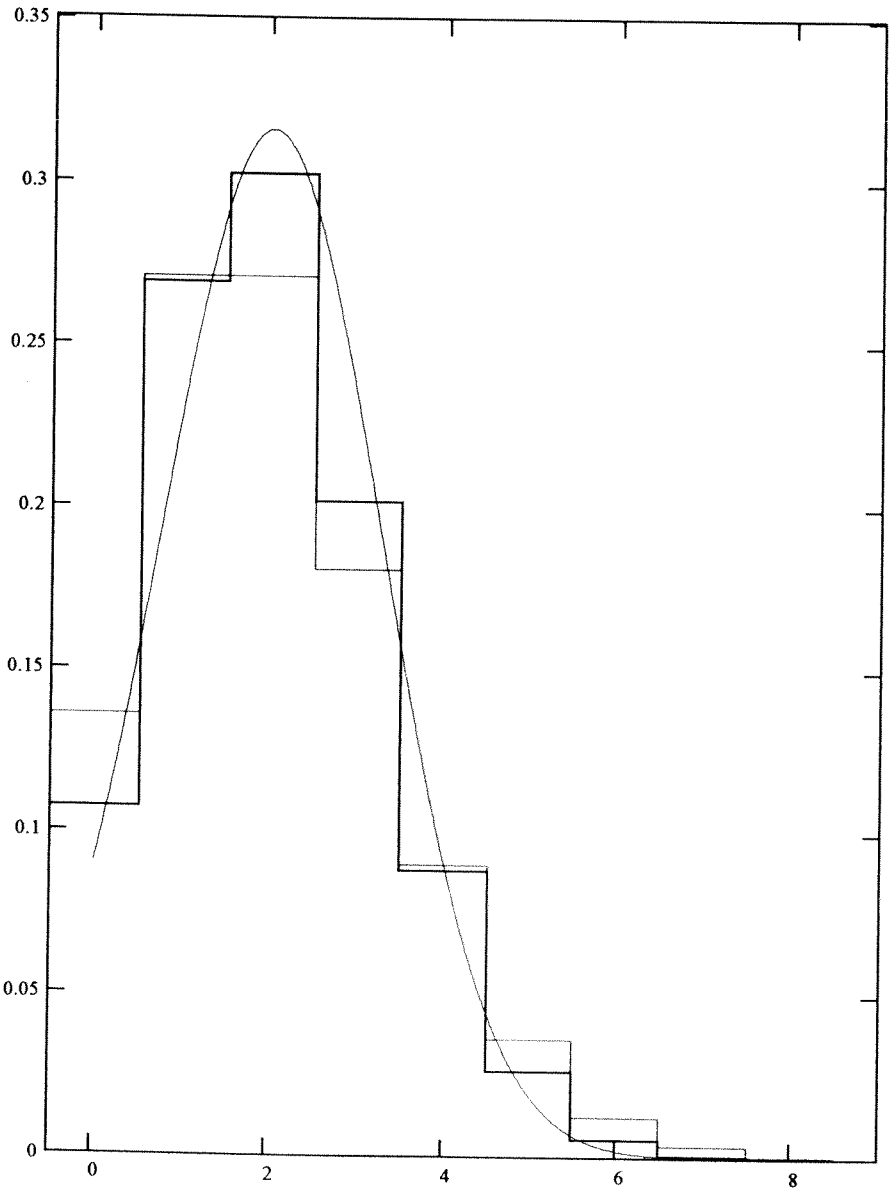
$n = 10$ et $p = 0.2$ (traits pleins)

par une loi gaussienne $N(m, \sigma)$ avec :

$m = 2$ et $\sigma = 1.265$ (courbe)

et une représentation de l' approximation de cette même loi binomiale par une loi de Poisson de paramètre

$\lambda = 2$ (traits fins)



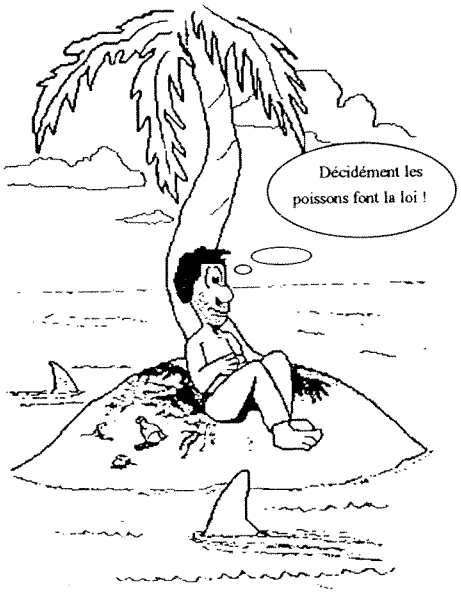
Document pour le professeur

Processus de Poisson

Objectif :

Réfléchir aux hypothèses nécessaires pour modéliser une situation concrète par un processus de Poisson.

Présenter les différentes définitions rencontrées dans les livres concernant le processus de Poisson.



PROCESSUS DE POISSON[1]

Dans le cas de la loi binomiale, on répète n fois une même épreuve sans se soucier de la durée globale de l'expérience. Cette durée a cependant une incidence décisive sur le coût de l'expérimentation.

De plus, dans de nombreux cas, l'expérimentateur qui observe certains faits n'est pas maître de leur nombre d'occurences au cours d'une période donnée. Pour l'étude de certains phénomènes le paramètre décisif sera la durée de l'observation et non le nombre d'épreuves.

On parlera de "signaux" chaque fois que l'on étudie des phénomènes instantanés d'un certain type (débuts de panne sur une machine, apparition d'une particule d'une certaine catégorie, désintégration d'un atome d'un élément radioactif, passage d'un véhicule en un point donné etc.)

On peut modéliser cette situation par une famille de variables aléatoires discrètes X_t où t décrit $[0, +\infty[$. Pour tout entier naturel n, $X_t = n$ traduit le fait que n signaux se produisent dans l'intervalle [0,t[. Une telle famille de variables aléatoires est appelée "processus temporel" Etant donné un intervalle $[t_1, t_2[$, le nombre de signaux observés dans l'intervalle $[t_1, t_2[$ est la variable aléatoire $Y_{[t_1, t_2[} = X_{t_2} - X_{t_1}$. Les variables aléatoires $Y_{[t_1, t_2[}$ sont appelées "accroissements" du processus $(X_t)_{t \in \mathbb{R}}$.

Première définition du processus de Poisson.
(La définition elle même est accompagnée de commentaires imprimés en italique.)

Un processus temporel est un processus de Poisson si les signaux S qui surviennent au cours du temps vérifient les conditions suivantes :

(P_1) La probabilité pour qu'un signal S se produise au cours d'une période de durée h est proportionnelle à h.

proba(S se produise entre t et t+h) = λh (où λ est une constante)

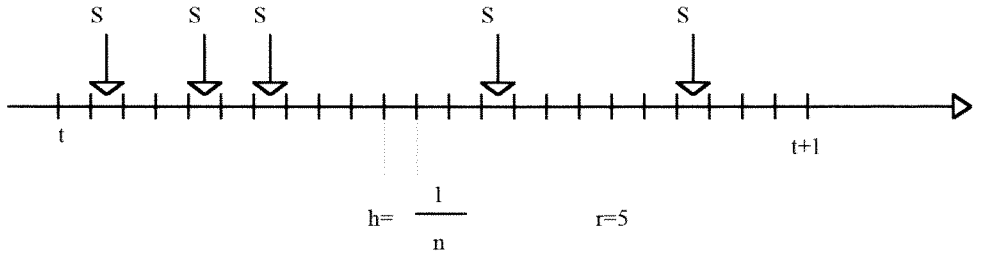
Cette probabilité ne dépend donc que de h (longueur de l'intervalle) et non de t (position sur l'axe des temps). On dit encore que le phénomène est homogène dans le temps, ou conserve la même intensité dans le temps.

(P_2) Etant donné deux intervalles de temps disjoints $[t_1, t_1 + h[$ et $[t_2, t_2 + h[$ les événements "S se produit dans l'intervalle $[t_1, t_1 + h[$" et "S se produit dans l'intervalle $[t_2, t_2 + h[$" sont indépendants.

(P_3) Il existe un réel h tel que la probabilité que le signal S survienne plus d'une fois au cours d'une période de durée h soit égale à 0.
Les signaux S ne sont jamais simultanés mais isolés; P_3 est vérifiée si on choisit h plus petit que l'écart minimal entre deux signaux consécutifs.

[1] POISSON (Siméon-Denis) mathématicien français né à Pithivier en 1781 et mort à Paris en 1840, élève (1798) puis professeur (1802) à l'Ecole Polytechnique. Il introduit la distribution qui porte son nom en 1830. En 1837 dans "Recherches sur la probabilité des jugements en matière criminelle et en matière civile" il la présente comme limite de la loi de Pascal. L'approximation de la loi binomiale est cependant déjà présente dans "La doctrine des chances" ouvrage publié dés 1817 par Abraham de Moivre.

La variable aléatoire Z prenant pour valeurs le nombre de signaux S observés au cours de l'unité de temps est une variable aléatoire de Poisson. Recherchons la loi de probabilité de la variable aléatoire Z.
Décomposons l'unité de temps en un nombre n d'intervalles dits élémentaires de même durée h.
On aura donc $h = \frac{1}{n}$. On choisira n assez grand pour que la propriété P_3 soit vérifiée.
D'après la propriété (P_3) le signal E se produit au plus une fois dans un intervalle de durée h. Il se produira r fois dans l'intervalle [t, t+1] s'il se produit au cours de r intervalles élémentaires. La probabilité pour qu'il se produise au cours de chacun de ces intervalles est constante et égale à λh donc à $\frac{\lambda}{n}$.



Le nombre Z de signaux observés durant l'unité de temps est une variable aléatoire de loi binomiale dont les paramètres sont n et $\frac{\lambda}{n}$ et on a :

$$p(Z = r) = C_n^r \left(\frac{\lambda}{n}\right)^r \left(1 - \frac{\lambda}{n}\right)^{n-r}$$

soit

$$p(Z = r) = \frac{\lambda^r}{r!} \frac{n(n-1)(n-2)\ldots(n-r+1)}{n^r} \left(1 - \frac{\lambda}{n}\right)^{n-r}$$

$$= \frac{\lambda^r}{r!} \frac{n}{n} \frac{n-1}{n} \frac{n-2}{n} \ldots \frac{n-r+1}{n} \left(1 - \frac{\lambda}{n}\right)^n \frac{1}{\left(1 - \frac{\lambda}{n}\right)^r}$$

$$= \frac{\lambda^r}{r!} A(n)B(n)C(n) \quad \text{en posant:}$$

$$A(n) = \frac{n}{n} \frac{n-1}{n} \frac{n-2}{n} \ldots \frac{n-r+1}{n}$$

$$B(n) = \left(1 - \frac{\lambda}{n}\right)^n$$

$$C(n) = \frac{1}{\left(1 - \frac{\lambda}{n}\right)^r}$$

Recherchons la limite quand $n \to +\infty$:
A(n) est le produit r termes qui ont tous pour limite 1. A(n) a pour limite 1.
C(n) a pour limite 1.
B(n) est une forme indéterminée du type 1^∞.

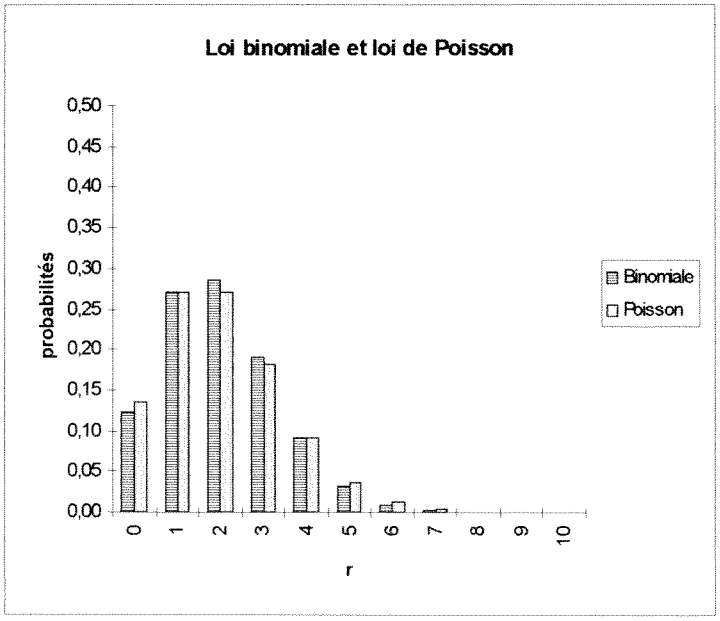
$$\ln(B(n)) = n.\ln\left(1 - \frac{\lambda}{n}\right) = n\left(-\frac{\lambda}{n} + \frac{\lambda}{n}\varepsilon\left(\frac{\lambda}{n}\right)\right) = -\lambda + \lambda.\varepsilon\left(\frac{\lambda}{n}\right)$$

ln(B(n)) a pour limite -λ, donc B(n) a pour limite $e^{-\lambda}$

$$\text{d'où :} \qquad p(Z=r) = e^{-\lambda}\frac{\lambda^r}{r!}$$

Les propriétés P_1, P_2 et P_3 sont celles qui conduisent le plus aisément à l'établissement de la formule donnant p(Z=r).
Notons qu'au passage on a démontré le résultat suivant :

Une loi de Poisson de paramètre λ peut-être utilisée comme approximation d'une loi binomiale de paramètres n et p lorsque n est suffisamment grand et p suffisamment petit avec np=λ.



Loi binomiale et loi de Poisson		
n=	20	
p=	0,1	
lamda=	2	
r	Binomiale	Poisson
0	0,121577	0,135335
1	0,270170	0,270671
2	0,285180	0,270671
3	0,190120	0,180447
4	0,089779	0,090224
5	0,031921	0,036089
6	0,008867	0,012030
7	0,001970	0,003437
8	0,000356	0,000859
9	0,000053	0,000191
10	0,000006	0,000038

Un autre type de situation correspond à ce même modèle. Soit par exemple l'expérience suivante : on saupoudre "uniformément" un nombre N donné de "particules" sur une surface plane limitée par exemple par un carré et d'aire totale S. L'expérience étant pratiquement instantannée, on ne va pas subdiviser le temps en intervalles élémentaires, mais plutôt la surface. On délimite dans la surface de départ n zones de même aire (n grand). Le "signal" observé est ici la présence d'une particule sur une certaine zone. On suppose que les propriétés suivantes sont vérifiées :

(P'_1) La probabilité pour qu'une particule tombe sur une zone d'aire petite ΔS=h est proportionnelle à h.

$$\text{proba("une particule tombe sur une zone d'aire h")} = \lambda h = \frac{N}{S}h$$

C'est la traduction de l'expression "saupoudrage uniforme".

(P'_2) Etant données deux zones disjointes S_1 et S_2, les événements "une particule tombe sur S_1" et "une particule tombe sur S_2" sont indépendants.

(P'_3) Il existe un réel h tel que la probabilité que plus d'une particule tombe sur une zone d'aire h est égale à 0.
Le nombre de zones est grand par rapport au nombre de particules.

La démonstration précédente montre que la variable aléatoire Z prenant pour valeur le nombre de particules sur une zone d'aire unité est une variable aléatoire de Poisson.

Ce point de vue permet d'utiliser la loi de Poisson dans des situations comme : nombre de bombes tombées sur une case d'un quadrillage superposé à la région bombardée, nombre de bactéries par case d'un hématimètre, nombre d'erreurs dans une page de livre, nombre d'anniversaires un jour donné pour un groupe de personnes, nombre de pannes par jour pour les machines d'un parc, nombre de défauts dans des bouteilles en verre (voir à ce propos les exemples en fin de chapitre).

Deuxième définition du processus de Poisson

En fait, l'hypothèse de proportionnalité P_1 ,P'_1 peut être remplacée par une hypothèse moins forte et l'on retrouve néanmoins la formule

$$p(Z = r) = e^{-\lambda} \frac{\lambda^r}{r!} \quad (1)$$

En utilisant les notations définies ci-dessus supposons que :

Des signaux E se produisent au cours du temps en vérifiant les conditions suivantes.
(p_1) La probabilité pour qu'un signal E se produise pendant un intervalle [t, t+h[ne dépend que de h.
(le processus est homogène dans le temps)
(p_2) Etant donné deux intervalles de temps disjoints $[t_1, t_1 + h[$ et $[t_2, t_2 + h[$ les variables aléatoires $Y_{[t_1, t_1 + h[}$ et $Y_{[t_2, t_2 + h[}$ sont indépendantes.
On dit que le processus est "à accroissements indépendants".
(p_3) La probabilité d'obtenir plus d'un signal E dans un intervalle de durée h devient négligeable devant celle d'en obtenir un seul lorsque h tend vers 0.

Ces hypothèses permettent également de démontrer la formule (1). L'idée de la démonstration est la suivante :

En notant $p_n(h)$ la probabilité que E se produise n fois durant un intervalle de longueur h, c'est à dire $p_n(h) = p\left(Y_{[t, t+h[} = n\right)$, on montre successivement que :
$p_0(t + h) = p_0(t).\, p_0(h)$, équation fonctionnelle dont la solution est $p_0(t) = e^{-\lambda t}$ puis que
$p'_n(t) = \lambda\left(p_{n-1}(t) - p_n(t)\right)$, d'où l'on peut déduire $p_n(t) = \frac{(\lambda t)^n}{n!} e^{-\lambda t}$.
Le lecteur intéressé trouvera la démonstration détaillée dans Engel[2].

[2] L'enseignement des probabilités, Arthur Engel, CEDIC

On trouvera dans cette même source le calcul de l'espérance mathématique d'une variable aléatoire qui suit une loi de Poisson :
$E\left(Y_{[t,t+h[}\right) = \lambda t$ qui pour t=1 donne la signification de λ, valeur moyenne du nombre de signaux durant une unité de temps.

EXEMPLES DE SITUATIONS MODELISABLES PAR UNE LOI DE POISSON

Désintégration radioactive (radioactivité naturelle).

Les noyaux d'un élément radioactif se désintègrent à des instants aléatoires.
Pour un noyau non désintégré à l'instant t, la probabilité qu'il se désintègre entre t et t+h ne dépend que de h (les physiciens disent "les noyaux ne vieillissent pas").
Les désintégrations sont indépendantes les unes des autres.
Ces hypothèses ne sont vérifiées que si la durée de l'observation est petite devant la période de la substance radioactive[3].
Le nombre de désintégrations par unité de temps est dans ce cas une variable aléatoire qui suit une loi de Poisson.
L'activité d'un échantillon radioactif s'exprime en curies, ou en bequerel. Une activité de 1 curie correspond à 27.10^9 désintégrations par seconde, un bequerel à une désintégration par seconde.

Exemple :

On a compté le nombre de désintégrations par minute dans un échantillon de radium (^{226}Ra, période 1600 ans environ).

Observations :

Nombre de désintégrations	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Nombre de minutes	1	9	23	48	80	112	133	139	129	107	81	56	36	22	12	6	3	2	1

En choisissant comme modèle une loi de Poisson de paramètre la moyenne du nombre de désintégration par minute calculée à partir des résultats expérimentaux ci dessus (8,347), on obtient une répartition théorique très proche de celle observée.

Modèle théorique :

Nombre de désintégrations	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Nombre de minutes	1	9	23	48	80	111	133	139	129	107	81	56	36	22	12	6	3	1	1

Présence de défauts dans un matériau.

Lors de la fabrication de bouteilles en verre, par exemple, se présente le problème suivant. Dans le verre en fusion, qui servira à faire les bouteilles, il subsiste des corpuscules solides que nous appellerons uniformément "pierres". Si une telle pierre est incorporé au verre d'une bouteille, celle-ci est inutilisable. Les pierres sont réparties aléatoirement dans le verre fondu, mais, dans des conditions de fabrication constantes, une même masse de verre liquide contient en moyenne le même nombre de pierres.

[3] La période d'un élément radioactif est le temps au bout duquel la moitié des noyaux sera désintégrée.

On suppose que chaque pierre a la même probabilité de se trouver incluse dans le matériau de chacune des bouteilles et que les répartitions individuelles des pierres sont indépendantes les une des autres.
Si, par exemple, le verre fondu destiné à la fabrication de 1000 bouteilles contient 200 pierres, le nombre de pierres dans une bouteille est une variable aléatoire qui suit une loi de Poisson de paramètre 200/1000=0,2. Avec ce modèle, la répartition théorique des bouteilles en fonction du nombre de pierres qu'elles contiennent est :

Nombre de pierres	0	1	2	3	4
Nombre de bouteilles	819	164	16	1	0

On peut s'attendre, lors d'une durée de fabrication assez longue, à une proportion de rebuts de 18%.

Estimation d'une concentration de bactéries ou numération globulaire par hématimètre.

Dans une solution fortement diluée on prélève une goutte que l'on étale sur l'hématimètre (plaque observable au microscope) qui est subdivisé en en un grand nombre N de cases. Si la solution est parfaitement homogène, le nombre X de bactéries (ou de globules) dans une case donnée est une variable aléatoire de Poisson de paramètre λ. Si on compte le nombre de bactéries dans chacune des N cases de l'hématimètre, on obtiendra N réalisations indépendantes de la variable aléatoire X et le paramètre λ peut alors être estimé par le nombre moyen de bactéries par case.

Bombardement sur Londres.

Un exemple tout aussi sanglant, rapporté par Jean-Louis Boursin[4], relève du même modèle.
Un quartier de Londres qui a subi un bombardement de V1 durant la seconde guerre mondiale a été subdivisé en 576 carreaux de 500m de côté et on a réparti ces carreaux en fonction du nombre d'impacts reçus :

Nombre d'impacts	0	1	2	3	4	5
Nombre de carreaux	229	211	93	35	7	1

Un total de 535 bombes tombant alèatoirement et indépendamment sur 476 cases, le nombre d'impact X pour un carreau donné est une variable aléatoire de Poisson et le paramètre est 535/576. En utilisant ce modèle théorique, on trouve le tableau ci-dessous :

Nombre d'impacts (r)	0	1	2	3	4	5
Nombre de carreaux 576*P(X=r)	227	211	98	30	7	1

[4]Les structures du hasard, Jean-Louis Boursin, Editions du Seuil 1986

Erreurs typographiques dans les pages d'un livre.

Un livre de 400 pages contient 60 coquilles typographiques réparties aléatoirement (la probabilité de présence d'une erreur dans une partie du livre ne dépend que de la quantité de texte c'est à dire du nombre de pages). Le nombre d'erreurs dans une page donnée du livre est une variable aléatoire de Poisson de paramètre 60/400=0,15. En utilisant ce modèle, on obtient la répartition suivante des pages en fonction du nombre d'erreurs qu'elles contiennent :

Nombre d'erreurs	0	1	2	3
Nombre de pages	344	52	4	0

Nombre d'anniversaires un jour donné dans un groupe fixé

Dans cet exemple cité par Engel[5], on a, à partir des dates d'anniversaires d'un groupe d'étudiants de Francfort, établi le tableau suivant :

Nombre d'anniversaires x	0	1	2	3	4	>4
Nombres de jours avec x anniversaires	182	138	39	6	1	0

En répartissant de manière aléatoire 238 anniversaires sur 365 jours, le nombre d'anniversaires un jour donné est une variable aléatoire de Poisson de paramètre 238/365 et on obtient le tableau suivant concernant la répartition (arrondie) des 365 jours :

Nombre d'anniversaires x	0	1	2	3	4	>4
Nombres de jours avec x anniversaires	190	124	40	9	1	0

Les morts par accident de cheval dans l'armée prussienne (Bortkiewicz)

Il s'agit d'un exemple historiquement célèbre. Bortkiewicz a relevé le nombre de tués par accident de cheval dans l'armée prussienne sur une période de 200 ans

Nombre d'accidents x	0	1	2	3	4
Nombres d'années avec x accidents	109	65	22	3	1

Si on suppose que le nombre d'accidents par année est une variable aléatoire de Poisson avec un paramètre de 122/200, on obtient la répartition théorique suivante :

[5] Les certitudes du hasard, Arthur Engel, Aleas Editeur, 1990

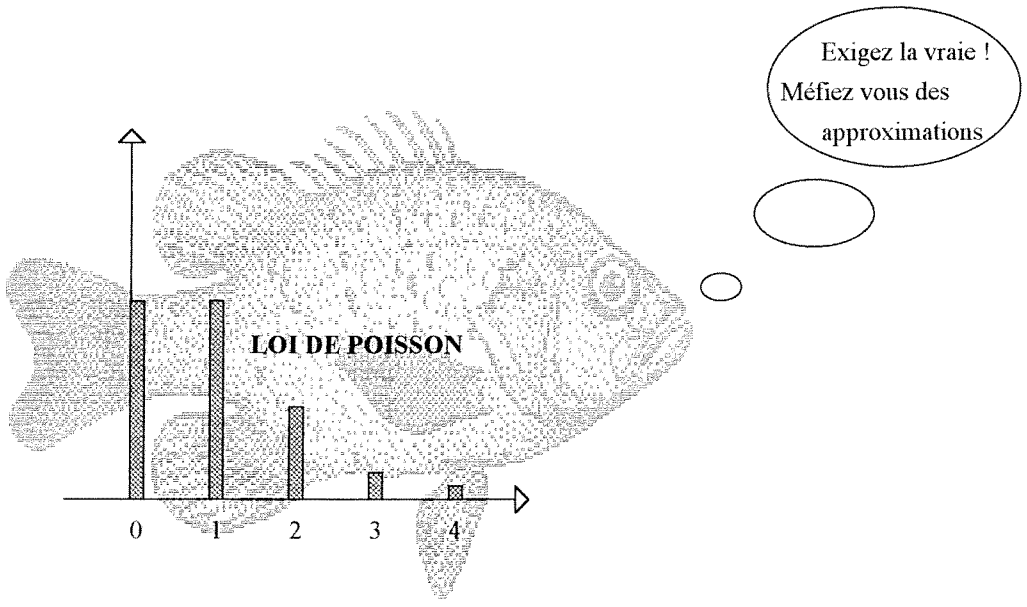
Nombre d'accidents x	0	1	2	3	4
Nombres d'années avec x accidents	109	66	20	4	1

File d'attente.

Imaginons un "service" qui sur une certaine période de 3 heures reçoit et doit "traiter" en moyenne 450 clients. Ces clients se présentent de manière aléatoire et indépendante. On subdivise les trois heures en 180 intervalles d'une minute. Le nombre de clients qui se présentent durant une minute donnée est une variable aléatoire de Poisson de paramètre 450/180 soit 2,5. On obtient la répartition théorique suivante pour 180 minutes en fonction du nombre de clients qui se présentent durant cette minute :

Nombre de clients	0	1	2	3	4	5	6	7	8
Nombre de périodes	15	37	46	38	24	12	5	2	1

En supposant que le service peut absorber sans gène pour l'usager jusqu' à 5 clients par minute, quelle est la probabilité pour un client d'être gêné, c'est à dire quelle est la probabilité que pour une minute donnée il se présente plus de 5 clients ?



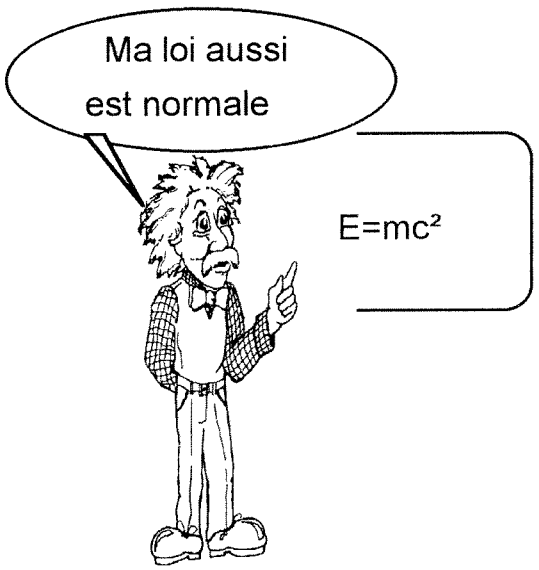
Document pour le professeur

Le caractère universel de la loi normale, la théorie des erreurs , le théorème de la limite centrée

Objectif:

apporter quelques éléments historiques qui ont conduit à la découverte du caractère universel de la loi normale.

présenter brièvement l'aspect expérimental concernant la loi des erreurs (approche de Gauss) et les justifications plus théoriques de Laplace.



Le caractère universel de la loi normale, la théorie des erreurs, le théorème de la limite centrée.

I) Le cheminement de la théorie des erreurs.

La plupart des observations qui dépendent de mesures sont redevables de la théorie des erreurs, notamment en ce qui concerne les résultats de mesures opérées sur une même grandeur.
Le but de cette théorie est de décrire la loi des différences entre la valeur exacte d' une grandeur et les mesures effectuées.

1) les différents modèles proposés.

Nous reprenons ci-dessous partiellement et dans ses grandes lignes une étude faite page 27 et suivantes dans l' ouvrage " Histoire de la statistique ", par J.J. Droesbeke et P. Tassi, paru en 1990 aux Presses Universitaires de France.

Le mathématicien Thomas Simpson (1710-1761) souhaitait défendre l' usage de la moyenne arithmétique des valeurs observées comme " estimateur " d' une grandeur.
Il introduit deux lois pour rendre compte des erreurs :

la loi dite " uniforme et discrète " :

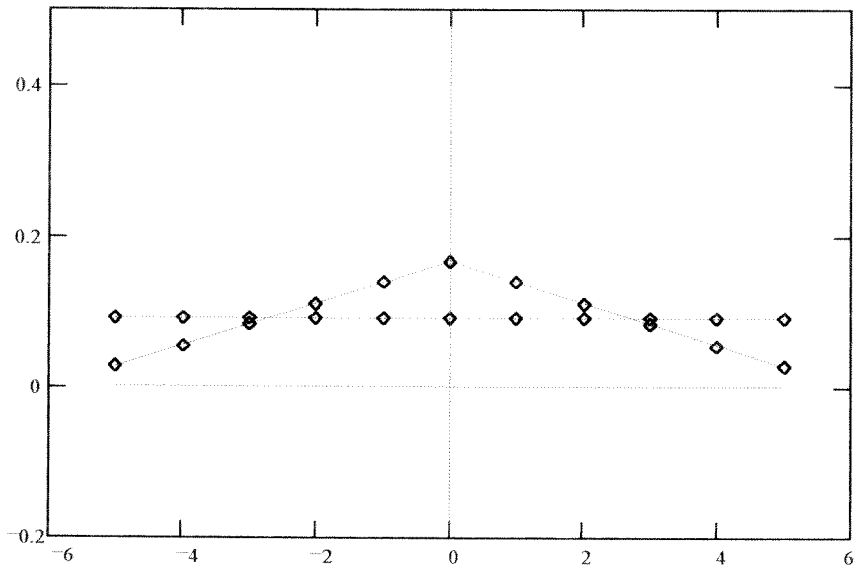
x prend les valeurs $-a, -a+1, \ldots, 0, \ldots, a-1, a$ avec chaque fois la probabilité $p_x = \frac{1}{2 \cdot a + 1}$

la loi " triangulaire discrète " :

x prend les valeurs $-a, -a+1, \ldots, 0, \ldots, a-1, a$ avec chaque fois la probabilité $p_x = \frac{(a+1) - |x|}{(a+1)^2}$

Nous donnons ci-dessous un exemple de représentation de ces lois dans le cas où

$a := 5$



A l' étude de ces deux lois deux observations s' imposent :
ces deux lois sont paires : on commet aussi souvent une erreur positive qu' une erreur négative, la seconde loi est fonction décroissante de |x| : on commet moins souvent de grandes erreurs que de petites.

Ces deux principes seront toujours respectés pour les futures densités envisagées plus loin.

Simpson envisage également pour ces deux lois la loi de probabilité de la somme de n erreurs indépendantes (et donc de leur moyenne) et par passage à la limite obtient la " loi triangulaire continue " :

Pour $-a \le x \le a$: $f(x)=\frac{a-|x|}{a^2}$.

Lagrange fait une étude similaire et étudie d' autres lois :

la " distribution uniforme " : pour $a \le x \le b$: $f(x)=\frac{1}{b-a}$

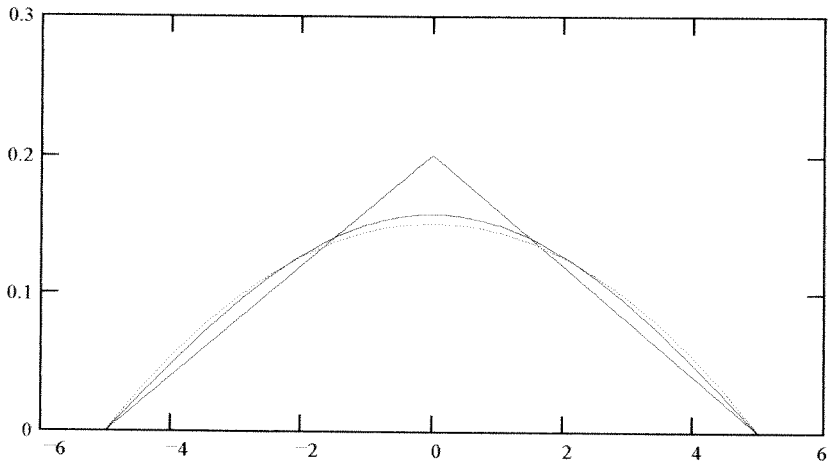
la " distribution parabolique " : pour $-a \le x \le a$: $f(x)=3\cdot\frac{a^2-x^2}{4\cdot a^3}$

la " distribution cosinusoïdale " : pour $-a \le x \le a$: $f(x)=\left(\frac{\pi}{4\cdot a}\right)\cdot\cos\left(\frac{\pi\cdot x}{2\cdot a}\right)$

L' erreur d' arrondi dans des calculs numériques où le nombre de décimales est fixé suit une distribution uniforme.

Les distributions triangulaire continue, parabolique (pointillés) et cosinusoïdale ont été représentées ci-dessous pour

$a := 5$



J.H. Lambert (1728-1777) introduit la loi semi-circulaire : pour $-a \le x \le a$: $f(x)=\frac{2}{\pi\cdot a^2}\cdot\sqrt{a^2-x^2}$

Pierre Simon de Laplace (1749-1827) s' intéressant à l' écart moyen pour mesurer les écarts des mesures à leur moyenne propose la loi dite aujourd'hui " exponentielle bilatérale " définie sur **R** par :

$$f(x)=\frac{k}{2}\cdot e^{-k\cdot|x|} \qquad \text{avec} \quad k>0$$

pour cette loi, une estimation du paramètre k est obtenue par : $\frac{1}{\sigma_1}$
où σ_1 désigne cet écart moyen.

La loi normale : $f(x)=\frac{1}{\sqrt{2\cdot\pi}}\cdot e^{-\frac{x^2}{2}}$ est étudiée par Laplace, mais introduite par Gauss pour étudier une distribution d' erreurs.

2) La démarche de Gauss (1777-1855) :

Dans le polycopié " Probabilités " (D. Foata et A. Fuchs, publication du département de Mathématiques de l' Université Louis Pasteur de Strasbourg, janvier 1994), on trouve en substance :

Gauss considère une grandeur θ qu' il cherche à estimer connaissant n mesures : $x_1, x_2, ...x_n$

Une première approximation est fournie par une mesure d' ajustement plus ancienne : elle consiste à prendre comme estimation de θ la valeur $\bar{\theta}$ réalisant le minimum de la fonction :

$$\theta \to \sum_i (x_i - \theta)^2 .$$

On obtient $\bar{\theta}=\bar{x}=\frac{x_1 + x_2 + ... + x_n}{n}$.

C' est la méthode " des moindres carrés ".

Gauss propose une autre méthode que nous appellerions aujourd' hui " méthode du maximum de vraisemblance ".

L' erreur δ commise lors d' une mesure possède une densité $f(\delta)$.

Sur cette fonction sont faites les hypothèses suivantes :

la fonction f est une densité de probabilité (i.e. $f \geq 0$, $\int_{-\infty}^{+\infty} f(\delta)\, d\delta = 1$),

la fonction f est fonction décroissante de $|\delta|$: on fait moins souvent de grandes erreurs que de petites,

la fonction f est paire : les erreurs positives sont aussi fréquentes que les négatives.

Lors de n mesures : $x_1, x_2, ...x_n$, les erreurs correspondantes sont $\delta_1 = x_1 - \theta$ ..., etc...

Ces erreurs ont supposées indépendantes, de sorte que leur densité soit :

$L(\theta)=f(x_1 - \theta)\cdot f(x_2 - \theta)\cdot ... f(x_n - \theta)$

Gauss recherche les fonctions f pour lesquelles la fonction L prend sa valeur maximale en $\bar{\theta}=\bar{x}$.

Gauss recherche donc les densités telles que l' estimation de θ par la " méthode des moindres carrés " coïncide avec son estimation par la méthode " du maximum de vraisemblance ".

Finalement, Gauss montre que les seules densités vérifiant toutes ces conditions sont les densités normales et centrées.

II) L' approche de Laplace : le théorème de la limite centrée :

1) Bref exposé du théorème de De Moivre - Laplace :

La loi normale porte souvent le nom de Gauss-Laplace, associant dans une même gloire ces deux mathématiciens. Laplace met en évidence le caractère universel de la loi normale d' une façon différente de celle de Gauss : il fait intervenir la loi normale dans le théorème de " De Moivre - Laplace " et éclaire d' un autre jour la théorie des erreurs.

Le théorème de " De Moivre - Laplace " affirme que si :

$$F_n(x)=P\left(\frac{S_n - n\cdot p}{\sqrt{n\cdot p\cdot q}} \leq x\right)$$

est la fonction de répartition du nombre normalisé de succès lors de n tirages de Bernoulli $X_1, X_2, \ldots X_n$ de paramètre p fixé et si :

$$F(x)=\frac{1}{\sqrt{2\cdot\pi}}\cdot\int_{-\infty}^{x} f(x)\,dx$$

est la fonction de répartition de la loi normale, alors :

$$F_n(x) \rightarrow F(x)$$

(lorsque n tend vers $+\infty$, et ce pour tout x).

Les lois des grands nombres renseignent sur le comportement de $\frac{S_n}{n}$ et le théorème que nous venons d' énoncer précise la façon dont est distribué $\frac{S_n}{n}$ autour de sa valeur centrale p, et donc la loi de probabilité limite de $S_n - n\cdot p$.

2) Les diverses étapes ayant amené au théorème :

Dans le livre posthume de Jacques Bernoulli (1654-1705) : " Ars conjectandi " datant de 1713, est clairement démontré que si (X_n) est une suite de tirages de Bernoulli, de paramètre p :

$P\left(\left|\frac{S_n}{n} - p\right| > \varepsilon\right)$ tend vers 0 quand n tend vers $+\infty$.

Ceci constitue une forme élémentaire de la loi des grands nombres.

Bernoulli ne donne en effet aucun renseignement sur la loi de probabilité limite de $S_n - n\cdot p$ ce qui est l' essentiel du théorème de De Moivre et Laplace.

Dans plusieurs traités, Abraham De Moivre (1667-1754) évoque ce problème dans le cas de tirages de Bernoulli de paramètre $p=\frac{1}{2}$.

Le théorème de De Moivre et Laplace énoncé ci-dessus est démontré par Laplace en 1812.

3) Théorèmes de la limite centrée :

On regroupe sous cette terminologie divers théorèmes qui expriment le fait que la loi d' une somme d' un grand nombre de variables aléatoires indépendantes se rapproche de la loi normale.

La forme la plus usitée est la suivante :

" Si $X_1, X_2, \ldots$ sont des variables aléatoires, indépendantes, équidistribuées, d' espérance m et d' écart-type σ , alors $\frac{(X_1 + X_2 + \ldots X_n) - n\cdot m}{\sigma\cdot\sqrt{n}}$ converge en loi vers une variable gaussienne, centrée et réduite."

Pour qu' une loi d' une variable aléatoire tende à être normale, il suffit donc que cette variable soit la somme d' un très grand nombre de causes vérifiant des conditions analogues à celles indiquées ci-dessus.

L' on peut alors concevoir que la théorie des erreurs développée par Gauss fasse également intervenir cette loi : toute erreur observée peut être considérée comme la somme d' erreurs élémentaires .

Diverses démonstrations ou généralisations sont dues à Lyapounov (1901), Lindeberg (1922), Paul Lévy, Kolmogorov (1932), Feller et Khintchine (1937).

La terminologie de théorème de la limite centrée est due à Polya qui l' utilise en 1920 dans l' un de ses articles.

4) D' autres lois limites :

On peut se poser plus généralement le problème suivant :

Si l' on considère des variables aléatoires $X_1, X_2, \ldots$, quelles sont les distributions limites possibles de sommes telles que $\frac{(X_1 + X_2 + \ldots + X_n) - b_n}{a_n}$?

Ce problème a été résolu par Paul Lévy en 1924 et 1937 par l' introduction des lois dites " stables ".

Une loi R non concentrée en un point est dite stable si pour tout entier naturel n il existe deux réels a_n, b_n tels que : $\frac{(X_1 + X_2 + \ldots + X_n) - b_n}{a_n}$ égale en loi une variable aléatoire X de loi R .

($X_1, X_2, \ldots$ désignant des variables aléatoires mutuellement indépendantes, de même loi R)

La loi R sera dite stable au sens strict si pour tout n : $b_n = 0$.

Il est possible de démontrer que pour une loi stable, a_n est nécessairement de la forme $n^{\frac{1}{\alpha}}$

le terme α est dit exposant caractéristique de la loi.

Toute loi stable peut être recentrée en une loi strictement stable.

Si R est stable au sens strict, avec exposant α , quels que soient les réels strictement positifs s, t :

$s^{\frac{1}{\alpha}} \cdot X_1 + t^{\frac{1}{\alpha}} \cdot X_2$ égale en loi $(s+t)^{\frac{1}{\alpha}} \cdot X$

(Cette relation reprend la règle d' addition des variances pour une loi normale).

Les lois stables sont les seules lois pouvant intervenir dans des théorèmes limites, et d' après le théorème de la limite centrée, la loi normale est la seule loi stable possédant une variance.

Comme autres lois stables, on peut citer celles de Cauchy, dont les densités sont de la forme $\frac{1}{\pi} \cdot \frac{c}{c^2 + (x-y)^2}$ et dont l' exposant vaut 1 .

(On pourra trouver un plus ample développement dans le livre de Feller, " An introduction to probability theory and its applications ", chez Wiley).

Démonstration du théorème de De Moivre - Laplace

Rappelons que, si :

$$F_n(x) = P\left[\frac{S_n - np}{\sqrt{npq}} \leq x\right]$$

est la fonction de répartition du nombre normalisé de succès lors de n tirages de Bernoulli $X_1, X_2, \ldots$ de paramètre p fixé et si :

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}u^2} du$$

est la fonction de répartition de la loi normale, alors :

$$F_n(x) \underset{n \to +\infty}{\to} F(x)$$

(pour tout x).

Démonstration :

Soit $S_n = X_1 + X_2 + \ldots X_n$.
Nous savons : $P(S_n = k) = C_n^k p^k q^{n-k}$, avec $q = 1 - p$.

On utilise la formule de Stirling : $n! = \sqrt{2\pi n} \cdot (\frac{n}{e})^n \cdot e^{\theta_n}$, avec : $0 < \theta_n \leq \frac{1}{12n}$.

Nous pouvons alors remplacer les différentes factorielles intervenant dans C_n^p pour obtenir :

$$P(S_n = k) = C_n^k p^k q^{n-k} = \frac{\sqrt{2\pi n} \cdot \left[\frac{n}{e}\right]^n \cdot e^{\theta_n}}{\sqrt{2\pi k} \cdot \left[\frac{k}{e}\right]^k \cdot e^{\theta_k} \cdot \sqrt{2\pi(n-k)} \cdot \left[\frac{(n-k)}{e}\right]^{(n-k)} \cdot e^{\theta_{n-k}}} \cdot p^k \cdot q^{n-k}$$

$$\boxed{P(S_n = k) = \frac{1}{\sqrt{2\pi}} \cdot \sqrt{\frac{n}{k.(n-k)}} \cdot \left[\frac{n.p}{k}\right]^k \cdot \left[\frac{n.q}{n-k}\right]^{n-k} \cdot e^{\theta_{n,k}}}$$

avec : $\theta_{n,k} = \theta_n - \theta_{n-k} - \theta_k$, et : $|\theta_{n,k}| \leq \frac{1}{12} \cdot \left[\frac{1}{n} + \frac{1}{k} + \frac{1}{n-k}\right] = \frac{1}{12} \cdot \left[\frac{1}{n} + \frac{n}{k.(n-k)}\right]$.

Tout entier k peut être considéré comme fonction d' un réel h tel que : $k = np + h\sqrt{npq}$ et l' on a alors : $n - k = nq - h\sqrt{npq}$.

Nous examinons les différents termes composant : $P(S_n = k)$

- Intéressons nous d' abord au terme : $e^{\theta_{n,k}}$:

Pour $h \in [a, b]$ nous avons :

$$\frac{\left[np + a\sqrt{npq}\right]\left[nq - b\sqrt{npq}\right]}{n.npq} \leq \frac{k(n-k)}{n.npq} \leq \frac{\left[np + b\sqrt{npq}\right] \cdot \left[nq - a\sqrt{npq}\right]}{n.npq},$$

Quand n tend vers $+\infty$, l' expression $\frac{k(n-k)}{n.npq}$ converge donc uniformément vers 1, et peut donc être minorée par $\frac{1}{2}$ pour toutes valeurs de n supérieures à un certain seuil n_0.

Nous avons alors :

$$\boxed{\left|\theta_{n,k}\right| \leq \frac{1}{12}\cdot\left[\frac{1}{n}+\frac{n}{k.(n-k)}\right] \leq \frac{1}{12}\cdot\left[\frac{1}{n}+\frac{2}{npq}\right]}.$$

- Etudions : $\left[\frac{n.p}{k}\right]^k$ et $\left[\frac{n.q}{n-k}\right]^{n-k}$:

$$\ln\left\{\left[\frac{np}{k}\right]^k\right\} = k\ln\frac{np}{k} = \left[np+h\sqrt{npq}\right]\ln\left\{1-\frac{h\sqrt{npq}}{np+h\sqrt{npq}}\right\}.$$

On utilise la formule de Taylor avec reste intégral :

$$\ln(1-y) = -y-\frac{y^2}{2}-\int_0^y \frac{t^2}{1-t}dt$$

$$\left|\ln(1-y)+y+\frac{y^2}{2}\right| \leq \frac{|y|^3}{3(1-|y|)}$$

On obtient, en faisant $y=\frac{h\sqrt{npq}}{np+h\sqrt{npq}}$, et en multipliant par $\left[np+h\sqrt{npq}\right]$:

$$\left[np+h\sqrt{npq}\right]\left|\ln\left\{1-\frac{h\sqrt{npq}}{np+h\sqrt{npq}}\right\}+\frac{h\sqrt{npq}}{np+h\sqrt{npq}}+\frac{\left[\frac{h\sqrt{npq}}{np+h\sqrt{npq}}\right]^2}{2}\right| \leq \frac{\alpha_n(h)}{\sqrt{n}};$$

où $\frac{\alpha_n(h)}{\sqrt{n}} = \frac{\left|\frac{h\sqrt{npq}}{np+h\sqrt{npq}}\right|^3}{3\left\{1-\left|\frac{h\sqrt{npq}}{np+h\sqrt{npq}}\right|\right\}}\left[np+h\sqrt{npq}\right]$

Lorsque n tend vers $+\infty$ la suite de fonctions $(\alpha_n(h))$ est uniformément convergente pour h.

Ainsi il est possible d' écrire une égalité de la forme :

$$\ln\left\{\left[\frac{np}{k}\right]^k\right\}+h\sqrt{npq}+\frac{h^2npq}{2\left[np+h\sqrt{npq}\right]} = \frac{\beta_n(h)}{\sqrt{n}},$$

où la suite de fonctions $(\beta_n(h))$ est uniformément convergente pour h.

On a aussi :

$$\frac{h^2npq}{2\left[np+h\sqrt{npq}\right]} = \frac{h^2q}{2}\left[1-h\sqrt{\frac{q}{np}}+\frac{h^2\frac{q}{np}}{1+h\sqrt{\frac{q}{np}}}\right].$$

On arrive ainsi à une relation de la forme :

$$\boxed{\ln\left\{\left[\frac{np}{k}\right]^k\right\}+h\sqrt{npq}+\frac{h^2q}{2} = \frac{\gamma_n(h)}{\sqrt{n}}}$$

la suite de fonctions $(\gamma_n(h))$ étant uniformément convergente pour h,

et l' on obtient de même :

$$\boxed{\ln\left[\left(\frac{nq}{n-k}\right)^{n-k}\right] - h\sqrt{npq} + \frac{h^2 p}{2} = \frac{\delta_n(h)}{\sqrt{n}}}$$

la suite de fonctions $(\delta_n(h))$ étant uniformément convergente pour h.

- Etudions : $\sqrt{\frac{n}{k.(n-k)}}$

$$\frac{n}{k.(n-k)} = \frac{n}{\left[np+h\sqrt{npq}\right]\left[nq-h\sqrt{npq}\right]} = \frac{1}{npq\left[1+h\sqrt{\frac{q}{np}}\right]\left[1-h\sqrt{\frac{p}{nq}}\right]}$$

Il vient donc que :

$$\boxed{\text{la suite } \left[\sqrt{\frac{n}{k.(n-k)}}\cdot\sqrt{npq}\right] \text{ converge donc uniformément vers 1.}}$$

On arrive finalement à une relation de la forme :

$$\boxed{P(S_n = k) = \frac{\varepsilon_n(h)}{\sqrt{2\pi npq}}\exp\left[-\frac{h^2}{2}\right]},$$

où la suite de fonctions $(\varepsilon_n(h))$ converge uniformément vers 1.

Il est alors possible de vérifier que, quels que soient les réels a, b vérifiant $-\infty \leq a < b \leq +\infty$:

$$\boxed{\lim_{n\to+\infty} P\left[a \leq \frac{S_n - np}{\sqrt{npq}} \leq b\right] = \frac{1}{\sqrt{2\pi}}\int_a^b \exp\left[-\frac{x^2}{2}\right]dx}$$

ce qui démontre le théorème.

(Pour de plus amples développements, se reporter au livre de Dacunha-Castelle, Probabilités et Statistiques, tome 1, aux éditions Masson.)

Document pour l'élève

Ce document est donné sous sa forme originalepour l'élève. Il est ensuite détaillé, commenté et généralisé pour le professeur.

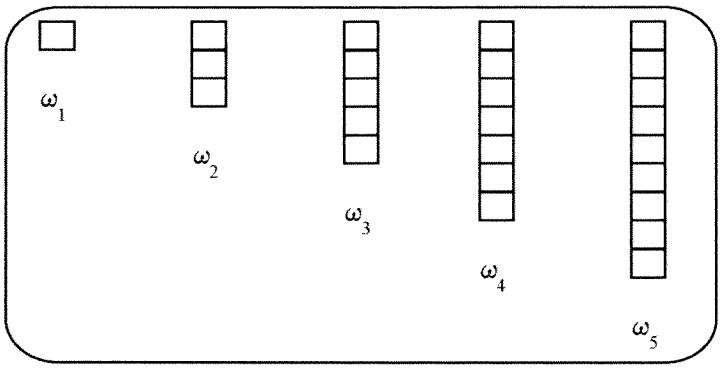
Echantillonage

Objectif:

Introduire la notion d'estimateur

ECHANTILLONNAGE

On dispose d'un ensemble Ω de cinq tiges dont les longueurs respectives sont 1, 3, 5, 7 et 9.



La population Ω

On choisit au hasard l'une de ces tiges. Soit X la variable aléatoire qui prend pour valeur la longueur de la tige choisie.

1) Calculer l'espérance mathématique et l'écart-type de la variable aléatoire X.

On tire successivement avec remise deux tiges dans l'ensemble Ω. On obtient un couple de tiges (ω_i, ω_j) qu'on appelle échantillon de taille 2.

On note X_1 la variable aléatoire qui prend pour valeur la longueur de la première tige tirée et X_2 la variable aléatoire qui prend pour valeur la longueur de la deuxième tige tirée.

2) Le tableau ci-dessous (à compléter) donne les valeurs de X_1 et X_2 pour chacun des 25 échantillons.

barre 1→ barre 2↓	ω_1	ω_2	ω_3	ω_4	ω_5
ω_1	1,1	1,3	1,5		
ω_2	3,1				
ω_3	5,1				
ω_4					
ω_5					

3) On appelle moyenne d'échantillon et on note $\overline{X}$ la variable aléatoire définie par :

$$\overline{X} = \frac{X_1 + X_2}{2}.$$

a) Calculer les valeurs de $\overline{X}$ pour chacun des 25 échantillons.

b) Calculer l'espérance mathématique, la variance et l'écart-type de la variable aléatoire $\overline{X}$.

c) Montrer que $E(\overline{X}) = E(X)$ et que $V(\overline{X}) = \frac{V(X)}{2}$.

4) On appelle variance d'échantillon et on note S^2 la variable aléatoire définie par :

$$S^2 = \frac{1}{2}\left[\left(X_1 - \overline{X}\right)^2 + \left(X_2 - \overline{X}\right)^2\right]$$

a) Calculer les valeurs de S^2 pour chacun des 25 échantillons.

b) Calculer l'espérance mathématique de S^2.

c) Montrer que $E(S^2) = \frac{1}{2} V(X)$

5) On appelle variance d'échantillon corrigée et on note S_c^2 la variable aléatoire définie par :

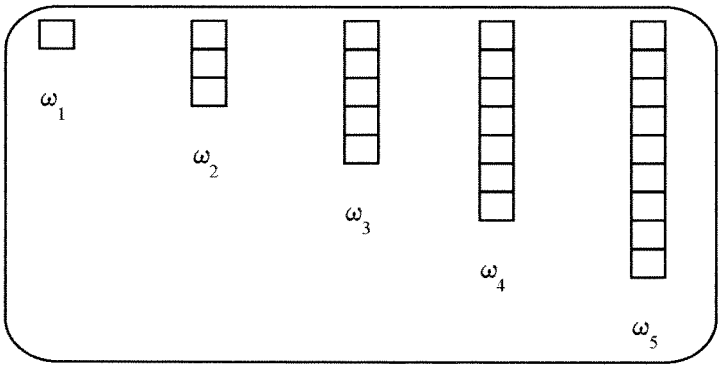
$$S_c^2 = \frac{1}{1}\left[\left(X_1 - \overline{X}\right)^2 + \left(X_2 - \overline{X}\right)^2\right]$$

Montrer que $E(S_c^2) = V(X)$.

VERSION COMMENTEE POUR LE PROFESSEUR

Le document ci-dessous reprend l'ensemble de l'exemple présenté ci-dessus en introduisant le vocabulaire et les notations habituellement utilisées et en énonçant les résultats dans le cas général des échantillons de taille n. Ces commentaires, destinés au professeur, sont en italique.

On dispose d'un ensemble Ω de cinq tiges dont les longueurs respectives sont 1, 3, 5, 7 et 9.



La population Ω

On choisit au hasard l'une de ces tiges. Soit X la variable aléatoire qui prend pour valeur la longueur de la tige choisie.

1) Calculer l'espérance mathématique et l'écart-type de la variable aléatoire X.

Dans le cas général, on a une ***population*** Ω *de taille N :*

$$\Omega = \{\omega_1, \omega_2, \omega_3, \ldots, \omega_N\}$$

les événements élémentaires étant équiprobables et une variable aléatoire X définie sur Ω *:*

$$X : \Omega \rightarrow R$$

$$\omega_i \rightarrow X(\omega_i) = x_i$$

L'espérance mathématique et la variance de X sont bien sûr définis par :

$$\mu = E(X) = \frac{1}{N}\sum_{i=1}^{N} x_i \quad \text{et} \quad \sigma^2 = \frac{1}{N}\sum_{i=1}^{N} (x_i - \mu)^2$$

Ces paramètres sont dans la pratique souvent inconnus et on cherche à les "estimer".

On tire successivement avec remise deux tiges dans l'ensemble Ω. On obtient un couple de tiges (ω_i, ω_j) qu'on appelle échantillon de taille 2.

Un ***échantillon*** *non exhaustif de taille n est un n-uplet de* Ω^n *("non exhaustif" signifie que le tirage se fait avec remise).*

$$e = (\alpha_1, \alpha_2, \ldots, \alpha_n)$$

On note X_1 la variable aléatoire qui prend pour valeur la longueur de la première tige tirée et X_2 la variable aléatoire qui prend pour valeur la longueur de la deuxième tige tirée.

On définit sur Ω^n un n-uplet de variables aléatoires (X_i) par :

$$X_i(\alpha_1, \alpha_2, \ldots, \alpha_n) = X(\alpha_i)$$

Ce n-uplet est appelé échantillon de la variable aléatoire X. Les variables aléatoires X_1, X_2, ...X_n sont indépendantes et de même loi que X.

2) Le tableau ci-dessous (à compléter) donne les valeurs de X_1 et X_2 pour chacun des 25 échantillons.

barre 1→ barre 2↓	ω_1	ω_2	ω_3	ω_4	ω_5
ω_1	1,1	1,3	1,5		
ω_2	3,1				
ω_3	5,1				
ω_4					
ω_5					

3) On appelle moyenne d'échantillon et on note $\overline{X}$ la variable aléatoire définie par :

$$\overline{X} = \frac{X_1 + X_2}{2}.$$

La variable aléatoire $\overline{X}$ est un ***estimateur*** *de l'espérance μ. de X.*

a) Calculer les valeurs de $\overline{X}$ pour chacun des 25 échantillons.
b) Calculer l'espérance mathématique, la variance et l'écart-type de la variable aléatoire $\overline{X}$.
c) Montrer que $E(\overline{X}) = E(X)$ et que $V(\overline{X}) = \frac{V(X)}{2}$.

On montre que dans le cas général (échantillon de taille n d'une population de N éléments, ce qui donne N^n échantillons possibles) que pour la variable aléatoire $\overline{X}$ définie par : $\overline{X} = \frac{X_1 + X_2 + \ldots + X_n}{n}$ *on a encore :* $E(\overline{X}) = E(X) = \mu$ *et l'on traduit ce résultat en disant que $\overline{X}$ est un* ***estimateur sans biais*** *de μ.*

De plus, la variance de $\overline{X}$ est égale à : $V(\overline{X}) = \frac{V(X)}{n} = \frac{\sigma^2}{n}$ *c'est à dire que* $\lim_{n \to \infty} V(\overline{X}) = 0$ *ce que l'on traduit en disant que $\overline{X}$ est un* ***estimateur sans biais et convergent*** *de μ.*

4) On appelle variance d'échantillon et on note S^2 la variable aléatoire définie par :

$$S^2 = \frac{1}{2}\left[\left(X_1 - \overline{X}\right)^2 + \left(X_2 - \overline{X}\right)^2\right]$$

a) Calculer les valeurs de S^2 pour chacun des 25 échantillons.

b) Calculer l'espérance mathématique de S^2.

c) Montrer que $E(S^2) = \frac{1}{2} V(X)$

Dans le cas général, la variance d'échantillon S^2 est définie par :

$$S^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - (\overline{X})^2$$

et l'on démontre que $E(S^2) = \left(\frac{n-1}{n}\right)\sigma^2$ *. La variable aléatoire S^2 est un* ***estimateur biaisé*** *de σ^2 car son espérance mathématique n'est pas égale à σ^2. C'est la raison pour laquelle on introduit une variance d'échantillon corrigée comme détaillé ci-dessous.*

5) On appelle variance d'échantillon corrigée et on note S_c^2 la variable aléatoire définie par :

$$S_c^2 = \frac{1}{1}\left[\left(X_1 - \overline{X}\right)^2 + \left(X_2 - \overline{X}\right)^2\right]$$

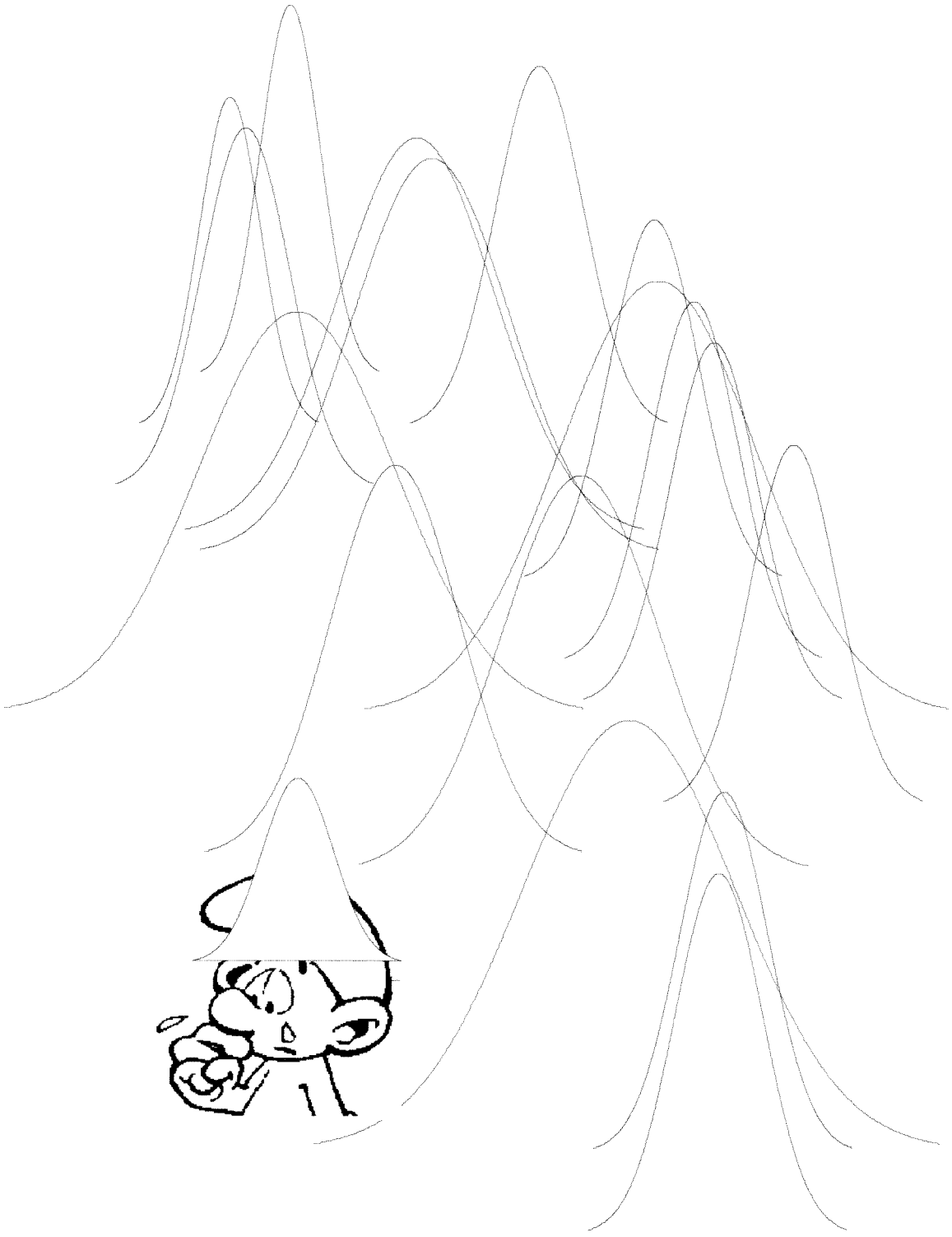
Montrer que $E(S_c^2) = V(X)$.

La variance d'échantillon corrigée S_c^2 est définie par :

$$S_c^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$$

et on démontre que $E(S_c^2) = V(X)$ *et que* $\lim_{n \to \infty} V(S_c^2) = 0$. *La variable aléatoire S_c^2 est donc un estimateur sans biais et convergent de $V(X) = \sigma^2$. Malheureusement, S_c n'est qu'un estimateur biaisé de σ (l'espérance mathématique d'une racine carrée n'étant pas, en général, la racine carrée de l'espérance !).*

ECHANTILLON DE LOIS NORMALES



Document pour le professeur

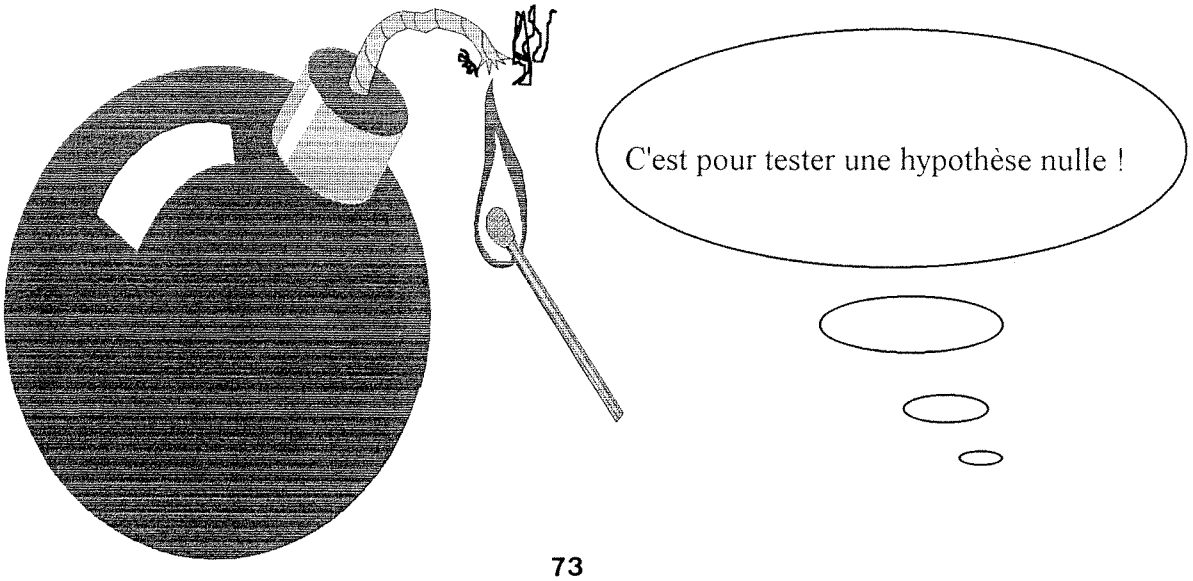
Ce document comporte aussi des exemples utilisables aves les élèves

Tests d'hypothèses

Objectif:

Présenter la problématique des tests d'hypothèses

Introduire une réflexion sur les qualités souhaitables pour un test, notamment la puissance notion qui par ailleurs n'est pas au programme des classes de techniciens supérieurs, mais qui est essentielle pour la compréhension de la problématique.



TESTS D'HYPOTHESE

On utilise un test d'hypothèse dans des situations du type suivant : un "vendeur" propose un produit dont il précise certaines caractéristiques (masse, dimension, durée de vie, proportion d'objets défectueux etc.) ; un "acheteur" désire s'assurer que les déclarations du vendeur sont "crédibles", en examinant un échantillon des objets fournis. Au vu de l'examen de l'échantillon, l'acheteur rejette ou non l'hypothèse qu'il a choisi de tester. Comme il n'examine qu'un échantillon, il ne peut avoir de certitude concernant l'ensemble de la population. En agissant ainsi, il y a deux risques qui correspondent aux 2 cases "erreur" du tableau ci-dessous :

Décision / Marchandise	**conforme**	**non conforme**
rejet	erreur	rejet justifié
acceptation	acceptation justifiée	erreur

1. rejeter à tort une marchandise conforme à ce qui est annoncé (rejet à tort)
2. ne pas rejeter une marchandise non conforme (acceptation à tort)

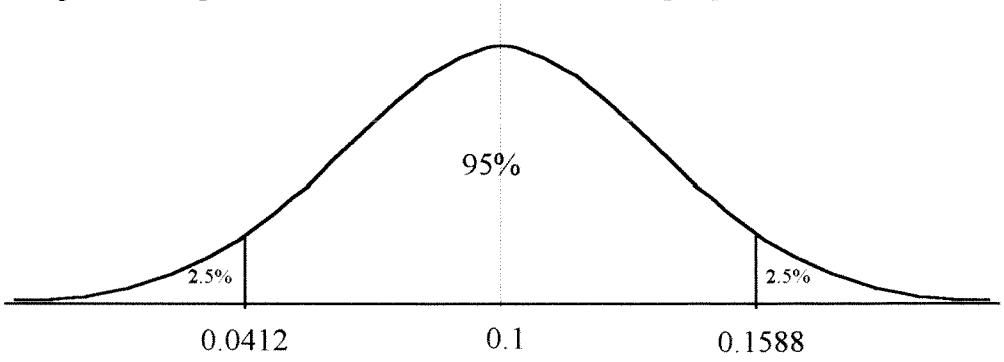
Dans le premier cas, c'est le vendeur (qui fournit une marchandise conforme qu'on lui refuse) qui subit un préjudice. La probabilité de cette éventualité est appelé pour cette raison "risque du vendeur" (ou encore risque de première espèce)
Dans le deuxième cas l'acheteur accepte une marchandise non conforme, le préjudice est pour lui. La probabilité de cet événement est appelé "risque de l'acheteur" ou encore risque de seconde espèce.
A coût égal, par exemple pour une même taille d'échantillon, il n'est pas possible de diminuer les deux risques. Il faut donc trouver un compromis acceptable par le vendeur et l'acheteur.

Pour bâtir un test, on choisit tout d'abord l'hypothèse à tester que l'on appelle (pour une raison mystérieuse) "hypothèse nulle", on choisit un seuil (risque) qui sera le risque de rejet à tort (comme on contrôle totalement cette probabilité, on choisira comme hypothèse nulle celle dont le rejet à tort aurait les conséquences les plus graves), on choisit une taille d'échantillon et on détermine l'intervalle d'acceptation. Si la moyenne observée sur l'échantillon, ou la proportion observée dans l'échantillon sont hors de cet intervalle, on rejette l'hypothèse nulle (et on connaît exactement la probabilité que l'on a de se tromper : c'est le risque de rejet à tort), si, en revanche, le paramètre observé sur l'échantillon est dans l'intervalle d'acceptation on accepte l'hypothèse nulle sans connaître la probabilité d'accepter à tort (qui dépend de la valeur réelle du paramètre testé qui est inconnue). La situation est donc moins "contrôlée" qu'en cas de rejet.

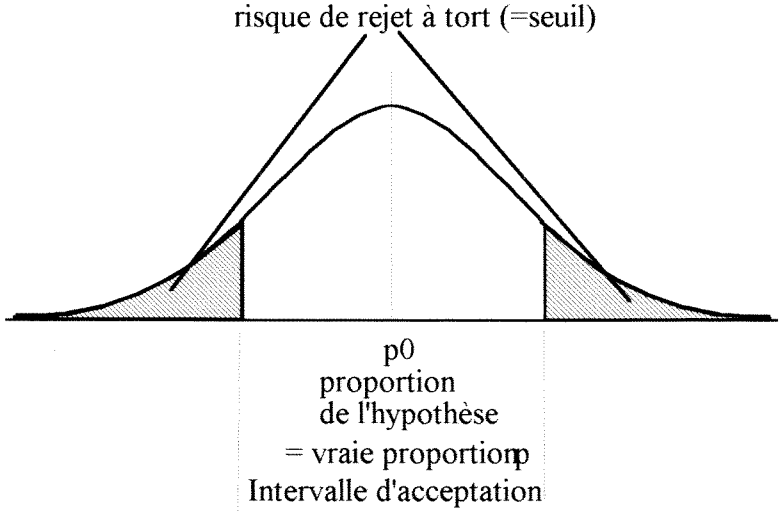
Nous allons présenter dans la suites deux exemples génériques : le test sur une proportion et le test sur une moyenne à partir, dans chacun des deux cas, de grands échantillons.

TEST SUR UNE PROPORTION

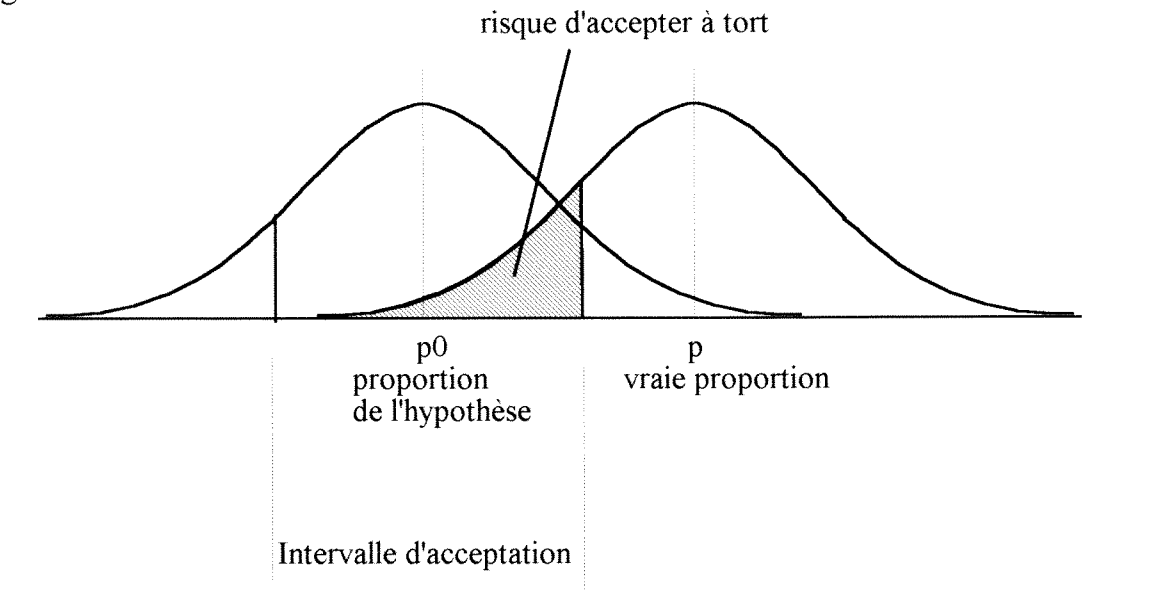
La proportion d'objets défectueux dans un grand lot : le vendeur affirme : "il y a au plus 10% d'objets défectueux dans le lot de 10000 objets que je vous livre".
L'acheteur désire tester si cette affirmation est vraie. Comme il tient à ne pas indisposer inutilement le vendeur il choisit comme hypothèse nulle "la proportion d'objet défectueux est 10%" et comme seuil 5%. Il sait ainsi que la probabilité de rejet à tort est de 5%. Les objets testés ne pouvant être ultérieurement réutilisés (imaginez qu'il s'agisse de fusées de feu d'artifice), il choisit une taille d'échantillon de 100. Pour déterminer l'intervalle d'acceptation, on se place dans le cas où l'hypothèse nulle est vraie (on dit que l'on fait les calculs "sous l'hypothèse H_0"). Dans ce cas, le nombre d'objets défectueux dans un échantillon de taille 100 (choisi au hasard avec ou sans remise puisque l'échantillon est très petit par rapport à la population) suit la loi binomiale B(100, 0.1) que l'on peut approximer par la loi normale N(10, 3). La proportion P d'objets défectueux dans un échantillon de taille 100 suit approximativement la loi normale N(0.1, 0.03). L'intervalle d'acceptation sera tel que la probabilité pour la proportion P d'en sortir soit de 0.05. C'est le cas de l'intervalle A centré en 0.1 et défini par : A=[0.1 - 1.96x0.03 , 0.1 + 1.96x0.03] = [0.0412 , 0.1588]

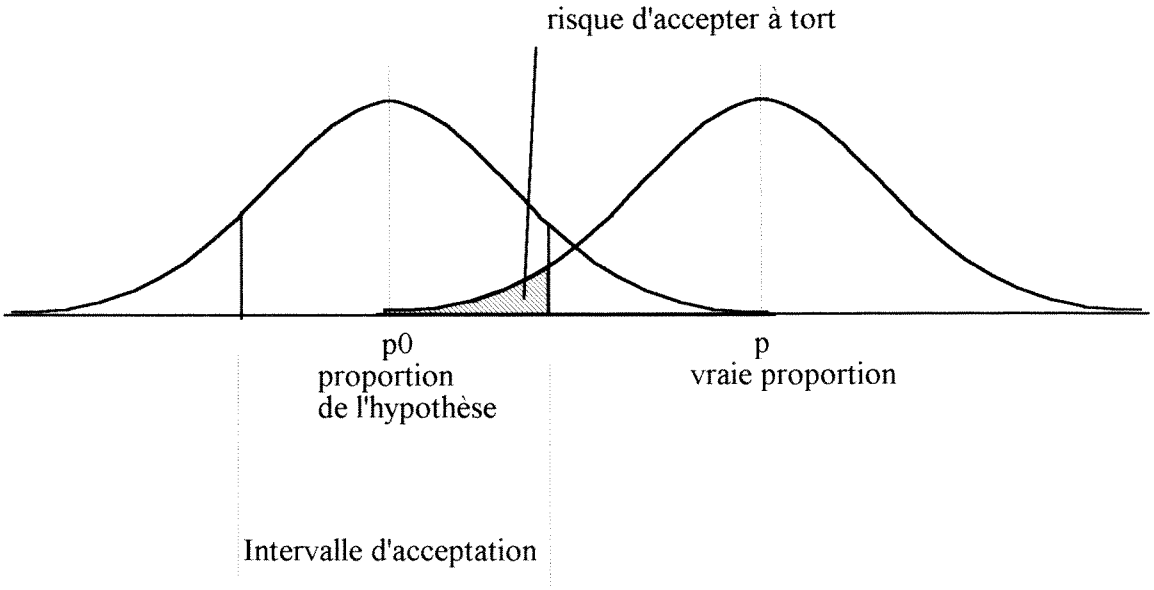


On prélève effectivement l'échantillon, la proportion d'objets défectueux est de 0.18, elle est située hors de l'intervalle d'acceptation A, on rejette l'hypothèse que la proportion d'objet défectueux est de 10% (la probabilité que l'on se trompe est de 5%).
Si la proportion d'objets défectueux observée dans l'échantillon est de 0.15 on sera amené à accepter l'hypothèse nulle p=0.1
Dans ce type de test, si la proportion est conforme à ce qui est annoncé, il est possible que l'on rejette à tort l'hypothèse. La probabilité correspondante est le risque de première espèce ou risque du vendeur et elle est égale au seuil. C'est aussi l'aire de la région hachurée sur le graphique ci-dessous.



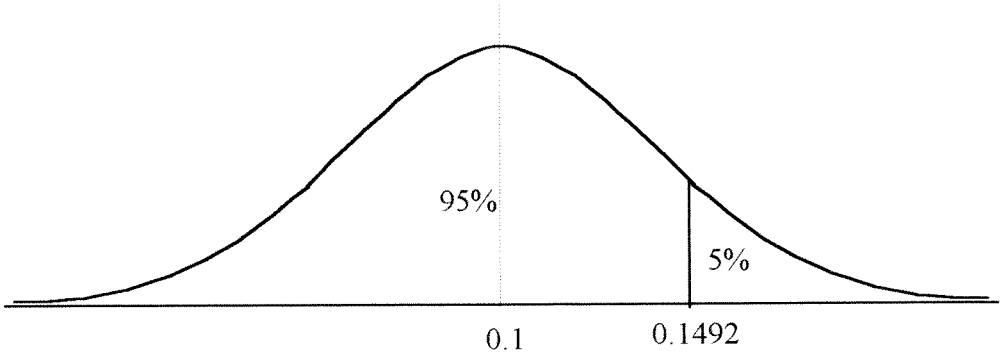
Si en revanche, la proportion réelle p n'est pas conforme à celle annoncée p_0, on peut être amené à accepter à tort l'hypothèse $p=p_0$. La probabilité correspondante est le risque de deuxième espèce ou risque de l'acheteur. Ce risque est égal à l'aire hachurée sur le graphique ci-dessous. On remarque sur ce graphique que le risque d'accepter à tort décroît lorque l'écart entre p et p_0 augmente. Le risque de seconde espèce est d'autant plus faible que l'hypothèse testée est fausse, ou encore : plus l'hypothèse est fausse, plus la probabilité de la rejeter augmente.





Dans le cas de cet exemple, on peut se demander s'il est bien raisonnable pour un acheteur de rejeter un lot d'objet parce que la proportion d'objets défectueux est trop faible (inférieure à 0.0412). L'acheteur testera plutôt l'hypothèse $p \leq 0.1$ en adoptant un intervalle d'acceptation illimité à gauche (commençant à 0 dans la pratique)
On obtient dans ce cas l'intervalle :

$$A' = [\ 0\ ,\ 0.1 + 1.64x0.03] = [\ 0\ ,\ 0.1492]$$



On remarquera que si la proportion observée est de 0.15 on sera conduit cette fois à rejeter l'hypothèse $p \leq 0.1$ alors que l'on acceptait p=0.1 !
Ce test, en acceptant les proportions inférieures à celle annoncées est plus "puissant" pour détecter les proportions supérieures à celles annoncées. Nous reviendrons par la suite sur la notion de puissance d'un test.

TEST SUR UNE MOYENNE

Moyenne des masses des objets d'un grand lot :

Une machine (règlée sur 20g) fabrique des sachets d'une masse moyenne de 20g. On sait que la masse réelle des sachets produits est une variable aléatoire de loi normale dont l'écart type est σ=0.5g (quelle que soit la moyenne). Un acheteur se demande si la machine est bien règlée. Il va tester l'hypothèse H_0 : μ=20. L'acheteur décide de prélever un échantillon de taille 200 (ce qui donne au test un coût qu'il considère comme acceptable) et un seuil de 10% (risque de rejet à tort qui lui semble raisonable)
En supposant que H_0 est vraie, "sous l'hypothèse H_0", la moyenne m des masses des sachets d'un échantillon de taille 200 est une variable aléatoire qui suit la loi normale de moyenne μ et d'écart type $\frac{\sigma}{\sqrt{n}}$. L'intervalle d'acceptation centré en 20 est tel que la probabilité que m en sorte soit de 0,1.

$$A = \left[\mu - 1.64 \frac{\sigma}{\sqrt{n}} \ , \ \mu + 1.64 \frac{\sigma}{\sqrt{n}} \right] = \left[20 - 1.64 \frac{0,5}{\sqrt{200}} \ , \ 20 - 1.64 \frac{0,5}{\sqrt{200}} \right]$$
$$\approx [\ 19.942 \ , \ 20.058 \ ]$$

On réalise le prélèvement et on constate que la moyenne des masses des sachets de l'échantillon est de 19.935g. Cette valeur est extérieure à l'intervalle A. On rejette l'hypothèse H_0 : μ=20.

NOTION DE PUISSANCE D'UN TEST

On dispose d'un dé cubique dont les faces sont numérotées de 1 à 6. On désire "tester" si ce dé est truqué ou non (truqué pour donner plus ou moins de "6" qu'un dé normal)
Pour un dé normal, la probabilité d'obtenir la face "6" est $p=1/6$ et celle de l'événement contraire que nous noterons <6 dans la suite est de 5/6.
On veut tester l'hypothèse $H_0 : p=1/6$

Test 1 :

On lance deux fois le dé et on adopte la règle de décision suivante : si on obtient deux fois le "6" on rejette l'hypothèse que $p=1/6$.

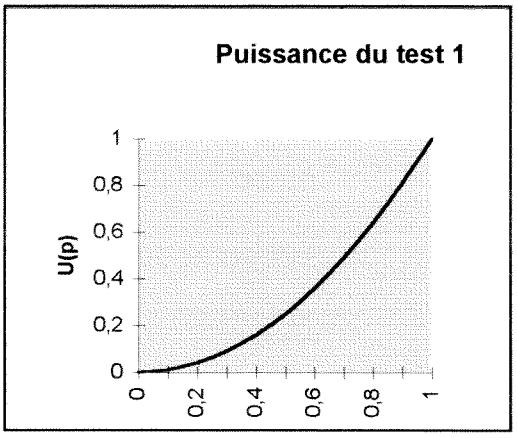
Les quatre résultats possibles lorsqu'on lance deux fois un dé sont :

Premier jet	Deuxième jet	Probabilités pour un dé normal	Probabilités pour un dé quelconque
6	6	1/36	p^2
6	<6	5/36	$p(1-p)$
<6	6	5/36	$p(1-p)$
<6	<6	25/36	$(1-p)^2$

En choisissant cette procédure, on prend, comme dans tout test d'hypothèse, deux risques :

- rejeter (à tort) un dé pour lequel $p=1/6$. La probabilité de rejeter à tort une hypothèse correcte est ici de 1/36 (c'est la probabilité d'obtenir deux "6" avec un dé normal). Cette probabilité est le risque de première espèce.
- accepter (à tort) un dé truqué pour lequel $p \neq 1/6$. La probabilité d'accepter à tort un dé truqué dépend bien sûr de la vrai valeur de p pour le dé testé. Elle est appelée risque de seconde espèce. Ici, par exemple, si pour le dé testé $p=0,3$ alors le risque d'accepter l'hypothèse $p=1/6$ est de $1-p^2=0.91$. Plus généralement, le risque de seconde espèce R est une fonction de la vraie valeur de p et dans le cas du test 1 on a $R(p)=1-p^2$.

La probabilité U de rejeter (avec raison) un dé truqué est le complément à 1 du risque de seconde espèce. Elle est appelée puissance du test car elle traduit la capacité du test à déceler un dé truqué. Ici $U(p)=p^2$.



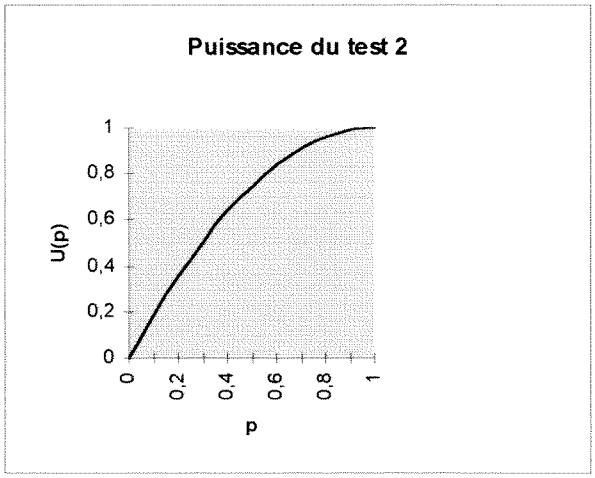
Test 2 : une règle de décision plus sévère.

On lance toujours deux fois le dé, mais on rejette l'hypothèse p=1/6 si l'on a obtenu au moins une fois le "6".

La probabilité de rejet à tort passe à :
1/36+5/36+5/36 = 11/36.
La puissance du test est :

$$U(p) = p^2+2p(1-p)$$

On gagne en puissance par rapport au test 2, mais la probabilité de rejet à tort augmente.



On remarque qu'en conservant le même nombre de jets, il est impossible, en modifiant la règle de décision d'augmenter la puissance du test sans augmenter aussi la probabilité de rejeter à tort une hypothèse vraie. Un "bon" test devrait être puissant et avoir un risque de rejet à tort faible. On peut améliorer ces deux critères en augmentant le nombre de jets.

Lorsqu'on lance trois fois un dé, le nombre de fois où l'on obtient "6" est une variable aléatoire qui suit une loi binomiale donnée dans le tableau ci-dessous :

Nombre de "6"	Dé normal (p=1/6)	Dé quelconque (p)
3	1/216	p^3
2	15/216	$3p^2(1-p)$
1	75/216	$3p(1-p)^2$
0	125/216	$(1-p)^3$

Test 3 :

On lance 3 fois le dé. On rejette l'hypothèse p=1/6 si on obtient au moins une fois le "6".

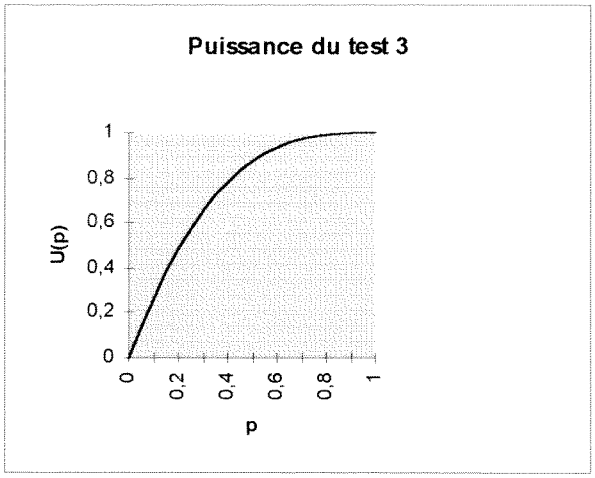
Le risque de rejet à tort est ici :

(1+15+75)/216 soit environ 0,41.

La puissance du test est :

$$U(p) = p^3 + 3p^2(1-p) + 3p(1-p)^2$$

Le test est relativement puissant mais c'est au prix d'un risque de rejet à tort très important. La règle de décision est trop sévère.



Test 4 :

On lance 3 fois le dé. On rejette l'hypothèse p=1/6 si on obtient au moins deux fois le "6".

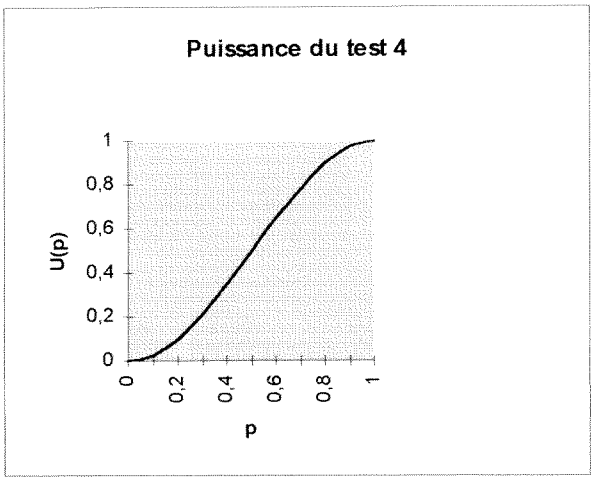
Le risque de rejet à tort est à présent :

(1+15)/216 soit environ 0,08.

La puissance du test est :

$$U(p) = p^3 + 3p^2(1-p)$$

On obtient un risque de rejet à tort supportable mais la puissance en souffre.



Test 5 :

On lance 10 fois le dé. On rejette l'hypothèse p=1/6 si on a obtenu au moins 4 fois le "6".

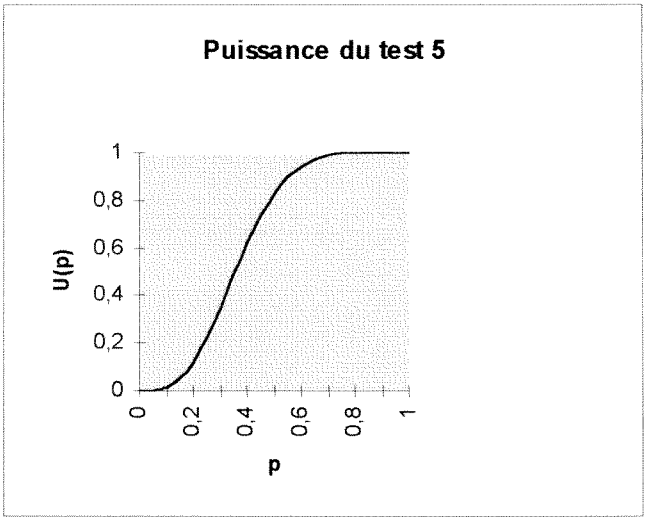
La dixième ligne du triangle de Pascal est : 1 10 45 120 210 252 210 120 45 10 1

La probabilité de rejet à tort avec le test 6 est donc :

$$1-(5/6)^7((5/6)^3+10(1/6)(5/6)^2+45(1/6)^2(5/6)+120(1/6)^3)$$

soit approximativement 0,07.

La puissance du test est : $U(p) = 1-(1-p)^7((1-p)^3+10p(1-p)^2+45p^2(1-p)+120p^3)$



CONCLUSION

Bâtir un "bon" test n'est pas chose aisée. Il faut trouver un compromis acceptable entre le risque de rejet à tort et la puissance du test. Les deux paramètres ne peuvent être améliorés tous les deux que si l'on augmente la taille de l'échantillon testé.
Augmenter la taille de l'échantillon peut rapidement rendre le coût du test prohibitif, la procédure utilisée pouvant être destructive ou très longue (test de fusées de feu d'artifice, test du nombre de ré-enregistrements supportés par des disquettes etc.). Le statisticien se trouve devant de nombreux choix :

1 Choix de l'hypothèse nulle (H_0) : On choisit généralement celle dont le rejet à tort aurait les conséquences les plus graves parce qu'on choisit également le risque de rejet à tort de l'hypothèse nulle qui est le seuil. Si l'on teste, par exemple, l'efficacité d'un médicament sans effets secondaires on choisira sans doute comme hypothèse nulle qu'il est efficace car le rejet à tort de cette hypothèse conduira à ne pas traiter des malades avec un produit susceptible de les soulager.

2 Choix de l'hypothèse alternative (H_1) : L'hypothèse alternative n'est pas toujours le contraire de l'hypothèse nulle. Si p est, par exemple, une proportion de défauts que l'on annonce égale à p_0 on testera l'hypothèse nulle H_0 : $p=p_0$ opposée à l'hypothèse alternative H_1 : $p>p_0$ (seule éventualité gênante).
Choisir l'hypothèse alternative c'est aussi choisir la forme de l'intervalle d'acceptation :
H_0 : $p=p_0$ et H_1 : $p \neq p_0$ donneront des intervalles centrés
H_0 : $p=p_0$ et H_1 : $p>p_0$ donneront des intervalles illimités à gauche
H_0 : $p=p_0$ et H_1 : $p<p_0$ donneront des intervalles illimités à droite etc.
Choisir un intervalle illimité permet d'obtenir un test plus puissant (du "bon" côté) sans nuire aux autres qualités.

3 Choix du seuil : La valeur de 5%, très souvent utilisée dans les exercices d'école peut dans certains cas être inadaptée (imaginer qu'en l'absence de traitement les malades de l'exemple ci-dessus sont condamnés) En augmentant le seuil, c'est à dire la probabilité de rejeter à tort, on obtient un test plus puissant.

4 Choix de la taille de l'échantillon : Il faudra tenir compte du coût qui en résultera mais aussi de la puissance voulue pour le test. Augmenter la taille de l'échantillon rendra le test plus puissant.

5 Choix de l'estimateur : On choisit l'estimateur que l'on utilisera pour le paramètre à tester. En cas de tests sur une proportion ou une moyenne à partir de grands échantillons on utilise les estimateurs usuels (proportion ou moyenne de l'échantillon). Lorsque les échantillons sont plus petits, la situation est moins simple. Les mêmes techniques s'appliquent encore si la variable aléatoire parente est normale. Dans le cas contraire on recherchera des tests moins sensibles au fait que les lois parentes s'écartent de celles espérées. Un tel test est dit "robuste", il donne des indications utilisables même si la loi parente n'est pas très "normale".

On s'aperçoit qu'il y a de nombreux paramètres à fixer et que la recherche du meilleur compromis est délicate.

TEST A PROPOS DES TESTS

En début de première année de BTS, la question suivante a été soumise aux étudiants : on vous propose de jouer à un jeu avec un dé suspect (il est peut-être truqué de manière à obtenir moins de six qu'attendu). Quelle procédure de vérification utiliseriez-vous pour vérifier si le dé est truqué, ou après quelles vérifications accepteriez-vous de jouer ?

Répartition des 56 réponses proposées :

Lancer le dé un certain nombre de fois : (42 réponses, 81%)

6 fois, si 1 fois le six, le dé est bon
6 fois, si la même face sort plus de 3 fois le dé est certainement truqué
6 fois, truqué si le six sort de façon évidente trop souvent
6 fois, il suffit de lancer 6 fois, si on obtient pas de six le dé est faussé

10 lancers p(A)=1/6 1/6x1/6x....=$(1/6)^{10}$; le dé est truqué si p(B)>p(A)
10 fois, si aucun chiffre n'apparaît un nombre exagéré de fois je joue
10 fois, pense non truqué si 1 ou 2 six, truqué si 0 ou 5 six
10 fois, pense truqué si plus de 5 six
10 fois, truqué si aucun 6

12 fois, suppose truqué si plus de 6 six

15 fois, considère non truqué si au moins une fois chaque nombre
15 fois, dé normal si résultats varié, sûrement truqué si 15 six

18 fois, non truqué si trois fois le même chiffre
18 fois, à peu près correct si 3 fois le six

20 fois, si aucun six je penserais que le dé est truqué. Normalement tous les chiffres devraient apparaître au moins une fois.
20 fois, probablement truqué si au moins 10 six
20 fois, si plus de 12 six je penserai que le dé est truqué
20 fois, on ne doit pas avoir de six
20 fois, fortes chances que truqué si au moins 15 six
20 fois, si les résultats sont variés il n'est apparemment pas truqué
au moins 20 fois non truqué si autant de chaque

30 fois certainement truqué si 15 six
30 fois, non truqué si répartition presque identique
30 fois et je regarde si la probabilité d'avoir un six et $\approx$ 1/6
une trentaine de fois, si 9/10 de six je me poserais la question
une trentaine de fois on devrait obtenir environ 5 six

36 fois, semble non pipé si le six sort dans 16.66% (+-3.33%) des cas
36 fois, pas truqué si au moins un six
36 fois, normal si à peu près le même nombre de fois chaque chiffre

40 lancers plus de 25 six cela prouve que le dé est truqué
40 fois si répartition à peu près égale pour chaque valeur il y a de fortes chances qu'il soit non
50 lancers, si on obtient plus de 15 six on suppose le dé truqué
50 fois, on compare les statistiques pour définir la probabilité de tirer un six avec ce dé. Si la probabilité est différente de 1/6 alors le dé est certainement pipé
50 fois, truqué si le six apparaît le plus souvent
50 fois, non truqué si autant de fois chaque chiffre

60 fois accepte si 10 six à 10% près

100 fois, truqué si plus de 50 six
100 fois j'étudie les résultats obtenus, c'est à dire le nombre de fois que le six est sorti, par une loi binomiale et par une loi de Poisson. Si les probabilités sont quasi identiques c'est que le dé est sûrement truqué.
100 fois, non truqué si chaque nombre est apparu à peu près 17 fois
100 lancers et 100/6 fois le six, plus on lance le dé, plus la probabilité d'obtenir un six se rapproche de 1/6

un grand nombre de fois si possible divisible par 6, on peut considérer que le dé est truqué si la probabilité est différente (relativement) de 1/6

n fois et calcul de n pour que $(5/6)^n$ <0.001, probas de 0.999 que le dé soit truqué
pipé

Vérifications mécaniques et géométriques : (6 réponses, 11%)

angles a peu près égaux
immersion dans l'eau, s'il est truqué la face six coule avant les autres
examen des surfaces et des angles
mesure des 6 côtés (surtout pour la face 1 opposée à la six)
vérification de la forme
je regarde toutes les faces

Comparaison avec un dé normal : (3 réponses, 5%)

je lance 30 fois les deux dés et je compare les résultats, s'ils se ressemblent à un 6 obtenu près, je joue.

Procédure à plusieurs niveaux : (4 réponses, 7%)

lancer 12 fois si pas de six, relancer 12 fois
lancer 30 fois, si trop peu de six, relancer 30 fois
lancer 6 fois si le six sort plus de 3 fois, relancer 12 fois, si plus de 6 fois le six le dé doit être truqué
lancer 20 fois, si pas de six relancer 20 fois si toujours pas de six on peut en déduire dé truqué

Baisser les bras : (2 réponses, 3%)

on ne peut pas savoir si le dé est truqué même après 100 lancers par exemple à moins d'obtenir toujours le même chiffre.
faire jouer un ami pour observer.

Document pour le professeur

Petit herbier de lois

Lois utiles pour l'estimation

Petit herbier de lois

Loi binomiale $\mathfrak{B}(n,p)$

Vous connaissez tous la loi binomiale $\mathfrak{B}(n,p)$ qui est la loi de la variable aléatoire Sn : le nombre de boules blanches obtenues lors de n tirages d'une boule au hasard et avec remise dans une urne contenant une proportion p de boules blanches, autrement dit le nombre de "succès" lors de **n répétitions indépendantes d'une même expérience de Bernoulli** $\mathfrak{B}(1,p)$, .

Le paramètre p désigne la probabilité de "succès" à chaque expérience de Bernoulli.

Rappellons que :

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

$$E(S_n) = np = \sum_{k=0}^{n} k \, P(S_n = k)$$

et

$$Var(S_n) = np(1-p) = \sum_{k=0}^{n} (k - E(S_n))^2 \, P(S_n = k))$$

Loi hypergéométrique $\mathfrak{H}(n,N_1,N)$

$\mathfrak{H}(n,N_1,N)$ est la loi du nombre S'_n de boules blanches obtenues lors de n tirages au hasard, successifs sans remise ou un tirage au hasard simultané de n boules dans une urne contenant N_1 boules blanches et N_2 boules non-blanches parmi N boules ($N = N_1 + N_2$).

$$P(S'_n = k) = \frac{C_{N_1}^k \, C_{N_2}^{n-k}}{C_N^n}$$

Posons $p = \frac{N_1}{N}$:

$$E(S'_n) = n\,p \qquad \text{et} \qquad Var(S'_n) = n\,p\,(1-p)\,(1 - \frac{n-1}{N-1})$$

Loi de Poisson $\mathcal{P}(\lambda)$

$\mathcal{P}(\lambda)$ est, fréquemment, la loi du comptage du nombre d'unités arrivant dans un système durant un intervalle de temps. C'est aussi la loi des "événements rares", comme loi limite d'une loi binomiale sous certaines conditions.

Une variable aléatoire X suit une loi de Poisson de paramètre $\lambda > 0$, notée $\mathcal{P}(\lambda)$, si , pour tout entier naturel k :

$$P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$$

$$E(X) = Var(X) = \lambda$$

Approximation binomiale

Si n est petit par rapport à N (par exemple : $n \leq \frac{N}{10}$), la loi hypergéométrique peut être assimilée à une loi binomiale.

$$\mathcal{H}(n,N_1,N) \cong \mathcal{B}(n,p) \quad \text{où } p = \frac{N_1}{N}$$

Approximation de Poisson

Si n est grand et p petit, la loi binomiale $\mathcal{B}(n,p)$ peut être assimilée à une loi de Poisson , ceci pouvant être admis en principe dès que $n \geq 50$ et $np \leq 10$; dans ce cas :

$$\mathcal{B}(n,p) \cong \mathcal{P}(np)$$

Loi normale $\mathcal{N}(\mu, \sigma)$

Disons simplement qu'il s'agit de la loi d'une variable aléatoire continue définie à partir d'une fonction de densité f par :

$$P(a \leq X \leq b) = \int_a^b f(x)\,dx$$

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

$$\sigma^2 = Var\,(X) = \int_{-\infty}^{+\infty} (x-\mu)^2\, f(x)\, dx$$

On utilise aussi très souvent la fonction de répartition F, qui est une primitive de la densité f

$$F(x) = P\,(\,X \leq x\,) = \int_{-\infty}^{x} f(t)\, dt$$

La densité d'une loi normale centrée réduite $\mathcal{N}(0,1)$, dite aussi loi standard, est :

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

La densité d'une loi normale générale $\mathcal{N}(\mu,\sigma)$ d'espérance μ et d'écart-type σ, est :

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

Propriété de standardisation des lois normales

$$\mathcal{L}(X) = \mathcal{N}(\mu,\sigma) \Leftrightarrow \mathcal{L}(X^*) = \mathcal{L}\left(\frac{X-\mu}{\sigma}\right) = \mathcal{N}(0,1)$$

C'est cette propriété qui permet de calculer les probabilités d'une loi normale quelconque à partir des valeurs (tabulées) de la loi $\mathcal{N}(0,1)$

Propriétés de la famille des lois normales

1) Si $\mathcal{L}(X) = \mathcal{N}(\mu,\sigma)$, alors :

$\mathcal{L}(aX + b) = \mathcal{N}(a\mu + b,\ |a|\sigma)$ où a et b sont des constantes réelles.

2) Si $\mathfrak{L}(X_1) = \mathfrak{N}(\mu_1,\sigma_1)$, $\mathfrak{L}(X_2) = \mathfrak{N}(\mu_2,\sigma_2)$ et si de plus X_1 et X_2 sont indépendantes, alors : $\mathfrak{L}(X_1+X_2) = \mathfrak{N}(\mu_1+\mu_2, \sqrt{\sigma_1^2 + \sigma_2^2})$

3) Si $\mathfrak{L}(X_i) = \mathfrak{N}(\mu,\sigma)$ pour i = 1,.....,n et si de plus les X_i sont indépendantes, alors :

$$\mathfrak{L}(\sum_{i=1}^{n} X_i) = \mathfrak{N}(n\mu ,\sigma\sqrt{n})$$

Posons $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$; il en résulte que

$$\mathfrak{L}(\overline{X}) = \mathfrak{N}(\mu , \frac{\sigma}{\sqrt{n}})$$

Théorème central limite

Cette expression désigne en fait, toute une famille de théorèmes affirmant que, sous certaines conditions et notamment l'indépendance, la loi de probabilité d'une somme de n variables aléatoires, individuellement petites par rapport à la somme, converge vers une loi normale lorsque n tend vers l'infini.
Ainsi, si les X_i, pour i = 1,...,n, sont des variables aléatoires indépendantes et de même loi quelconque, d'espérance μ et d'écart-type σ, alors, lorsque n est "suffisamment" grand :

$$\mathfrak{L}(\sum_{i=1}^{n} X_i) \cong \mathfrak{N}(n\mu , \sigma\sqrt{n})$$

et

$$\mathfrak{L}(\overline{X}) \cong \mathfrak{N}(\mu , \frac{\sigma}{\sqrt{n}})$$

Approximation normale de la loi binomiale

Si $\mathfrak{L}(S_n) = \mathfrak{B}(n,p)$, S_n peut être considérée comme la somme de n variables aléatoires X_i indépendantes et de même loi de Bernoulli $\mathfrak{B}(1,p)$.
Donc si n est assez grand :

$$\mathfrak{L}(S_n) = \mathfrak{L}(\sum_{i=1}^{n} X_i) \cong \mathfrak{N}(np, \sqrt{np\,(1-p)})$$

On peut admettre que cette approximation est acceptable dès que $np \geq 10$ et $n(1-p) \geq 10$

Il faudra tenir compte du fait que l'on approche une loi discrète par une loi continue, en posant :

$$P(S_n = k) = P(S_n \in (k - \frac{1}{2} ; k + \frac{1}{2}))$$

Approximation normale de la loi de Poisson

Si $\lambda \geq 10$, on peut admettre $\mathcal{P}(\lambda), \cong \mathcal{N}(\lambda, \sqrt{\lambda})$

Lois du χ^2 (Chi-deux).

Définition.

X_1 , X_2 ...., X_p étant p variables aléatoires de même loi $\mathcal{N}$ (0 ; 1) indépendantes, on appelle loi du chi-deux à p degrés de liberté, χ^2_p , la loi de la variable aléatoire $\sum_{i=1}^{p} X_i^2$

Conséquence.

La somme de deux variables indépendantes χ^2_p et χ^2_q est une variable χ^2_{p+q}.

Paramètres de la loi χ^2_p

Si Y suit une loi χ^2_p , alors $E(Y) = p$ et $V(Y) = 2p$

Loi de Student $\mathcal{T}_p$

Soient X et Y deux variables aléatoires indépendantes

X de loi $\mathcal{N}(0 , 1)$,

Y de loi χ^2_p ,

on définit la variable aléatoire de Student à p degrés de liberté T_p comme suit :

$$T_p = \frac{X}{\sqrt{\frac{Y}{p}}}$$

Propriétés.

Si p=1, la loi de Student est la loi de Cauchy qui ne possède aucun moment fini.

Si $p > 1$, $E(T_p) = 0$

Si $p > 2$, $V(T_p) = \frac{p}{p-2}$.

Si $p \to \infty$, $T_p \xrightarrow{L} \mathcal{N}(0 ; 1)$.

Utilisation des lois du Chi-deux et de Student

Ces lois nous intéressent parce qu'elles sont, sous certaines conditions, les lois des estimateurs des paramètres μ et σ^2 de la variable aléatoire X qui nous intéresse. Ces estimateurs : $\overline{X}$, la moyenne d'échantillon et S^2, la variance d'échantillon, ont été introduits dans le chapitre "échantillonage". Rappelons que ces estimateurs sont des variables aléatoires ; nous avons besoin de leurs lois pour déterminer des intervalles de confiance ou faire des tests d'hypothèses.

Loi de $\overline{X}$ lorsque σ est connu.

On distingue 2 cas :

i) si la loi de X est une loi normale $\mathcal{N}(\mu,\sigma)$ alors la loi de $\overline{X}$ est encore une loi normale de même moyenne μ mais d'écart-type $\frac{\sigma}{\sqrt{n}}$. Autrement dit, la loi de $\frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$ est une loi normale centrée réduite $\mathcal{N}(0,1)$.

ii) si X est de loi quelconque, de moyenne μ et d'écart-type σ connus et si n est assez grand, alors la loi de $\overline{X}$ est approximativement une loi normale de même moyenne μ mais d'écart-type $\frac{\sigma}{\sqrt{n}}$. Autrement dit, la loi de $\frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$ est approximativement une loi normale centrée réduite $\mathcal{N}(0,1)$.

Loi de S^2.

On se place dans le cas où la variable X qui nous intéresse suit une loi normale $\mathcal{N}(\mu,\sigma)$.

La variance d'échantillon S^2 peut aussi s'écrire :

$$S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2-(\overline{X}-\mu)^2$$

En divisant par σ^2 et multipliant par n, on obtient :

$$\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2 = \frac{nS^2}{\sigma^2} + \left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\right)^2$$

Le premier membre est une variable qui suit une loi du χ^2_n. En vertu du théorème de Cochran, qui est une version probabiliste étendue du théorème de Pythagore, on peut montrer que les deux termes du deuxième membre sont indépendants et suivent des lois du χ^2 avec, respectivement, n-1 et 1 degrés de liberté.

Résultat 1. $\frac{nS^2}{\sigma^2}$ suit une loi du χ^2_{n-1}

Résultat 2. $\overline{X}$ et S^2 sont des variables indépendantes.

Résultat 3.

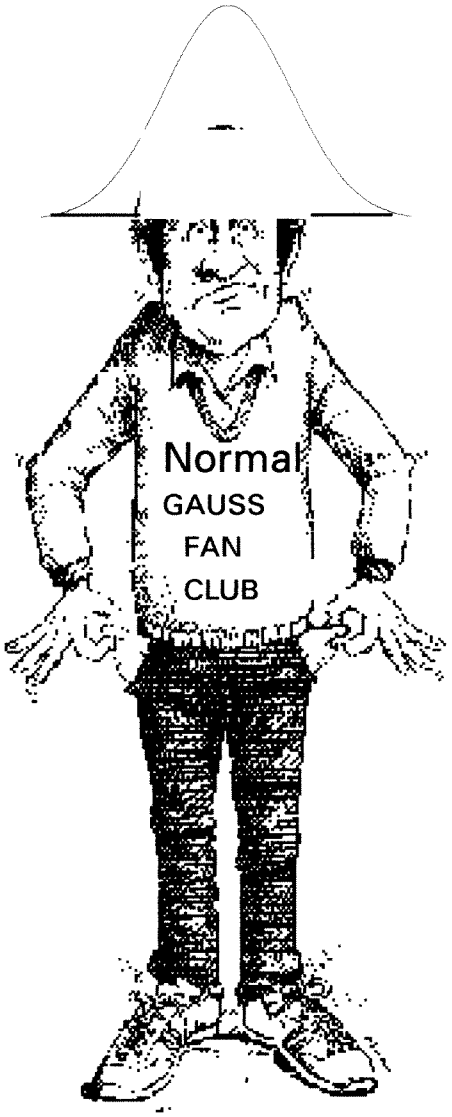
Puisque $\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$ suit une loi normale $\mathcal{N}(0\,;\,1)$, indépendamment de $\frac{nS^2}{\sigma^2}$ qui suit une

loi du χ^2_{n-1}, la variable T_{n-1} définie par :

$$T_{n-1} = \frac{\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}}{\sqrt{\frac{nS^2}{(n-1)\sigma^2}}} = \frac{\overline{X}-\mu}{S/\sqrt{n-1}}$$

suit une loi de Student à n-1 degrés de liberté.
Cette variable sera très utile car elle ne dépend pas de σ et pourra donc être utilisée dans tous les cas où σ est inconnu.
Lorsque le nombre de degrés de liberté est assez grand (on préconise souvent supérieur à 30), la loi de Student peut être assimilée à une loi normale $\mathcal{N}(0,1)$.

CE DOCUMENT S'INSPIRE D'UN COURS ÉLABORÉ À L'UNIVERSITE LOUIS PASTEUR DANS LE CADRE DU CENTRE D'ETUDES STATISTIQUES.


Normal
GAUSS
FAN
CLUB

Document pour le professeur

Estimation
estimation ponctuelle
estimation par intervalle

Objectif:

Présenter divers aspects de l'estimation des paramètres :

- moyenne
- variance
- écart type
- proportion

Estimation

Un problème d'estimation.

Une fabrique de bouteilles en verre utilise une machine pour la production d'un certain type de bouteilles ; on admet que leur masse suit une loi normale $\mathcal{N}(\mu,\sigma)$, μ et σ étant des constantes attachées à la machine.
20 bouteilles sont prélevées au hasard de la production ; voici leurs masses en g :

402	405	405	412	411	406	407	408	402	404
400	399	401	408	398	402	407	406	403	400.

i) Calculer la moyenne, la variance, l'écart-type observés sur cet échantillon.
Quelle est l' estimation (ponctuelle) de μ ?
Donner une estimation (ponctuelle) de σ^2 et de σ

ii) Le constructeur affirme que la vraie valeur de σ est 3,72 g.
Donner un intervalle de confiance pour μ aux seuils de 1%, 5%, 10% et 20%.

iii) Même question si l'on se méfie du constructeur !

iv) On appelle p la proportion de bouteilles produites par cette machine et présentant un défaut.
On choisit au hasard 1200 bouteilles dans la production : 6 d'entre-elles présentent un défaut.
Quelle est l' estimation (ponctuelle) de p ?
Donner un intervalle de confiance pour p aux seuils de 1%, 5%, 10% et 20%.

Ce problème servira à illustrer les diverses notions dont il va être question.

Estimation ponctuelle

Rappelons qu'un **estimateur** est une fonction des variables aléatoires $X_1,.....,X_n$ constituant l'échantillon. Un estimateur est utilisé pour estimer un paramètre.
On appelle **estimation** une réalisation de l'estimateur obtenue à partir d'une réalisation de l'échantillon , notée : $x_1,.....,x_n$

Moyenne observée, variance observée et écart-type observé

Ces expressions désignent les notions classiques de statistique descriptive, qui permettent de donner des éléments de position centrale et de dispersion de l'échantillon observé.

La moyenne observée $\overline{x}$ vaut $\overline{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$

La variance observée vaut $s^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \overline{x})^2$

L'écart-type observé s est évidemment la racine carrée de s^2 .

 Dans notre exemple : $\overline{x}$ = 404,3 g s^2 = 14,31 s= 3,78 g

La quantité s est souvent désignée par σ_n sur les calculatrices.

Estimateur et estimation de μ, moyenne théorique

$\overline{X}$ est un estimateur sans biais et convergent de μ.

Dans le cadre d'un problème d'estimation, la moyenne observée $\overline{x}$ est considérée comme une estimation de μ ; en effet, $\overline{x}$ est une réalisation de la variable aléatoire $\overline{X}$.

Cette phrase contient les trois notions de moyenne qui sont à utiliser lorsqu'on fait de l'estimation.
Il y a μ, la moyenne dite théorique, qui est l'espérance de la variable aléatoire X qui nous intéresse. C'est un paramètre fixe mais en général inconnu. Dans notre exemple, X est la masse en grammes d'une bouteille et μ son espérance. On pourrait connaître μ si on était capable de mesurer les masses de toutes les bouteilles produites.

Il y a $\overline{X}$, la moyenne d'échantillon, qui est une variable aléatoire ; sous certaines conditions, on peut déterminer sa loi (voir le chapitre petit herbier des lois). Son espérance est aussi μ .

Il y a enfin $\overline{x}$, la moyenne observée, calculée sur l'échantillon des 20 bouteilles effectivement observées. Cette moyenne dépend évidemment de l'échantillon : d'autres bouteilles nous donneraient vraisemblablement une autre valeur.

Estimateur et estimation de σ^2, variance théorique

La variance d'échantillon corrigée $S_c^2 = \frac{n}{n-1} S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})^2$ est un estimateur sans biais et convergent de σ^2.

La variance observée corrigée $s_c^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \overline{x})^2$ est une estimation de σ^2

Dans notre exemple $s_c^2 = 15{,}06$ et $s_c = 3{,}88$.
La quantité s_c est souvent désignée par σ_{n-1} sur les calculatrices.

Estimateur et estimation de p, proportion théorique, paramètre d'une loi de Bernoulli $\mathfrak{B}(1,p)$

Soit Y la variable définie sur l'ensemble des bouteilles produites, qui prend la valeur 1 lorsque la bouteile est défecteuse, 0 sinon. On dit que Y est une variable de Bernoulli de loi $\mathfrak{B}(1,p)$. Le paramètre p est l'espérance de Y, et l'on pourrait se contenter de rappeler que $\overline{Y}$ est donc un estimateur sans biais et convergent de p.

En général, on ne se contente pas de ce rappel, en particulier parce que, si p est bien l'espérance de Y, sa variance est p (1 - p) et le paramètre p y intervient donc aussi. Nous allons "reconstruire" l'estimateur de p, que nous appellerons F.

Soient $Y_1,....,Y_n$ un n-échantillon issu de Y, c'est-à-dire n variables indépendantes et de même loi $\mathfrak{B}(1,p)$ que Y.

Nous avons vu , dans le § "approximation normale de la loi binomiale" que si S_n est la somme de n variables aléatoires Y_i indépendantes et de même loi de Bernoulli $\mathfrak{B}(1,p)$ alors la loi de S_n est une loi binomiale $\mathfrak{B}(n,p)$, .
Donc si n est assez grand, la loi de S_n peut être approchée par une loi normale $\mathfrak{N}(np, \sqrt{np\ (1-\ p)}$.
On peut admettre que cette approximation est acceptable dès que $np \geq 10$ et $n\ (1-p) \geq 10$
La variable $F = \frac{S_n}{n}$ suit donc dans les mêmes conditions une loi normale d'espérance p et de variance $\frac{p(1-p)}{n}$.
F est un estimateur sans biais et convergent de p. Sa réalisation f sur l'échantillon effectivement prélevé est donc une estimation de p.

Dans notre exemple (question iv), l'estimation (ponctuelle) de p est :

$$f = \frac{6}{1200} = 0{,}005$$

Estimation par intervalles.

Estimation par intervalle : intervalle de confiance pour μ

PREMIER CAS : X SUIT UNE LOI NORMALE DE MOYENNE μ INCONNUE ET D'ÉCART-TYPE σ CONNU.

Première présentation :

On choisit un niveau de confiance $1 - \alpha$. Ceci nous permet de lire dans une table de la loi normale $\mathcal{N}(0,1)$ une valeur z_α, qui dépend de α, et telle que

$$P\left(- z_\alpha < \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} < + z_\alpha \right) = 1 - \alpha$$

Cette relation est équivalente à

$$P\left(\mu - z_\alpha \frac{\sigma}{\sqrt{n}} < \overline{X} < \mu + z_\alpha \frac{\sigma}{\sqrt{n}} \right) = 1 - \alpha$$

et aussi à la suivante, compte tenu de la symétrie de l'intervalle

$$P\left(\overline{X} - z_\alpha \frac{\sigma}{\sqrt{n}} < \mu < \overline{X} + z_\alpha \frac{\sigma}{\sqrt{n}} \right) = 1 - \alpha$$

Pour $\alpha = 5\ \%$; $z_\alpha = 1{,}96$

L'intervalle de confiance i_α pour μ est l'ensemble des valeurs μ "compatibles" avec l'observation $\overline{x}$ c'est à dire telles que

$$\overline{x} \in \left] \mu - z_\alpha \frac{\sigma}{\sqrt{n}} \ ; \ \mu + z_\alpha \frac{\sigma}{\sqrt{n}} \right[$$

Toutes les valeurs de μ qui rendent vraie cette relation sont telles que

$$\mu \in \left] \overline{x} - z_\alpha \frac{\sigma}{\sqrt{n}} \ ; \ \overline{x} + z_\alpha \frac{\sigma}{\sqrt{n}} \right[$$

L'intervalle de confiance i_α est donc $i_\alpha = \left] \overline{x} - z_\alpha \frac{\sigma}{\sqrt{n}} \ ; \ \overline{x} + z_\alpha \frac{\sigma}{\sqrt{n}} \right[$

Dans notre exemple (question ii), si σ est connu et vaut 3,72 :

Pour $\alpha = 1\ \%$; $z_\alpha = 2{,}58$ et $i_\alpha =]402{,}1\ ;\ 406{,}5[$

Pour $\alpha = 5\ \%$; $z_\alpha = 1{,}96$ et $i_\alpha =]402{,}6\ ;\ 406[$

Pour $\alpha = 10\ \%$; $z_\alpha = 1{,}645$ et $i_\alpha =]402{,}9\ ;\ 405{,}7[$

Pour $\alpha = 20\ \%$; $z_\alpha = 1{,}282$ et $i_\alpha =]403{,}2\ ;\ 405{,}4[$

On ne peut pas dire que.....

Il est clair que l'on ne peut pas affirmer que la vraie valeur de μ appartient à un intervalle de confiance donné associé à un échantillon donné. D'une part , à chaque valeur observée $\bar{x}$, on associe un intervalle de confiance i_α .Un autre échantillon observé peut donner une autre valeur $\bar{x}$ et un autre intervalle de confiance pour le même seuil α. D'autre part , la longueur de l'intervalle dépend du seuil α qui est arbitrairement choisi.

On peut dire que.....

Si on fait un très grand nombre N d'expériences indépendantes, chaque expérience nous donnant un intervalle de confiance i_α, on obtient N intervalles i_α éventuellement différents. Si N est très grand, on peut dire que

- la proportion de ces intervalles qui contiennent effectivement μ est à peu près $1-\alpha$
- si on tire au hasard un intervalle i_α parmi les N intervalles obtenus, la probabilité pour qu'il contienne effectivement μ est à peu près $1-\alpha$
- l'objectif est atteint dans environ 95% ($1-\alpha$) des cas.

Deuxième présentation :

L'intervalle de confiance $\mathbf{I}_\alpha$ pour μ est un **intervalle aléatoire** qui doit être

centré sur $\overline{X}$

et tel que

$$\mathbf{P(I_\alpha \ni \mu) = 1 - \alpha}$$

$1 - \alpha$ est le niveau de confiance choisi (souvent $1 - \alpha = 95\ \%$).
α est le seuil ou risque choisi

La solution de ce problème est l'intervalle aléatoire I_α

$$I_\alpha = \left]\overline{X} - z_\alpha \frac{\sigma}{\sqrt{n}} \ ; \ \overline{X} + z_\alpha \frac{\sigma}{\sqrt{n}}\right[$$

dont les bornes sont des variables aléatoires.
Chaque échantillon observé peut donner une autre valeur $\bar{x}$ et donc une autre réalisation i_α de l'intervalle de confiance (aléatoire) I_α pour le même seuil α.

$$i_\alpha = \left]\overline{x} - z_\alpha \frac{\sigma}{\sqrt{n}} \ ; \ \overline{x} + z_\alpha \frac{\sigma}{\sqrt{n}}\right[$$

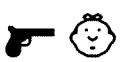
Par abus de dénomination, cette réalisation de l'intervalle de confiance est aussi appelée intervalle de confiance .

Ce qu'on peut dire et ce qu'on ne peut pas dire .....

Les remarques faites dans la première présentation restent valables. Ainsi, si l'on peut écrire

$$P(\mu \in I_\alpha) = P\left(\overline{X} - z_\alpha \frac{\sigma}{\sqrt{n}} < \mu < \overline{X} + z_\alpha \frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$$

puisque les expressions entre parenthèses constituent bien des événements.
On ne peut par contre pas écrire



$$P(\mu \in i_\alpha) = 1 - \alpha$$

ni

$$P\left(\overline{x} - z_\alpha \frac{\sigma}{\sqrt{n}} < \mu < \overline{x} + z_\alpha \frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$$

Les expressions entre parenthèses ne contiennent que des constantes ; ce ne sont donc pas des événements mais des relations entre constantes qui sont soit vraies soit fausses.

DEUXIÈME CAS : X SUIT UNE LOI NORMALE DE MOYENNE μ ET D'ÉCART-TYPE σ TOUS DEUX INCONNUS.

Nous avons vu que dans ce cas la loi de $\dfrac{\overline{X} - \mu}{\dfrac{S}{\sqrt{n-1}}}$ est une loi de Student $\mathcal{T}_{n-1}$ à n-1 degrés de liberté.

On choisit un niveau de confiance $1 - \alpha$. Ceci nous permet de lire dans une table de la loi de Student $\mathcal{T}_{n-1}$ à n-1 degrés de liberté une valeur t_{n-1}, qui dépend de α, et telle que

$$P\left(- t_{n-1} < \frac{\overline{X} - \mu}{\dfrac{S}{\sqrt{n-1}}} < + t_{n-1} \right) = 1 - \alpha$$

Cette relation est équivalente à

$$P\left(\overline{X} - t_{n-1} \frac{S}{\sqrt{n-1}} < \mu < \overline{X} + t_{n-1} \frac{S}{\sqrt{n-1}} \right) = 1 - \alpha$$

On obtient ainsi l'intervalle de confiance (aléatoire) I_α

$$I_\alpha = \left]\overline{X} - t_{n-1}\frac{S}{\sqrt{n-1}} \; ; \; \overline{X} + t_{n-1}\frac{S}{\sqrt{n-1}}\right[$$

dont les bornes sont des variables aléatoires.

La réalisation i_α de l'intervalle de confiance (aléatoire) I_α pour le même seuil α est

$$i_\alpha = \left]\overline{x} - t_{n-1}\frac{s}{\sqrt{n-1}} \; ; \; \overline{x} + t_{n-1}\frac{s}{\sqrt{n-1}}\right[$$

Par abus de dénomination, cette réalisation de l'intervalle de confiance est aussi appelée intervalle de confiance.

Rappel : Si le nombre de degrés de liberté est assez grand (en pratique à partir de 30), t_{n-1} peut être lu dans une table $\mathcal{N}(0,1)$.

Dans notre exemple (question iii) :

Pour $\alpha = 1\ \%$; $t_{19} = 2{,}861$ et $i_\alpha =]401{,}817 \; ; \; 406{,}783[$

Pour $\alpha = 5\ \%$; $t_{19} = 2{,}093$ et $i_\alpha =]402{,}48 \; ; \; 406{,}12[$

Pour $\alpha = 10\ \%$; $t_{19} = 1{,}729$ et $i_\alpha =]402{,}8 \; ; \; 405{,}8[$

Pour $\alpha = 20\ \%$; $t_{19} = 1{,}328$ et $i_\alpha =]403{,}15 \; ; \; 405{,}45[$

Estimation par intervalle : intervalle de confiance pour p, proportion théorique, paramètre d'une loi de Bernoulli $\mathfrak{B}(1,p)$

Première présentation :

On choisit un niveau de confiance 1 - α . Ceci nous permet de lire dans une table de la loi normale $\mathfrak{N}(0,1)$ une valeur z, qui dépend de α, et telle que

$$P\left(-z_\alpha < \frac{F-p}{\sqrt{\frac{pq}{n}}} < +z_\alpha \right) = 1-\alpha$$

avec q=1-p.

Cette relation est équivalente à

$$P\left(p - z_\alpha \sqrt{\frac{pq}{n}} < F < p + z_\alpha \sqrt{\frac{pq}{n}} \right) = 1-\alpha$$

et aussi à la suivante, compte tenu de la symétrie de l'intervalle

$$P\left(F - z_\alpha \sqrt{\frac{pq}{n}} < p < F + z_\alpha \sqrt{\frac{pq}{n}} \right) = 1-\alpha$$

Pour $\alpha = 5\ \%$; $z_\alpha = 1{,}96$

L'intervalle de confiance i_α pour p est l'ensemble des valeurs p "compatibles" avec l'observation f c'est à dire telles que

$$f \in \left]p - z_\alpha\sqrt{\frac{pq}{n}} \ ; \ p + z_\alpha\sqrt{\frac{pq}{n}}\right[$$

Toutes les valeurs de p qui rendent vraie cette relation sont telles que

$$p \in \left]f - z_\alpha\sqrt{\frac{pq}{n}} \ ; \ f + z_\alpha\sqrt{\frac{pq}{n}}\right[$$

On souhaite que les bornes de l'intervalle de confiance ne dépendent pas de p, qui est inconnu.

1ère méthode : majoration.

Le produit p(1-p) étant toujours inférieur à 1/4, on pourrait proposer comme intervalle de confiance "maximum"

$$i_\alpha = \left]f - \frac{z_\alpha}{2\sqrt{n}} \ ; \ f + \frac{z_\alpha}{2\sqrt{n}}\right[$$

2ème méthode : résoudre l'inéquation.

$$p \in \left]f - z_\alpha\sqrt{\frac{pq}{n}} \ ; \ f + z_\alpha\sqrt{\frac{pq}{n}}\right[$$

équivaut à l'inéquation suivante d'inconnue p

$$|f - p| < z_\alpha\sqrt{\frac{pq}{n}}$$

On élève les deux membres au carré et on résout l'inéquation du second degré en p. On obtient

$$\frac{f + \frac{z_\alpha^2}{2n} - z_\alpha\sqrt{\frac{f(1-f)}{n} + \frac{z_\alpha^2}{4n^2}}}{1 + \frac{z_\alpha^2}{n}} < p < \frac{f + \frac{z_\alpha^2}{2n} + z_\alpha\sqrt{\frac{f(1-f)}{n} + \frac{z_\alpha^2}{4n^2}}}{1 + \frac{z_\alpha^2}{n}}$$

On pourrait s'arrêter là mais personne (ou presque) ne le fait et on préfère utiliser une valeur approchée des deux bornes, ce qui donne l'encadrement suivant :

$$f - z_\alpha\sqrt{\frac{f(1-f)}{n}} < p < f + z_\alpha\sqrt{\frac{f(1-f)}{n}}$$

d'où

$$i_\alpha = \left]f - z_\alpha\sqrt{\frac{f(1-f)}{n}} \ ; \ f + z_\alpha\sqrt{\frac{f(1-f)}{n}}\right[$$

Dans notre exemple (question iv)

Pour $\alpha = 1\ \%$; $z_\alpha = 2{,}58$ et $i_\alpha = [0 \ ; \ 0{,}01026[$

Pour $\alpha = 5\ \%$; $z_\alpha = 1{,}96$ et $i_a =]0{,}001 \ ; \ 0{,}009[$

Pour $\alpha = 10\ \%$; $z_\alpha = 1{,}645$ et $i_a =]0{,}00165 \ ; \ 0{,}00835[$

Pour $\alpha = 20\ \%$; $z_\alpha = 1{,}282$ et $i_\alpha =]0{,}00239 \ ; \ 0{,}00761[$

Deuxième présentation :

L'intervalle de confiance I_α pour p est un **intervalle aléatoire** qui doit être

centré sur F

et tel que

$$\mathbf{P(I_\alpha \ni p) = 1 - \alpha}$$

$1 - \alpha$ est le niveau de confiance choisi (souvent $1 - \alpha = 95\ \%$).

α est le seuil ou risque choisi

La solution de ce problème est l'intervalle aléatoire I_α

$$I_\alpha = \left]F - z_\alpha\sqrt{\frac{pq}{n}} \ ; \ F + z_\alpha\sqrt{\frac{pq}{n}}\right[$$

dont les bornes sont non seulement aléatoires mais aussi dépendent de p qui est inconnu.

Chaque échantillon observé peut donner une autre valeur f et donc une autre réalisation i_α de l'intervalle de confiance (aléatoire) I_α pour le même seuil α

$$i_\alpha = \left]f - z_\alpha\sqrt{\frac{pq}{n}} \ ; \ f + z_\alpha\sqrt{\frac{pq}{n}}\right[$$

et on est ainsi ramené exactement au problème de la 1ère présentation : la présence de p (inconnu) dans cette expression.

Par abus de dénomination, cette réalisation de l'intervalle de confiance est aussi appelée intervalle de confiance .

On entend dire .....

"puisque, dans l'expression de $i_\alpha = \left]f - z_\alpha\sqrt{\frac{pq}{n}} \ ; \ f + z_\alpha\sqrt{\frac{pq}{n}}\right[$ on retrouve p qui est inconnu, remplaçons le par son estimation f"

Pourquoi ? pourquoi pas, puisqu'on retrouve l'intervalle que nous avons justifié précédemment.

On entend dire aussi .....

"il vaut mieux remplacer $\frac{p(1-p)}{n}$ par son estimation sans biais $\frac{f(1-f)}{n-1}$"

pourquoi pas ?

Il est exact que l'espérance de la variable aléatoire $\frac{F(1-F)}{n-1}$ est $\frac{p(1-p)}{n}$

En effet :

$$\frac{E(F)-E(F^2)}{n-1} = \frac{E(F)-Var(F)-E(F)^2}{n-1} = \frac{1}{n-1}\left(p-\frac{p(1-p)}{n}-p^2\right) = \frac{1}{n-1}\left(\frac{(n-1)p(1-p)}{n}\right)$$

Mais d'une part, l'espérance de $\sqrt{\frac{F(1-F)}{n-1}}$ n'est pas $\sqrt{\frac{p(1-p)}{n}}$ et d'autre part on ne voit pas en vertu de quel critère l'intervalle ainsi obtenu serait meilleur. Il est certain qu'il est un peu plus grand, qu'il contient le précédent et que, lorsque n est grand il n'est pas très différent. De là à être meilleur.....

Attention !

Les réalisations i_α des intervalles de confiance I_α (qu'on appelle aussi les fourchettes les soirs d'élections) dépendent de l'échantillon effectivement prélevé et ne sont plus du tout aléatoires. On ne peut donc pas dire, comme on l'entend souvent :

"Le pourcentage de voix du candidat  est compris entre 49% et 51%, avec une probabilité de 95%" ou

"Nous sommes sûrs, à 95%, que le pourcentage de voix du candidat  est compris entre 49% et 51%".

En effet p est fixe, l'intervalle i_α est fixe et donc p lui appartient ou ne lui appartient pas, mais on ne peut pas probabiliser ceci.

La seule chose que l'on puisse affirmer est que : sur un très grand nombre d'échantillons effectivement réalisés, la proportion des intervalles i_α comprenant p sera approximativement égale à 1 - α (95%)

CE DOCUMENT S'INSPIRE D'UN COURS ÉLABORÉ À L'UNIVERSITE LOUIS PASTEUR DANS LE CADRE DU CENTRE D'ETUDES STATISTIQUES.

Document pour le professeur

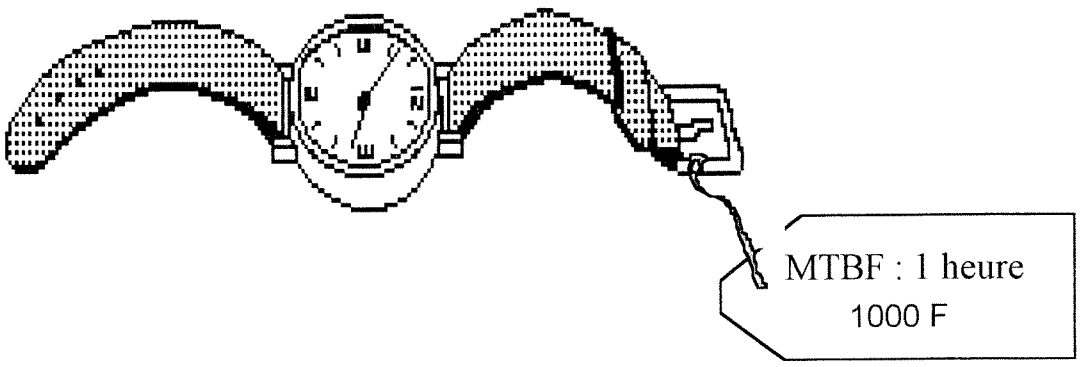
Ce document comporte aussi des exemples utilisables aves les élèves

Fiabilité

Objectif:

Présenter le vocabulaire de la fiabilité utilisé par les professionnels de la maintenance

Utiliser le modèle de Weibull dans des cas simples.



A PROPOS DE LA FIABILITE

Introduction :

Les défaillances sont à la maintenance ce que les maladies sont à la médecine : leur raison d'exister. (La fonction maintenance. François Monchy).

L'AFNOR donne la définition suivante de la fiabilité : La fiabilité est la caractéristique d'un dispositif qui s'exprime par la probabilité pour ce dispositif d'accomplir une fonction requise, dans des conditions déterminées, pendant une période donnée.

La terminologie de la fiabilité, s'applique aussi bien à des grands nombres de dispositifs identiques, tels que les résistances, qu'à un seul et unique dispositif.

La définition précédente comporte quatre concepts :

1. La **Probabilité** : ce concept justifie le présent travail. Pour un dispositif donné, on est amené à évaluer sa probabilité de *bon fonctionnement* ou celle d'*accomplir une mission.*

2. La **Fonction requise** : cette définition implique un seuil d'admissibilité en deça duquel *le service attendu n'est pas rendu.*

3. La **Condition d'utilisation** : ce concept met en évidence l'importance de l'environnement et de ses variations, des contraintes mécaniques ou autres (un même matériel placé dans deux environnements de fonctionnement différents n'aura pas la même fiabilité).

4. La **Période de temps** : cette notion est à utiliser dans un sens plus général. Elle peut être remplacée par un *nombre de cycles*, une *distance* ... etc.

I-**Principales définitions :**

1.1 : **Fonction de fiabilité R(t) :**

Soit T une variable aléatoire continue, à valeurs dans $\mathbb{R}^+$ et de densité de probabilité f.
Lorsque la variable aléatoire T mesure la durée entre deux réparations, appelée TBF (temps de bon fonctionnement) ou mesure la durée de vie s'il n'y a pas réparation, sa fonction de répartition définie par :

$$F(t) = \mathbb{P}\{T \leq t\}$$

est appelée **la fonction de défaillance.**
Son complément à la probabilité totale est *la fonction de survie* ou *fonction de fiabilité* du matériel :

$$R(t) = 1 - F(t) = \mathbb{P}\{T > t\}.$$

C'est donc la probabilité que le matériel fonctionne jusqu'au temps t.

La fonction densité $f(t)$ et la fonction de répartition vérifient (si F est dérivable) : $F'(t) = f(t)$.
La probabilité de panne entre les instants 0 et t est donc :

$$F(t) = \int_0^t f(x)dx \ .$$

1.2 : **Taux de défaillance instantané.**

Définition : le taux d'avarie (ou taux de défaillance) au cours d'une période $[t, t + \Delta t]$ est le nombre de défaillants au cours de cette période rapporté au nombre de survivants au début de la même période et à la durée de cette période.

Exemple : on a étudié 100 moteurs durant une période allant de 9 000 heures à 10 000 heures. 60 défaillances ont été enregistrées. Quel est le taux de défaillance durant cette période ?

Réponse : Désignons par $N(t)$, le nombre de survivants à l'instant t. Durant la période considérée le taux de défaillance est :

$$\lambda = \frac{N(t) - N(t + \Delta t)}{\Delta t \, N(t_0)} = \frac{60}{1000 \cdot 100} = 6.10^{-4} \, panne/heure.$$

Mais pourquoi introduire un tel taux ?

Comme la probabilité de défaillance entre les instants t et $t + \Delta t$ est approximativement $f(t)\Delta t$. On a :

$$\begin{aligned} f(t)\Delta t &\approx \mathbb{P}(t \leq T \leq t + \Delta t) \\ &\approx \mathbb{P}\{(0 \textit{ panne sur } [0, t[) \textit{ et } (\textit{panne dans } [t, t + \Delta t])\}. \end{aligned}$$

Or :

$$P\{(0 \textit{ panne sur } [0, t[)\} = R(t),$$

donc

$$f(t)\Delta t \approx R(t)\,\mathbb{P}\{(\textit{panne dans }]t, t + \Delta t])|(0 \textit{ panne sur } [0.t])\}$$

et

$$\frac{f(t)}{R(t)} \approx \frac{\mathbb{P}\{(\textit{panne dans }]t, t + \Delta t])|(0 \textit{ panne sur } [0.t])\}}{\Delta t}$$

d'où l'utilité d'introduire un taux instantané de panne $\lambda(t)$ tel que :

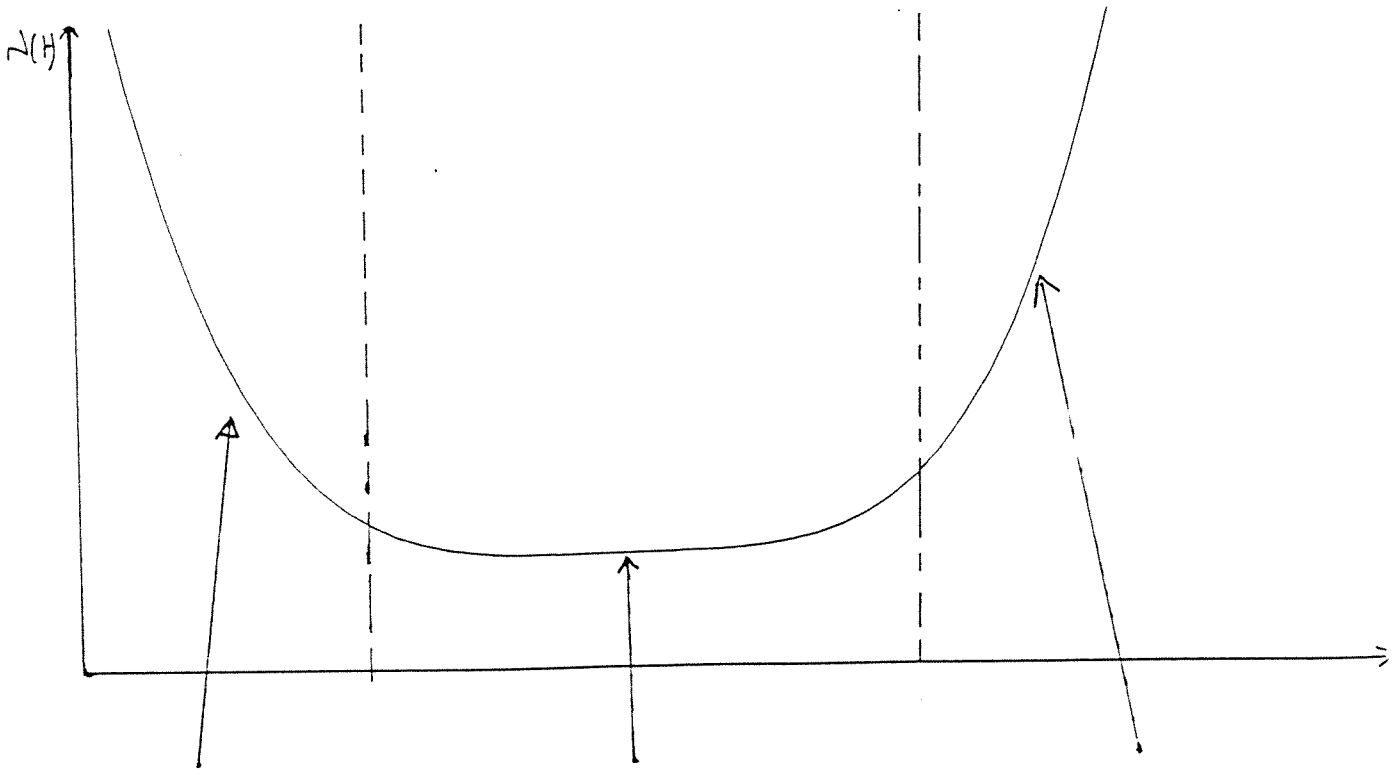
$$\lambda(t) = \lim_{\Delta t \longrightarrow 0} \frac{\mathbb{P}\{(\textit{panne dans }]t, t + \Delta t])|(0 \textit{ panne sur } [0.t])\}}{\Delta t}.$$

On a :

$$\lambda(t) = \frac{f(t)}{R(t)} = \frac{f(t)}{1 - F(t)} = -\frac{R'(t)}{R(t)}.$$

On conçoit que l'on doit pratiquer la maintenance préventive, si la fiabilité $R(t)$ est en dessous d'un seuil minimum (en général 90 %) ou si l'on détecte une évolution du taux de défaillances vers un niveau maximal (seuil d'usure).

On constate que pour la plupart des matériels, la courbe représentive du taux d'avarie en fonction du temps a la forme suivante dite **courbe en baignoire** :



Période de rodage Période de maturité Période d'usure

Le modèle exponentiel :

Le modèle le plus simple est obtenu lorsque le taux d'avarie est constant. Tous les dispositifs sont concernés par cette situation durant la période dite de maturité. Dans ce cas :

$$\lambda(t) = \lambda = -\frac{R'(t)}{R(t)}$$

donne $R(t) = \exp(-\lambda t)$, $F(t) = 1 - \exp(-\lambda t)$ et $f(t) = F'(t) = \lambda \exp(-\lambda t)$.

1.3.Etude de la fiabilité et MTBF* :

On a vu que $\lambda(t) = \frac{F'(t)}{1-F(t)}$. Par intégration sur l'intervalle $[0,t]$, on obtient :

$$\begin{aligned}\int_0^t \lambda(x)dx &= \int_0^t \frac{F'(x)}{1-F(x)}dx \\ &= -\ln(1-F(t)) \cdot\end{aligned}$$

Donc

$$F(t) = 1 - \exp\Big(-\int_0^t \lambda(x)dx\Big),$$

$$R(t) = \exp\Big(-\int_0^t \lambda(x)dx\Big),$$

et

$$f(t) = \lambda(t)\exp\Big(-\int_0^t \lambda(x)dx\Big)\cdot$$

Nous avons ici les expressions les plus générales liant le taux de défaillance aux lois de fiabilité : F, f et R.

MTBF : c'est la moyenne des temps de bon fonctionnement. Elle représente l'espérance mathématique de la variable aléatoire T :

$$\begin{aligned}MTBF &= \int_0^\infty tf(t)dt \\ &= \int_0^\infty \mathbb{P}\{T > t\}dt \\ &= \int_0^\infty R(t)dt\cdot\end{aligned}$$

La deuxième égalité est une autre manière de définir l'espérance mathématique de la variable aléatoire T.

*Moyenne des Temps de Bon fonctionnement

Exemples :

1)*Cacul de la MTBF dans le cas de la loi exponentielle :*

Comme on l'a vu ci-dessus, pour la loi exponentielle :

$$f(t) = \lambda \exp(-\lambda t).$$

Donc

$$MTBF = \int_0^\infty t f(t) dt = \int_0^\infty \lambda t \exp(-\lambda t) dt.$$

Une intégration par parties donne :

$$\begin{aligned} MTBF &= [-te^{-\lambda t}]_0^\infty + \int_0^\infty \exp(-\lambda t) dt \\ &= 0 + [-\frac{1}{\lambda} e^{-\lambda t}]_0^\infty = \frac{1}{\lambda}. \end{aligned}$$

Donc

$$MTBF = \frac{1}{\lambda}.$$

Remarque :R(t)=$e^{-\lambda t}$ donc $R(\frac{1}{\lambda}) = e^{-1} \approx 0.368$. La probabilité que le matériel fonctionne jusqu'à sa MTBFest dans ce cas de 0.368.

2)*Cas non exponentielle*(BTS-Maintenance Industrielle 1992).

Des pièces métalliques de forme parallélépipédique sont fabriquées par des machines pour lesquelles l'étude des temps de bon fonctionnement, exprimés en mois, conduit à la fonction de fiabilité R telle que :

$$R(t) = (0.05t + 1) e^{-0.1t}.$$

a) Calculer $R'(t)$ et vérifier ainsi que la fonction R est bien décroissante sur $[0, \infty[$.

b) Calculer à 10^{-2} près :

-la probabilité qu'une machine fonctionne plus de 10 mois

-la probabilité qu'une machine tombe en panne au cours de la première année.

-Calculer la MTBF.

Réponses :

a) R est dérivable sur $\mathbb{R}^+$ et $R'(t) = (-0,005t - 0,05)e^{-0,1t}$. Pour $t \in [0, +\infty[$, $R'(t) < 0$. R est donc strictement décroissante sur $[0, +\infty[$.

b) Si on note T la variable aléatoire donnant le temps de bon fonctionnement de la machine, on sait que la fiabilité R s'exprime par $R(t) = \mathbb{P}(T > t)$. La probabilité que la machine fonctionne plus de 10 mois est $R(10) = \mathbb{P}(T > 10)$ =R(10)=(0,05 . 10+1)e$^{-0,1.10} = 1,5e^{-1} \approx 0,55$.

De même la probabilité qu'une machine tombe en panne au cours de la première année est : $\mathbb{P}(T \leq 12) = 1 - \mathbb{P}(T > 12) = 1 - R(12) \approx 0,52$.

$$\begin{aligned} MTBF &= \int_0^{\infty} R(t)dt \\ &= \int_0^{\infty} (0.05t + 1)\, e^{-0.1t} dt \\ &= \lim_{\alpha \longrightarrow \infty} [(-0,5t - 15)e^{-0,1t}]_0^{\alpha} \\ &= \lim_{\alpha \longrightarrow \infty} (-0,5\alpha - 15)e^{-0,1\alpha} + 15 \\ &= 15\ mois. \end{aligned}$$

II-Loi de WEIBULL.

Le modèle exponentiel n'est applicable qu'aux cas où le taux de défaillance λ est constant. Afin d'élargir le domaine d'utilisation du modèle aux différentes périodes de vie du matériel, Weibull a proposé une loi de probabilité comportant trois paramètres.

2.1. Fonction densité :

Une variable aléatoire T suit une loi de Weibull, si sa densité de probabilité est :

$$f(t) = \frac{\beta}{\eta}(\frac{t-\gamma}{\eta})^{\beta-1} \exp -(\frac{t-\gamma}{\eta})^{\beta} \qquad pour\ t > \gamma.$$

β : c'est le paramètre de forme, $\beta > 0$.

η : c'est le paramètre d'echelle, $\eta > 0$.

γ : c'est le paramètre de position, $-\infty < \gamma < +\infty$.

Un choix judicieux des trois paramètres permet à la loi de Weibull d'ajuster correctement une large classe de résultats expérimentaux et opérationnels. C'est pour cette qualité que la loi de Weibull a été largement utilisée en fiabilité.

2.2. Fonction fiabilité et taux de défaillance.

On a la fonction de répartition (fonction défaillance) :

$$F(t) = 1 - e^{-(\frac{t-\gamma}{\eta})^{\beta}}.$$

La fiabilité correspondante est donc :

$$R(t) = 1 - F(t) = e^{-(\frac{t-\gamma}{\eta})^{\beta}}.$$

Remarque : a) pour $\gamma = 0$ et $\beta = 1$, le modèle de Weibull se réduit au modèle exponentiel :

$R(t) = 1 - F(t) = \exp -\frac{(t}{\eta)}$ et $f(t) = \frac{1}{\eta} \exp -\frac{(t}{\eta)}, t \geq 0$ et dans ce cas $\lambda = \frac{1}{\eta} = \frac{1}{MTBF}$.

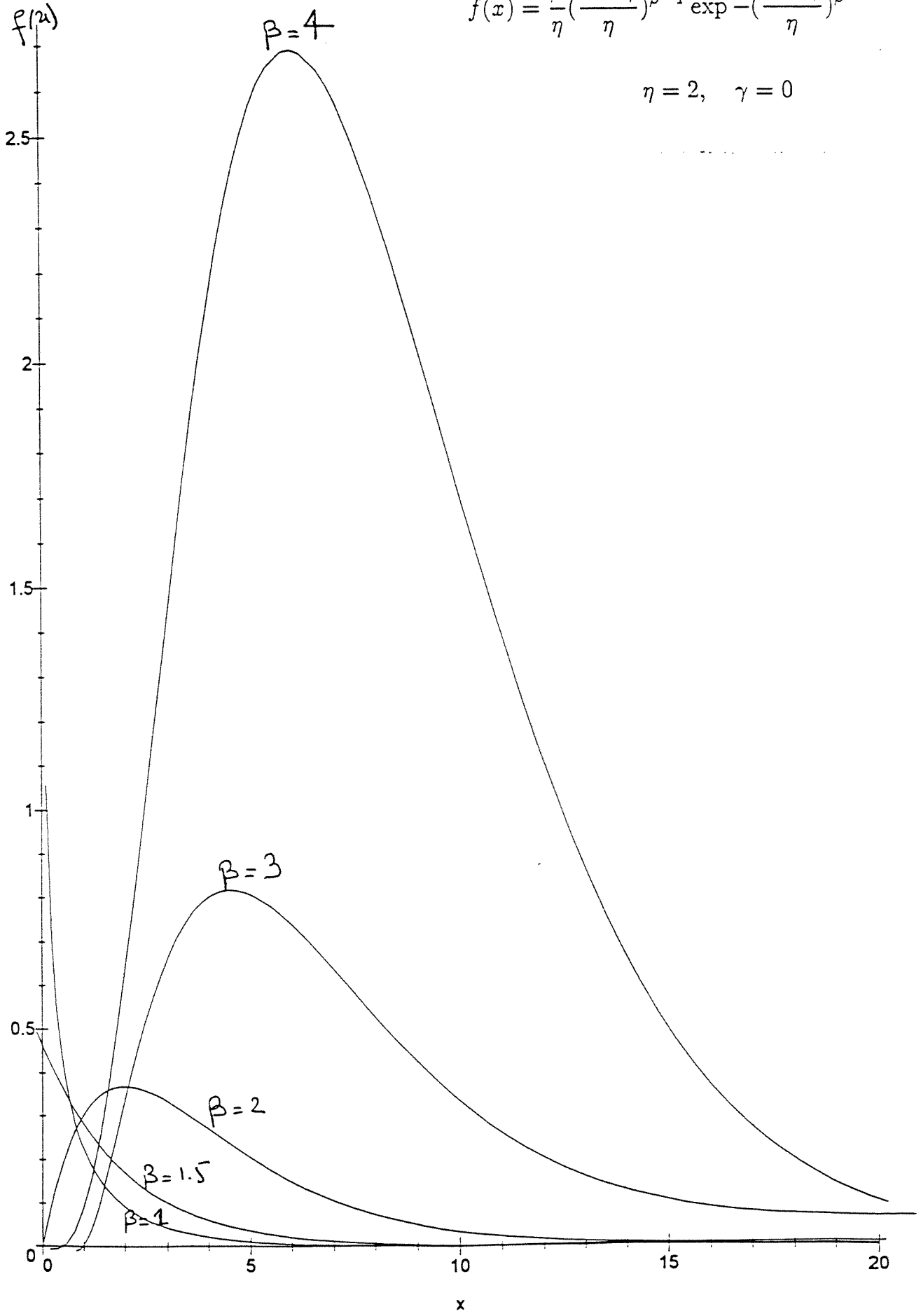
Le taux instantané de défaillance $\lambda(t)$ est donné pour tous $t \geq \gamma$, $\beta > 0$ et $\eta > 0$, par :

$$\lambda(t) = \frac{f(t)}{1 - F(t)} = \frac{\beta}{\eta}(\frac{t-\gamma}{\eta})^{\beta-1}.$$

Polymorphisme de la loi de Weibull sous l'influence du paramètre de forme β.

$$f(x) = \frac{\beta}{\eta}\left(\frac{x-\gamma}{\eta}\right)^{\beta-1} \exp -\left(\frac{x-\gamma}{\eta}\right)^{\beta}$$

$$\eta = 2, \quad \gamma = 0$$



On remarque que :
Si $\beta < 1$ alors $\lambda(t)$ décroît : période de jeunesse (rodage ou déverminage).
Si $\beta = 1$ alors le taux $\lambda(t)$ est constant : indépendance du processus par rapport au temps.
Si $\beta > 1$ alors $\lambda(t)$ croît : période d'usure, vieillissement du matériel.
Ce qui montre que la loi de Weibull peut s'adapter aux trois périodes de vie du matériel.

2.3. Espérance mathématique (MTBF) et écart-type :

Soit T une variable aléatoire distribuée suivant une loi de Weibull. On démontre alors et nous l'admettrons que l'espérance mathématique de T est :

$$\mathbb{E}(T) = A\eta + \gamma.$$

où $A = \Gamma(1 + 1/\beta)$, Γ étant la fonction d'Euler ; et

$$\mathbb{V}ar(T) = \eta^2 \left(\Gamma(1 + 2/\beta) - (\Gamma(1 + 1/\beta)^2\right).$$

Donc $\sigma(T) = B\eta$ où $B = \sqrt{\left(\Gamma(1 + 2/\beta) - (\Gamma(1 + 1/\beta)^2\right)}$.

En pratique, on dispose de tables numériques de la fonction d'Euler. La détermination des trois paramètres permettra, à partir de ces tables, d'évaluer la MTBF et l'écart-type. C'est l'objectif du paragraphe suivant.

Exemple : si $\beta = 1.5$, on lit dans la table $A = \Gamma(1 + 1/\beta) = 0.90$ et $B = \Gamma(1 + 2/\beta) - (\Gamma(1 + 1/\beta) = 0.61$. Si de plus $\gamma = 0$ et $\eta = 5.7.\ 10^5\ cycles$, alors la

$$MTBF = 0,90.\ 5,7 \quad 10^5\ cycles = 5,1.\ 10^5\ cycles.$$

et

$$\sigma(T) = \sqrt{Var(T)} = 5,7.10^5.0,61 = 3,5.\ 10^5\ cycles.$$

III-**Ajustement Graphique :** la détermination des paramètres.

3.1. Principe : on a vu que, par sa souplesse (voir polymorphisme de la loi de Weibull), le modèle de Weibull s'ajuste correctement à une large classe de données expérimentales. La détermination des trois paramètrtes β, η et γ permet d'ajuster la loi probabiliste à la statistique relevée.
L'historique de fonctionnement d'un matériel détermine ses temps de bon fonctionnement entre deux pannes (TBF) ou la durée de vie (pour les composants électronique).

Ceci donne des **fréquences cumulées de défaillances, constituant une approximation de la fonction défaillance** $F(t)$. Ces fréquences cumulées seront notées $\widetilde{F}(t_i)$.

On porte les points $M(t_i, \widetilde{F}(t_i))$ sur un papier spécial dit de Weibull.

Dans le cas où $\gamma = 0$ qui est le seul cas que nous envisagerons ici, le nuage de points ainsi formé sera ajusté par une droite dite de Weibull.

Avant d'utiliser le papier de Weibull, nous allons donner une justification mathématique de sa conception.

On a vu que la fonction de répartition d'une variable alétoire suivant une loi de Weibull est :

$$F(t) = 1 - \exp\left(-(\frac{t-\gamma}{\eta})^{\beta}\right)$$

et que la fonction fiabilité est :

$$R(t) = 1 - F(t) = \exp\left(-(\frac{t-\gamma}{\eta})^{\beta}\right).$$

Donc

$$\frac{1}{R(t)} = \frac{1}{1-F(t)} \geq 1.$$

Ceci entraîne que :

$$\begin{aligned} \ln\ln\frac{1}{R(t)} &= \ln(\frac{t-\gamma}{\eta})^{\beta} = \beta\ln(\frac{t-\gamma}{\eta}) \\ &= \beta\ln(t-\gamma) - \beta\ln\eta. \end{aligned}$$

En posant $Y = \ln\ln\frac{1}{R(t)}$, $X = \ln(t-\gamma)$ et $C = -\beta\ln\eta$, on obtient :

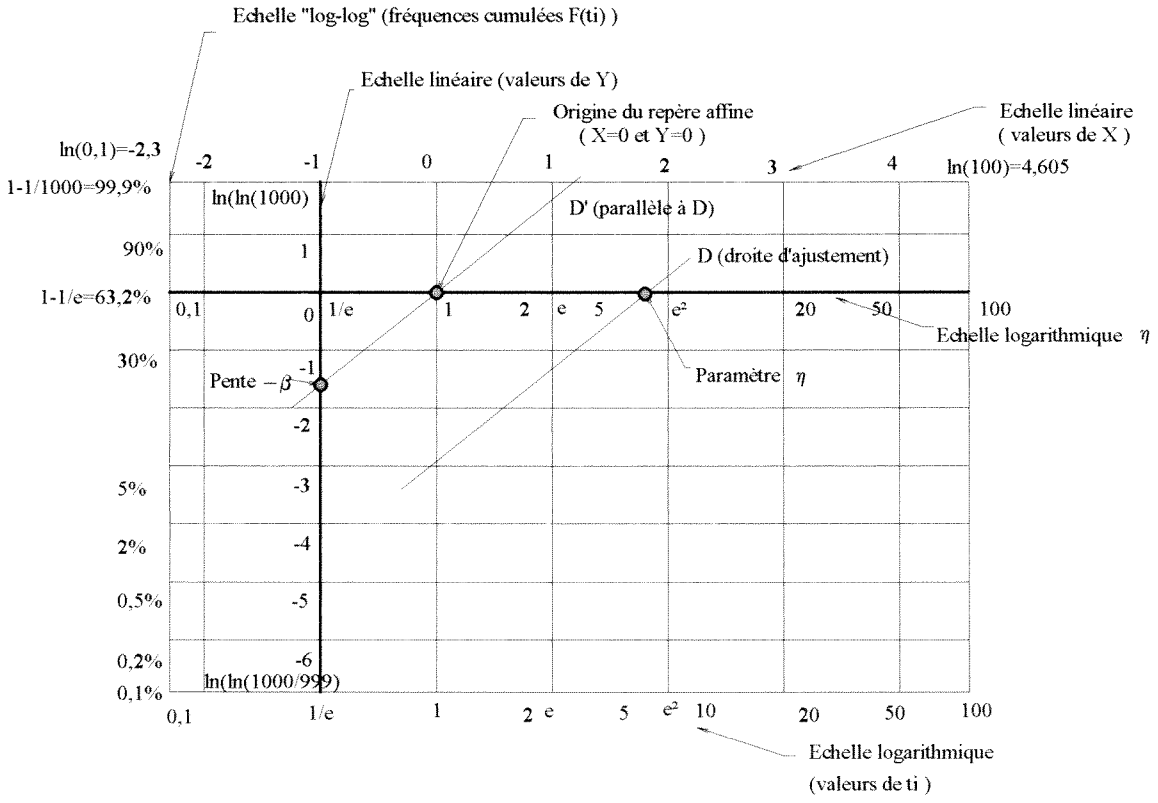
$$Y = \beta X + C$$

D'où une relation affine entre X et Y.

β est la pente de la droite D d'ajustement du nuage de points $\left(t_i, \widetilde{F}(t_i)\right)$ c'est donc aussi la pente de la droite D' parallèle à D.

$Y = 0 \quad \text{si} \quad X = \dfrac{C}{\beta \ln \eta}$.

PAPIER DE WEIBULL



Le papier de Weibull comporte essentiellement :

1. En abscisse, une échelle logarithmique (la distance à l'origine correspond au logarithme de la graduation) placée sur le bord inférieur de la feuille, sur laquelle on portera les temps de bon fonctionnement t_i (ou $t_i - \gamma$) observés.

2. En ordonnée, une échelle "log-log" (la distance à l'origine correspond au logarithme du logarithme de la graduation) placée sur le bord gauche de la feuille, sur laquelle on portera les estimations des fréquences cumulées $\widetilde{F}(t_i)$.

Pour permettre une lecture directe des paramètres d'une loi de Weibull, trois autres axes ont été gradués :

1. Une échelle linéaire verticale sur laquelle on lit bien sûr Y qui est le logarithme du logarithme de la graduation correspondante de l'échelle log-log et qui permettra une lecture de la pente β.

2. Une copie de l'échelle logarithmique du bord inférieur passant par l'origine sur laquelle on lira la valeur du paramètre η.

3. Une échelle linéaire horizontale placée au bord haut de la feuille sur laquelle on lit les valeurs de X.

Lorsque le paramètre $\gamma = 0$, le nuage de points $(t_i, \widetilde{F}(t_i))$ peut être ajusté par une droite D d'équation $Y = \beta X + C$ avec $C = -\beta \ln \eta$. Cette droite coupe l'axe Y=0 au point d'abscisse $X = \ln \eta$ ce qui permet de lire η sur l'axe gradué logarithmiquement.

La parallèle D' à la droite d'ajustement D passant par l'origine du repère affine coupe l'échelle Y placée en X = - 1 au point d'ordonnée $-\beta$ (sur le modèle commercial de papier de Weibull, les signes "-" devant les graduations de cet axe sont d'ailleurs ommis)

3.3. Un exemple d'utilisation du papier de Weibull :

L'exemple suivant est cité dans : F. Monchy, La fonction maintenance. Considérons un lot de six roulements, chargés dans des conditions spécifiques à un essai de durée de vie. Nous avons enregistré les résultats suivants :

N° de roulement	Nombre de cycles avant rupture
1	$4.0 \quad 10^5$
2	$1.3 \quad 10^5$
3	$9.8 \quad 10^5$
4	$2.7 \quad 10^5$
5	$6.6 \quad 10^5$
6	$5.2 \quad 10^5$.

Préparation des données : on commence toujours par ordonner les valeurs des TBF ou des durées de vie enregistrées puis on en déduit un tableau d'approximations de $F(t)$. Il y a 3 cas à considérer selon le nombre N de données.

Si $N > 50$: Par i nous notons le nombre de défaillants à l'instant t_i. Dans ce cas la fréquence cumulée

$$\widetilde{F}(t_i) = \frac{i}{N}$$

est très voisine de la fonction de répartition $F(t)$ de la loi de Weibull.

Quand N<50, on procède à une "correction" du rapport (i/N) qui fournit une meilleure estimation de la fréquence cumulée. Plus précisément :

Si $20 < N \leq 50$, nous utiliserons la formule dite des rangs moyens définie par :

$$\widetilde{F}(t_i) = \frac{i}{N+1},$$

i est toujours le nombre de défaillants et aussi le rang de la défaillance.

Si $N \leq 20$, nous utiliserons la formule d'approximation dite des rangs médians :

$$\widetilde{F}(t_i) = \frac{i - 0.3}{N + 0.4}.$$

Evaluation de la fréquence cumulée $\widetilde{F}(t_i)$: en ordonnant les valeurs des TBF enregistrées et en appliquant la formule des rangs médians : $F(t_i) = \frac{i-0.3}{N+0.4}$, on obtient :

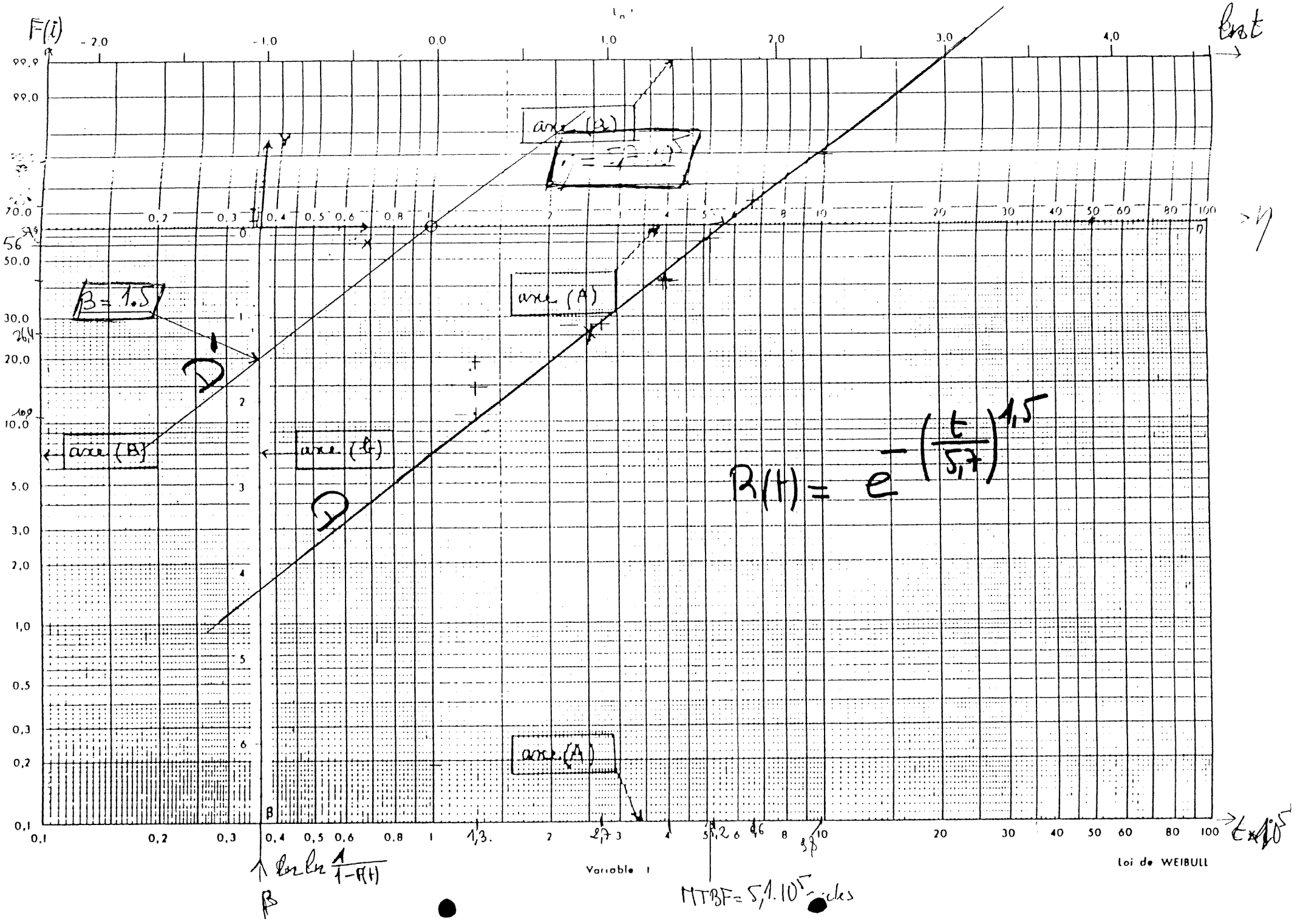
Ordre i	*TBF (cycles avant rupture)*	*F(i) (rangs médians)*	
1	$1,3.10^5$	0,109	10,9%
2	$2,7.10^5$	0,264	26,4%
3	$4,0.10^5$	0,421	42,1%
4	$5,2.10^5$	0,579	57,9%
5	$6,6.10^5$	0,736	73,6%
6	$9,8.10^5$	0,890	89%

Détermination des paramètres de Weibull :

En portant sur le papier de Weibull les couples de points $(t_i, \widetilde{F}(t_i))$, on obtient un nuage de points qui peut être ajusté par une droite D. Donc $\gamma = 0$. La droite D coupe l'axe (η) à l'abscisse $\eta = 5.7\ 10^5$ cycles. La droite D' parallèle à la droite D coupe l'axe (β) à l'ordonnée $\beta = 1.5$.

Exploitation : La table de MTBF donne A=0.90 et B=0.61. Donc la MTBF=5.1 10^5 cycles et l'écart-type vaut 3.5 10^5 cycles.

Fiabilité associée à la MTBF : sur le graphique, on lit $F(t) = 56\%$ i.e : $R(t) = 44\%$. Donc seuls 44% de roulements atteindront la MTBF.


F(i)
ln
ln t
axe (B)
axe (t)
axe (A)
axe (A)
B = 1.5
D
D
R(t) = e^-(t/5,7)^1,5
MTBF = 5,1.10^5 cycles
Variable 1
Loi de WEIBULL
t x 10^5

126

Loi de Weibull - Tables de calcul de la MTBF et de l'écart type					
β	**A**	**B**	β	**A**	**B**
0,2	120,0000	1901,1575	2,5	0,8873	0,3797
0,25	24,0000	199,3590	2,6	0,8882	0,3670
0,3	9,2605	50,0780	2,7	0,8893	0,3552
0,35	5,0291	19,9761	2,8	0,8905	0,3443
0,4	3,3234	10,4382	2,9	0,8917	0,3341
0,45	2,4786	6,4601	3	0,8930	0,3246
0,5	2,0000	4,4721	3,1	0,8943	0,3156
0,55	1,7024	3,3453	3,2	0,8957	0,3072
0,6	1,5046	2,6451	3,3	0,8970	0,2993
0,65	1,3663	2,1789	3,4	0,8984	0,2918
0,7	1,2658	1,8512	3,5	0,8997	0,2847
0,75	1,1906	1,6108	3,6	0,9011	0,2780
0,8	1,1330	1,4282	3,7	0,9025	0,2716
0,85	1,0880	1,2854	3,8	0,9038	0,2656
0,9	1,0522	1,1711	3,9	0,9051	0,2598
0,95	1,0234	1,0777	4	0,9064	0,2543
1	1,0000	1,0000	4,1	0,9077	0,2490
1,05	0,9808	0,9344	4,2	0,9089	0,2440
1,1	0,9649	0,8783	4,3	0,9102	0,2392
1,15	0,9517	0,8297	4,4	0,9114	0,2345
1,2	0,9407	0,7872	4,5	0,9126	0,2301
1,25	0,9314	0,7498	4,6	0,9137	0,2258
1,3	0,9236	0,7164	4,7	0,9149	0,2217
1,35	0,9170	0,6866	4,8	0,9160	0,2178
1,4	0,9114	0,6596	4,9	0,9171	0,2140
1,45	0,9067	0,6352	5	0,9182	0,2103
1,5	0,9027	0,6129	5,1	0,9192	0,2068
1,55	0,8994	0,5925	5,2	0,9202	0,2034
1,6	0,8966	0,5737	5,3	0,9213	0,2001
1,65	0,8942	0,5564	5,4	0,9222	0,1969
1,7	0,8922	0,5402	5,5	0,9232	0,1938
1,75	0,8906	0,5252	5,6	0,9241	0,1908
1,8	0,8893	0,5112	5,7	0,9251	0,1879
1,85	0,8882	0,4981	5,8	0,9260	0,1851
1,9	0,8874	0,4858	5,9	0,9269	0,1824
1,95	0,8867	0,4742	6	0,9277	0,1798
2	0,8862	0,4633	6,1	0,9286	0,1772
2,05	0,8859	0,4529	6,2	0,9294	0,1747
2,1	0,8857	0,4431	6,3	0,9302	0,1723
2,15	0,8856	0,4338	6,4	0,9310	0,1700
2,2	0,8856	0,4249	6,5	0,9318	0,1677
2,25	0,8857	0,4165	6,6	0,9325	0,1655
2,3	0,8859	0,4085	6,7	0,9333	0,1633
2,35	0,8862	0,4008	6,8	0,9340	0,1612
2,4	0,8865	0,3935	6,9	0,9347	0,1592

Fonction de répartition de la LOI NORMALE N(0,1)

z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
3	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990
3,1	0,9990	0,9991	0,9991	0,9991	0,9992	0,9992	0,9992	0,9992	0,9993	0,9993
3,2	0,9993	0,9993	0,9994	0,9994	0,9994	0,9994	0,9994	0,9995	0,9995	0,9995
3,3	0,9995	0,9995	0,9995	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9997
3,4	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9998
3,5	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998
3,6	0,9998	0,9998	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999
3,7	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999
3,8	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999

Titre : Probabilités et statistiques en classe de techniciens supérieurs

Auteurs : Claire DUPUIS, Mohamed ATLAGH, André BASTIAN, Bernard GOESEL, Christiane HEUSCH, Bernard KOCH, Dominique PERNOUX, Suzette ROUSSET-BERT.

Mots-clés : Enseignement - Probabilité - Statistique - STS

Date : 1996

Editeur : I.R.E.M. de Strasbourg (S. 167)

ISBN : 2-911446-03-8

Résumé : Cette brochure est un recueil de différents documents sur des parties qui nous ont semblé délicates dans l'enseignement des probabilités en classe de techniciens supérieurs, à savoir :

Opérations sur les variables aléatoires
Variable aléatoire continue
Diverses approximations de la loi binomiale
Processus de Poisson
Caractère universel de la loi normale, théorie des erreurs, théorème de la limite centrée
Echantillonage
Petit herbier de lois
Estimation ponctuelle, estimation par intervalle
Fiabilité

Pour chacun des thèmes abordés, nous avons précisé s'il s'agit d'un document pour l'élève, pour l'enseignant ou pour l'élève avec des compléments pour les enseignants.